INTERNATIONAL CONFERENCE ON TELECOMMUNICATION

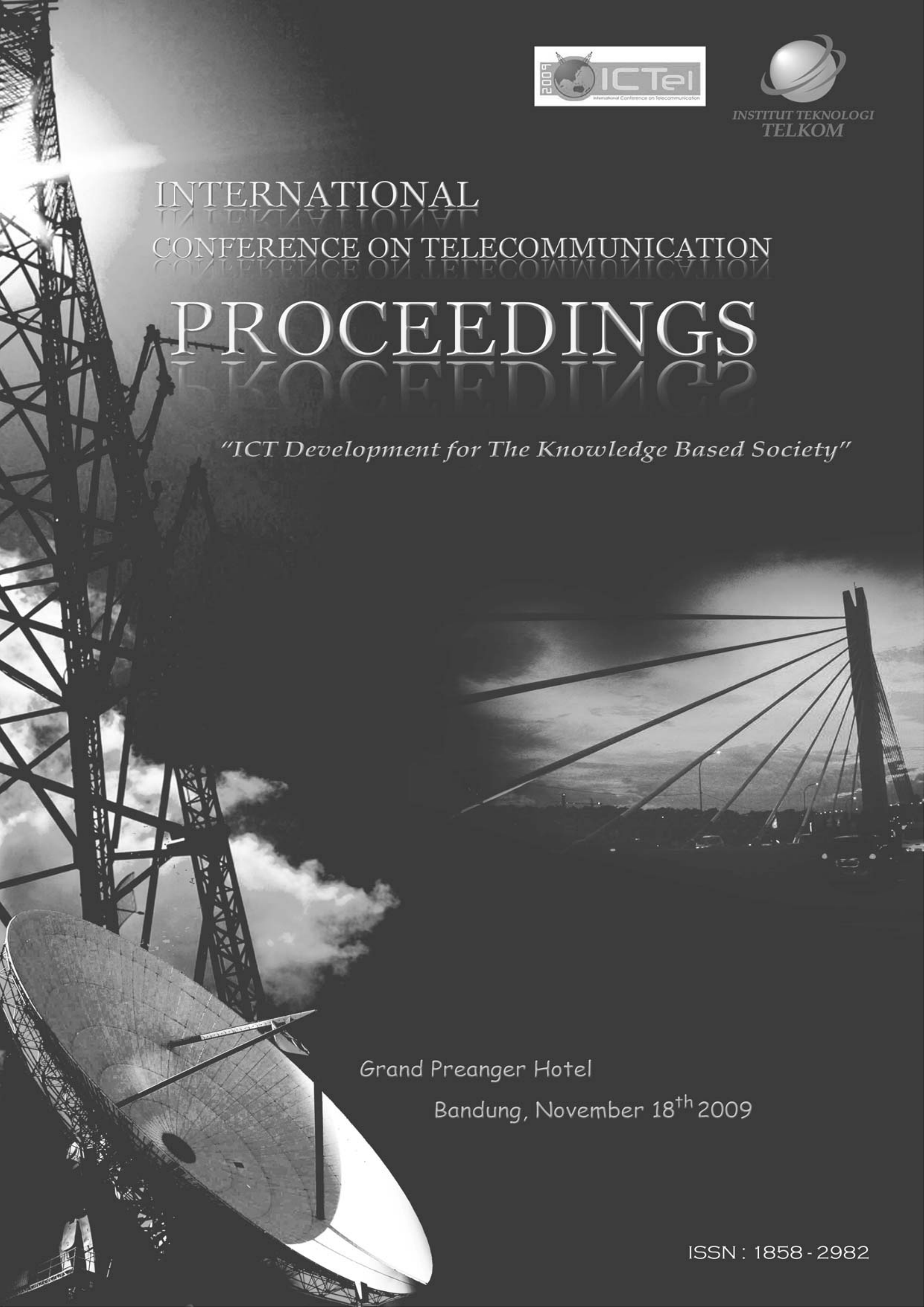
PROCEEDINGS

"ICT Development for The Knowledge Based Society"

Grand Preanger Hotel

Bandung, November 18th 2009

ISSN : 1858 - 2982



INTERNATIONAL CONFERENCE ON TELECOMMUNICATION

PROCEEDINGS

"ICT Development for The Knowledge Based Society"

Grand Preanger Hotel

Bandung, November 18th 2009

ISSN : 1858 - 2982

FOREWORD ICTEL 2009

International Conference on Telecommunication (ICTel) is an annual event hosted by IT Telkom since the year of 2005. Thus today ICTel is its fifth, highlighting on the theme of "ICT Development for the Knowledge Based Society". This theme is chosen as a response from our institution in seeing the current trend where ICT already enters various areas of human life. Those areas include education, health, banking and finance, mining, national defense, and many others. Most of business activities in those areas have been using ICT as their main assisting device. Therefore we can conveniently say that 'ICT for life' has become 'a jargon comes to life', in which lots of ICT applications are everyday phenomena.

Responding to that existing trend, ITTelkom as one educational institution is taking on active participation to link communication among people and communities in ICT development. The vast development in ICT has given more and more rooms to ideas and creative minds. This is exactly where ICTel is expected to facilitate communications and information exchanges, allowing faster and more integrated expansion of ICT knowledge. All ICTel participants coming from diverse background: academics, research and development, industry, and engineering are given widest opportunity to openly discuss the best solution for ICT for the society growing on pillars of knowledge.

Last but not least, since ICT development is not something to be built overnight, and that it needs continuous and sustainable communication, we hope that the discussions taken place in this event can be followed by real efforts to improve the quality of life of society. To all the participants, writers, committees, and all the contributing parties, we extend our sincere gratitude and appreciation. We hope that ICTel 2009 can give meaningful contribution to ICT development for society.

Yours,
Director of Academic Support

Suwandi

WELCOMING SPEECH

ICTel2009 is an annual international conference especially for researchers and academicians in the field of telecommunication to share and publish their works. ICTel2009 consists of plenary session featuring various presenters to expose their research and read condition of telecommunication world.

The theme “ICT Development for the Knowledge Based Society” is chosen because, as a matter of fact, the role of ICT nowadays it expected to be more evolved for the development of sophisticated and advanced society. Moreover, ICT will also enable this society to have better information access to close the economical and social gap. Thus, eventually prosperity shall be achieved by their society.

Mainly, the plenary session of ICTel2009 consists of presentations covering the topic of telecommunication industry, optic, radar, computer, communication system, artificial intelligence, and many others. There are 78 abstracts and papers sent to the committee and only 53 papers are accepted and to be presented. These include 4 international presenters, 16 national presenters, and 33 presenters from the IT Telkom. Participants and presenters of ICTel 2009 come from German, Malaysia, Korea, and Indonesia.

Good luck to all people involved in ICTel 2009. I hope that all of you will enjoy and gain invaluable benefits from all agenda of ICTel 2009.

Bandung, November 2009
Chair of Organizing Committee

Iswahyudi Hidayat

ICTel 2009 COMMITTEE

ORGANIZER

Directorate of Academic Support, Institut Teknologi Telkom Bandung

TECHNICAL PROGRAM COMMITTEE

Josaphat Tetuko Sri Sumantyo (Chiba University)
Joel Joseph Sacro Marciano (IEEE Phillipine Section Chair)
Toni Anwar (TGGS, Thailand)
Marzuki Khalid (University Technology of Malaysia)
Heiko Schroder (Royal Melbourne Institute of Technology)
Andrian Bayu Suksmono (Institut Teknologi Bandung)
Sugihartono (Institut Teknologi Bandung)
Suhartono Tjondronegoro (Institut Teknologi Bandung)
Aly Muayyadi (Institut Teknologi Telkom)
Deni Saepudin (Institut Teknologi Telkom)
Rendy Munadi (Institut Teknologi Telkom)
Rina Puji Astuti (Institut Teknologi Telkom)
Ary M. Barmawi (Institut Teknologi Telkom)
Heroe Wijanto (Institut Teknologi Telkom)
Dharu Arseno (Institut Teknologi Telkom)
Iwan Iwut TA (Institut Teknologi Telkom)

ORGANIZING COMMITTEE

Iswahyudi Hidayat
Yuliant Sibaroni
Erwin Budi Setiawan
Kinkin Kindawati
Leanna Vidya Yovita
Shaufiah
Adiwijaya
Iman Hedi Santoso
Indra Chandra
Deny Saepudin
Koredianto Usman
Rian Febrian Umbara
Florita Diana Sari
Istikmal
Hetti Hidayati
Heri Iman N.
Mediana Mayang Kencana
Ani Yuliani
Warsino
Aris Hartaman
Yanuar Firdaus AW
Chandra Purna Darmawan

LIST OF PARTICIPANT PAPERS

A. Social and Technical aspect of Computer Network

B. Computer Sience and Intelligent System

C. Communication System, Optics, Radar, and Applied Physics

D. Data Mining, Applied Mathematics, Image Processing and ICT

PERFORMANCE COMPARISON OF SCHEDULING ALGORITHM: ROUND ROBIN (RR), WEIGHTED ROUND ROBIN (WRR) AND DEFICIT ROUND ROBIN (DRR) ON WIMAX NETWORK WITH NS-2 SIMULATOR

Iman Hedi Santoso[1], Rian Febrian Umbara[2]

[1]Faculty of Electric and Communication, Institut Teknologi Telkom
[2]Faculty of Science, Institut Teknologi Telkom
[1]ihs@ittelkom.ac.id, [2]rfu@ittelkom.ac.id

Abstract

WIMAX offering us QoS-guaranteed at MAC level. In order to guarantee the QoS, base station (BS) has to allocate bandwidth according to certain algorithm. IEEE 802.16 standard has defined QoS signalling mechanism but not yet for the scheduling algorithm. This scheduling algorithm over WIMAX is opened, giving opportunity to everyone creating inovation in this field. WIMAX network is a high speed network, the simplicity of the scheduling algorithm become important to notice. Therefore, in this research three simple scheduling, Round Robin (RR), Weighted Round Robin (WRR) and Deficit Round Robin (DRR) will be analyzed and compared. The simulation shows that those scheduling can guarantee QoS requirement for every service classes, but the DRR performance slightly below WRR. We should not use RR scheduling, because every service class in WIMAX network have different QoS requirement. Therefore, in WIMAX network, this research suggests using WRR instead of DRR or RR

Index Terms--DRR, NS-2, WIMAX, WRR.

1. Introduction

WIMAX is an IEEE standard for supporting Broadband Wireless Access. The main advantage of WIMAX if compared with another access network is its coverage and its sophisticated in assuring QoS at MAC layer level. Broadband Wireless Access (BWA) system is expected can guarantee QoS for real-time application, such as: video conference, video streaming, and voice over IP (VoIP). Therefore, WIMAX support 5 service type: Unsolicited Grant Services (UGS), real-time Polling Service (rtPS), extended real-time Polling Service (ertPS), non real-time Polling Service (nrtPS), and Best Effort (BE). Each service type has different QoS *requirement*.

IEEE 802.16 standard have defined 2 basic mode: point to multipoint (PMP) and mesh. In mesh mode, each subscriber station (SS) will be able to communicate with other SS or BS. In PMP mode, SS is only allowed made communication through BS. The providers that connected users with internet backbone should use PMP mode because BS will be able to control every SS in order to fulfill the QoS *requirement*.

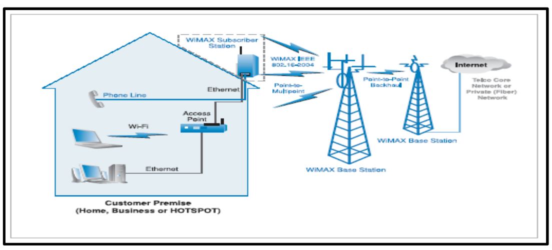


Figure 1. PMP Structure

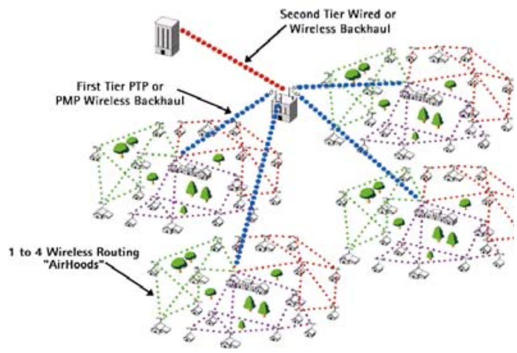


Figure 2. Mesh Structure

One thing making WIMAX become different from other access network is, it is a connection-oriented system. It means that an SS must register to BS before it can start to send or receive data. During this registration process, an SS can negotiate the initial QoS requirements with the BS. These requirements can be changed later, and a new connection may also be established on demand. The QoS requirements may be either per connection based (GPC) or per subscriber station based (GPSS). GPSS mode is more scalable and efficient as compared to GPC. It is also capable in providing lower delay to real-time applications because SS more intelligent in GPSS mode and can react quickly to the needs of real-time applications. This research using GPSS mode as well.

The basic approach in assuring QoS on WiMAX network is that the BS does the scheduling for both the uplink and downlink directions. The scheduling algorithm may be used for:

1. Translating the QoS requirements of SSs into the appropriate number of slots. The algorithm can also account for the bandwidth request size that specifies size of the SS input buffer. When the BS makes a scheduling decision, it informs all SSs about it by using the UL-MAP and DL-MAP messages in the beginning of each frame. These special messages define explicitly slots that are allocated to each SS in both the uplink and downlink directions.
2. Allocating bandwidth, namely time slot, among BS and SS queues.

The scheduling policy, i.e. an algorithm to allocate bandwidth, is not defined in the WiMAX specification, but rather is open for alternative implementations.

There are several articles on the WiMAX QoS scheduling that have presented architectures and scheduling disciplines to guarantee QoS. Several research works [6, 18] propose complex schedulers or even an hierarchy of schedulers, such as Earliest Deadline First (EDF), Deficit Round Robin (DRR) [17], Weighted Fair Queueing (WFQ) [10], and Worst-case Weighted Fair Queueing (W^2FQ) [11]. However, it is a challenging task to use an hierarchy of schedulers because the per-connection QoS requirements should be translated into the scheduler configuration at each level. Furthermore, it is not enough to calculate the scheduler configuration only once when an SS joins or leaves the network. As SSs send data, their request sizes change all the time. As a result, the scheduler at the BS should reassign slots. For exactly these reasons we suggest to use one level with a simple scheduling mechanism that is based conceptually on the round-robin (RR) approach. A simpler solution is better, because there is not much time to do the scheduling decision. For instance, one of the possible configuration values is 400 frames per second [1]. Thus, the BS should make 400 scheduling decision per one second to achieve the accurate and fair resource allocation.

Therefore, for WIMAX network this research propose simple scheduling technique based on round robin, IEEE 802.16: **Round Robin, Weighted Round Robin** [12, 13]. **dan Deficit Round Robin** [17].

The rest of the paper is organized as follow. In Section II, we discuss the overview of IEEE 802.16 standard. Basic theory about scheduling algorithm will be given in section III, especially for WRR and DRR. Furthermore, in Section IV, we explain the implementation of scheduling on NS-2. In Section V we explain the simulation scenario to test the performance of WRR and DRR on WIMAX network. The simulation result and discussion about it, will be descibed in this section. Finally,in Section VI, we give the conclusion and future work of this research.

2. IEEE 802.16 – WIMAX Standard

The IEEE 802.16 standard is released for the first time on April 2002, and afterward after some revision on June 2004 the released standard is named IEEE 802.16-2004. The next edition of this standard is IEEE 802.16-e that released on Desember 2005.

Physical layer of this standard operate at 10-66 GHz (IEEE 802.16) and 2-11 GHz (IEEE 802.16a-e) with data rate 32 – 130 Mbps depend on the frequency channel and modulation technique.

The making use of modulation technique depend on the distance between SS and BS, see fig. 4. In short distance, an SS receive large enough signal power so it may use modulation technique that support high bit rate (but less robust) like 64-QAM. If the distance grow farther and farther away, the receiving signal become weaker, so then it should use more robust modulation technique although the resulting data rate is lower.

Table 1.Modulation Technique

Modulation Technique	Useful bits in OFDM Symbol	Raw bitrate, [Mbps]	Ratio
BPSK 1/2	96	2.82	9
QPSK 1/2	192	5.65	4.5
QPSK 3/4	288	8.48	3
16-QAM 1/2	384	11.29	2.25
16-QAM 3/4	576	16.94	1.5
64-QAM 2/3	768	22.59	1.125
64-QAM 3/4	864	25.41	1

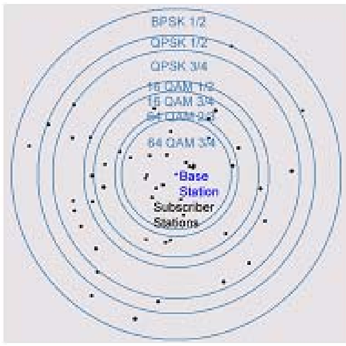


Figure 3. Relationship Between Distance and Modulation

IEEE 802.16 architecture consist of two kinds of stations: Subscriber Station (SS) and Base Station (BS). The Base Station controls all of the communication happened in network. There is 2 communication direction between BS and SS: uplink (from SS to BS) and downlink (from BS to SS).

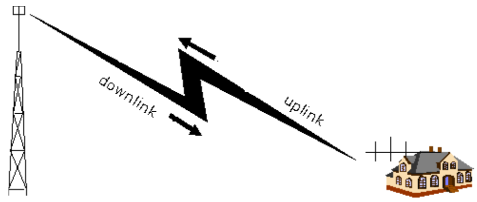


Figure 4. Communication Direction

If we use TDM mode for transmission, a frame consist of 2 section, *uplink subframe* and *downlink subframe* (Fig. 6). The duration of those subframes are defined dynamically by the BS. Each subframe consists of a number of time slots. SS and BS must be synchronized and the data transmitted based on predetermined time slots.

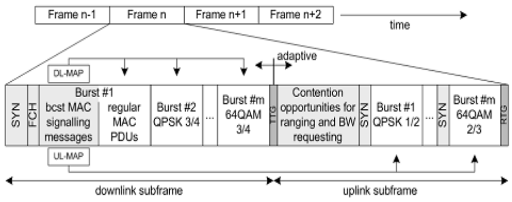


Figure 5. TDD Frame Structure

In introduction have been said that, IEEE 802.16 support multiple communication services (data, audio, and video) with different QoS requirement. Figure 7 shows some parameter like: reliability, delay, jitter , dan bandwidth requirement, where all of those parameters are the QoS *requirement* for sundry telecommunication application.

Table 2. QoS R*equirement*
for Telecommunication Applications [13]

Application	Reliability	Delay	Jitter	Bandwidth
E-mail	High	Low	Low	Low
File Transfer	High	Low	Low	Medium
Web Access	High	Medium	Medium	Medium
Remote Login	High	Medium	High	Low
Audio on Demand	Low	Low	High	Medium
Video on Demand	Low	Low	High	High
Telephony	Low	High	High	Low
Videoconferencing	Low	High	High	High

MAC layer in IEEE 802.16 have defined QoS signalling function and mechanism to control data transmission between BS and SS. On the downlink (from BS to SS), the transmission is relatively simple because the BS is the only one that transmits during the downlink subframe. The data packets are broadcasted to all SSs and an SS only picks up the packets destined to it. One of the modes of uplink arbitration (from SS to BS) uses a TDM MAC.The BS determines the number of time slots that each SS will be allowed to transmit in an uplink subframe. This information is broadcasted by the BS through the uplink map message (UL-MAP) at the beginning of each frame. UL-MAP contains information element (IE) which include the transmission opportunities, i.e. the time slots in which the SS can transmit during the uplink subframe. After receiving theUL-MAP,each SS will transmit data in the predefined time slots as indicated in IE.

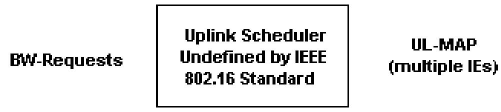


Figure 6 QoS Signalling

The BS uplink scheduling module determines the IEs using bandwidth request PDU (BW-request) sent from SSs to BS (see fig. 8). In IEEE 802.16 standard, there are two modes of transmitting the BW-Request: **contention mode** and **contention-free mode (polling)**. In contention mode, SSs send BW-Request during the contention period. Contention is resolved using back-off resolution. In contention-free mode, BS polls each SS and SSs reply by sending BW-request. Due to the predictable signaling delay of the polling scheme, contention-free mode is suitable for real time applications. IEEE 802.16 defines the required QoS signaling mechanisms described above such as BW-Request and UL-MAP, but it does not define the Uplink Scheduler, i.e. the mechanism that determines the IEs in the UL-MAP.

Recent IEEE 802.16 standard defines five types of service flows with different QoS requirements and corresponding scheduler policy:

1. **Unsolicited grant service (UGS)**, this service supports constant bit-rate (CBR) or CBR-like flonws such as Voice over IP. These applications require constant bandwidth allocation, therefore BW-request not required and scheduler allocates a fixed number of time slots in each time frame.
2. **Extended real-time Polling Service (ertPS)**, this service combine the superiority of UGS and rtPS and is made to support VoIP in silent-suppression mode. These applications require BW-request with contention-free mode.
3. **Real-time polling service (rtPS)**, this service is for real-time VBR-like flows such as MPEG video. These applications have specific bandwidth requirements as well as a deadline (maximum delay). Late packets that miss the deadline will be useless. So that, BW-request processes are needed in contention-free mode only. The current queue size that represents the current bandwidth demand is included in the BW-Request.
4. **Non-real-time polling service (nrtPS)**, this service is for non-real-time flows which require better than best effort service, e.g.

bandwidth intensive file transfer. These applications are time-insensitive and require minimum bandwidth allocation. The BW-request activities are used either in contention-free mode or contention mode. Current queue size is included in BW-request.

5. **Best effort service (BE)**, this service is for best effort traffic such as HTTP. There is no QoS guarantee. The applications in this service flow receive the available bandwidth after the bandwidth is allocated to the previous four service flows. BW-request uses only contention mode. Current queue size is included in BW-request.

3. Scheduling Algorithm

The next generation of the Internet supports two types of applications: best-effort and guaranteed-service applications. The best-effort applications, which are already common to the Internet, are content to accept whatever performance the network gives them. For example although a file transfer application would prefer to encounter zero end-to-end delay and infinite bandwidth, it can still adapt to the available network resources. These applications are called best-effort because the network promises to deliver their data without any guarantees on performance bounds.

Beside these applications, the Internet is expected to carry traffic from applications that require performance bounds in the future. For example, an application that contains voice as a 64 Kbps data stream will be no longer usable if the network provides a bandwidth less than 64kbps.

The performance received by a connection depends principally on the scheduling mechanism. These mechanisms are carried out by the switching nodes located along the path between the source and the destination of a connection. Scheduling mechanisms are implemented at output interfaces of switches or routers. At each output queue of an interface a scheduling mechanism is used to choose which packet to transmit to the outgoing link. The scheduler can allocate different queuing delays and different bandwidths to different connections. This is done by the scheduler's choice of service order and by serving a certain number of packets from a particular connection. The scheduler can also allocate different loss rates to connections by assigning a certain amount of buffer space to them. Furthermore, it is able to allocate resources, which are desired properties in the networks fairly among best-effort connections. Thus, the next generation of Internet needs scheduling disciplines in order to support:

- per-connection delay, bandwidth, and loss bounds needed for guaranteed-service applications.
- fair resource allocation needed for best-effort applications.

In the design of scheduling schemes, different trade-offs can be considered in terms of the following five requirements: complexity, fairness, isolation/protection, efficiency, and performancy. Depending on the specific situation, some of these requirements may be more important than others and the decision for the best choice is made given the particular situation.

Next, will be described briefly about some kinds of scheduling algorithm [13]:

A. First In First Out (FIFO)

First-in, first-out (FIFO) queuing is the most basic queue scheduling discipline. In FIFO queuing, all packets are treated equally by placing them into a single queue, and then servicing them in the same order that they were placed into the queue. FIFO queuing is also referred to as First-come, first-served (FCFS) queuing.

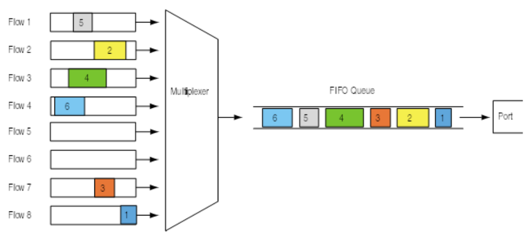


Figure 7. FIFO

B. Priority Queueing (PQ)

This scheduling algorithm is the basis for a class of queue scheduling algorithms that are designed to provide a relatively simple method of supporting differentiated service classes. In classic PQ, packets are first classified by the system and then placed into different priority queues. Packets are scheduled from the head of a given queue only if all queues of higher priority are empty. Within each of the priority queues, packets are scheduled in FIFO order.

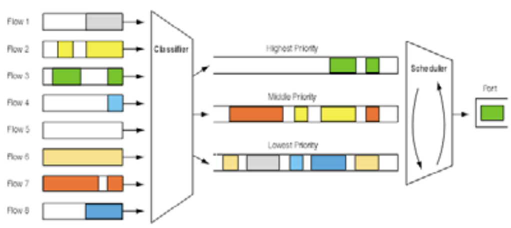


Figure 8. Priority Queueing (PQ)

C. Fair Queueing (FQ)

FQ is the foundation for a class of queue scheduling disciplines that are designed to ensure that each flow has fair access to network resources and to prevent a bursty flow from consuming more than its fair share of output bandwidth. In FQ, packets are first classified into flows by the system and then assigned to a queue that is specifically dedicated to that flow. Queues are then serviced one packet at a time in round-robin order. Empty queues are skipped. FQ is also referred to as per-flow or flow-based queuing.

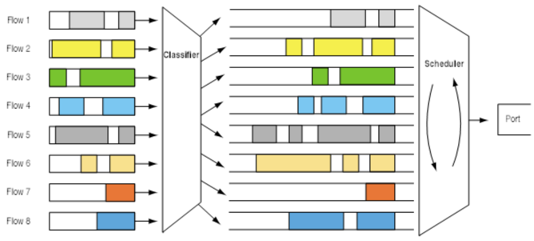


Figure 9. Fair Queueing (FQ)

D. Weighted Fair Queueing (WFQ)

WFQ is the basis for a class of queue scheduling disciplines that are designed to address limitations of the FQ. WFQ supports flows with different bandwidth requirements by giving each queue a weight that assigns it a different percentage of output port bandwidth. WFQ supports the fair distribution of bandwidth for variable-length packets by approximating a generalized processor sharing (GPS) system. GPS is a theoretical scheduler that cannot be implemented.

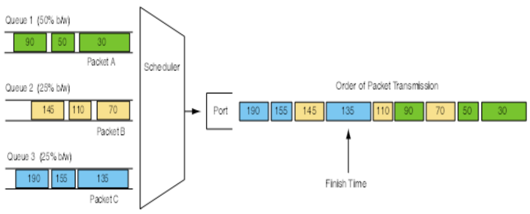


Figure 10. Weighted Fair Queueing

E. Round Robin (RR) and Weighted Round Robin (WRR)

WRR is the foundation for a class of queue scheduling disciplines that are designed to address the limitations of the FQ and PQ models. RR scheduling is just a special case of WRR where RR is a WRR scheduling with similar weight

WRR address the limitations of the FQ model by supporting flows with significantly different bandwidth requirements. With WRR queuing, each queue can be assigned a different percentage of the output port's bandwidth. WRR addresses the limitations of the strict PQ model by ensuring that lower-priority queues are not denied access to buffer space and output port bandwidth.

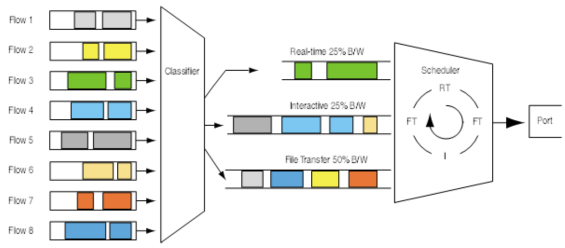


Figure 11. Weighted Round Robin

F. Deficit Round Robin (DRR)

DRR is the basis for a class of queue scheduling disciplines that are designed to address the limitations of the WRR and WFQ models. DRR addresses the limitations of the WRR model by accurately supporting the weighted fair distribution of bandwidth when servicing queues that contaion variable-length packets. DRR addresses the limitations of the WFQ model by defining a scheduling discipline that has lower computational complexity and that can be implemented in hardware. This allows DRR to support the arbitration of output port bandwidth on high-speed interfaces in both the core and a the edges of the network.

In DRR, each queue is configured with a number of parameters:

- **A weight** that defines the percentage of the output port bandwidth allocated to the queue
- **A Deficit Counter** that specifies the total number of bytes that the queue is permitted to transmit each time that it is visited by the scheduler. The Deficit Counter allows a queue that was not permitted to transmit in the previous round because the packet at the head of the queue was larger than the value of the Deficit Counter to save transmission "credits" and use them during the next service round.
- **A quantum** of service that is proportional to the weight of the queue and is expressed in terms of bytes. The Deficit Counter for a queue is incremented by the quantum each time that the queue is visited by the scheduler. If : **quantum[i] = 2*quantum[x]**, then queue *i* will receive twice the bandwidth as queue *x* when both queues are active.

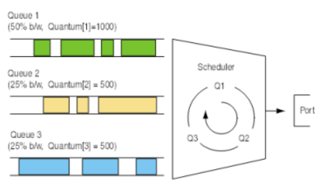


Figure 12. Deficit Round Robin (DRR)

The figure below describes how DRR scheduling serve packets in each queue. From that figure we can see the superiority of DRR in serving queues with variable packet-size.

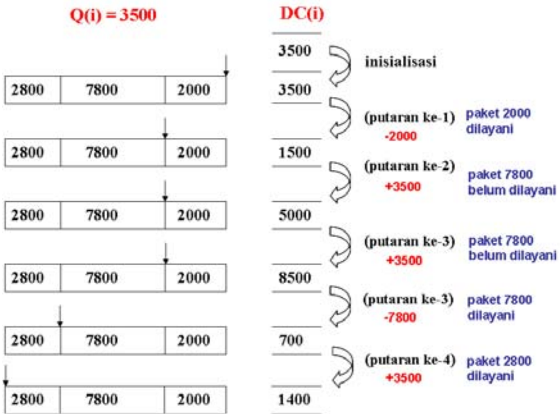


Figure 13. DRR procedure

Here it is three formulations for DRR [16]:

- **Formulation 1**
  For all *i*, the following invariant holds for every execution of the DRR algorithm:
  $$0 < DC(i) < Max.$$
  Max is maximum packet size in each queue.

- **Formulation 2**
  For an interval (t_1, t_2) in any execution of the DRR service discipline prevail:
  $$FM(t_1, t_2) \leq 2\ Max + Q\ ,\ \text{where } Q = Min[\ Q(i)\ ]$$

- **Formulation 3**
  The work for DRR is O(1), if for all *i*:
  $$Q(i) \geq Max$$

4. Scheduling Algorithm Implementation on NS-2 Simulator

The most popular network simulator used by the academia and industry is the network simulator 2 (ns-2), which has become the de facto standard for the simulation of packet-switched networks. Specifically, more and more published network studies and investigations take ns-2 as their evaluation tool to verify their work. Although there is another force that investigates the IEEE 802.16-based simulator, this simulator is not for public. The ns-2 is roughly composed of various traffic models, transport-layer protocols, network-layer protocols, medium access control (MAC) layer protocols, and more. These components enable ns-2 to simulate different types of networks and their topologies. Researchers can take benefit from these preliminary tests on their investigation and find out the drawbacks of their new design in the early stage efficiently.

This research uses the WiMAX module developed by Jenhui Chen[3,4] that focused on MAC protocol development which is inheriting from the original MAC class in ns-2. This module is based on IEEE802.16 point-to-multipoint (PMP) mode, which means that one BS can serve multiple subscriber stations (SSs) concurrently. For the physical (PHY) layer, this module uses the orthogonal frequency-division multiple access (OFDMA) scheme. OFDMA PHY become interesting technique for wireless applications due to its high date rate transmission capability and its robustness to multipath delay spread.

In NS-2, scheduling algorithms are done by a function, ie: **Scheduler()**. Scheduler() function has several task, namely:

1. Allocate bandwidth fairly among all of service types that resident in an SS queues.
2. Guaranteing QoS for all service types in an SS.

For the sake of QoS-guarantee, all service type in an SS will be granted a part of bandwidth system, at least as much as that requested by an SS.

To begin with, in the DL, we associate one percentage parameter with each classification as q5, q4, q3, q2, q1, which is corresponding to the UGS, rtPS, ertPS, nrtPS, and BE, respectively. In the first round, the expected serving quantity of each classification is calculated as $B_i = \min(R_i, B_{total} * q_i)$, $i \in \{1..5\}$ and $\sum qi = 1$, where R_i represents the total amount of requested type services and B_{total} represents the total available bandwidth of the system. The parameters {q5,q4,q3,q2,q1} are variables and can be regulated by any simulation need.

In the second round, the Scheduler() will serve the remaining, unserved services in priority order. If all remaining services in priority i are served, the Scheduler() will serve the next priority i+1 and so on. This process will be repeated until whole available bandwidth are exhausted or remaining required services are served. The adopted strategy is used to guarantee that lower priority traffic can still obtain a minimum bandwidth for transmission if the traffic load is extremely heavy.

5. Simulation and Analysis of Scheduling Algorithm on WIMAX Network

In this section will be done WIMAX network simulation using NS-2. In fact, NS-2 haven't support the WIMAX yet, therefore a number of party make effort to develop WIMAX module for NS-2, one of them available in reference [3], [4]. The simulation environment, shown in fig. 16, is set of one BS serving a number of SSs. Modulation technique used is 64-QAM 3/4. The amount of SSs involve in this simulation depend on simulation scenario.

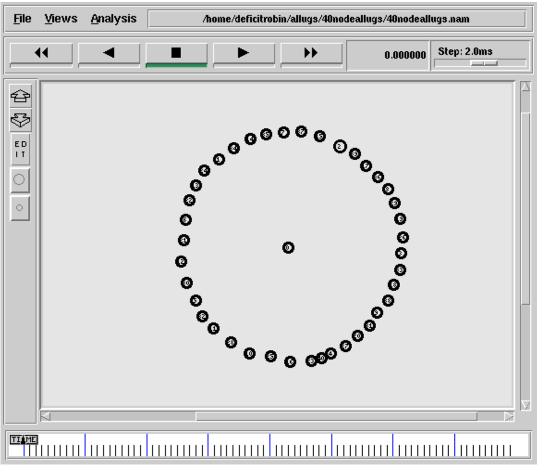


Figure 14. Tampilan Grafis pada Simulator NS-2

There are three kinds of service flows: UGS, nrtPS, and BE which are all generated from the traffic generating agent in both the SSs and the BS. The Internet traffic is treated as the DL traffic to the SSs; on the contrary, the UL traffic is the traffic from the SSs to the Internet. UGS connection has bit rate 448 kbps and packet size 210 byte. ertPS connection has 448 kbps, packet size 210 byte in normal situation and 128 byte while silent. rtPS traffic has 448 kbps bit rate and the packet size follows the uniform distribution model by setting Uniform(200, 980) and time interval Uniform(-0.5, 0.5), i.e. each connection occupies for about 1.5 Mbps. The traffic priority order is: UGS, ertPS, and rtPS.

In this simulation several matter will be analyzed, ie:

1. WIMAX network characteristics.
2. Performance comparison between WRR and DRR.
3. Determination of weight to get good performance of scheduling algorithm.

The used WIMAX parameters are shown in the table below:

Table 3. WIMAX parameters [3], [4]

Parameter	Value
Rasio DL/UL	3:2
OFDMA simbol per frame	48
OFDMA symbol time	100.84 μs
OFDMA frame length	5 ms
Subchannel	30
Bandwidth Request opp	12 OFDMA symbols
Initial ranging CID	0
Basic CIDs	1-1000
Primary CIDs	1001-2000
Transport/secondary Mgt. CIDs	2001-65278
Broadcast CID	65535
SFID range	1-4294967295
TTG	200 μs
RTG	200 μs

A. WIMAX Network Characteristics

In this section will be observed the behaviour of three service flow: UGS, ertPS, and rtPS, related to throughput and delay in case of increasing SS/node number. In this scenario, each node contain one service type. The simulation will be done using scheduling algorithm Deficit Round Robin.

Fig. 17 to 20 show that, based on throughput and queue delay observation, UGS service type has most stabil performance, folowwed by ertPS and rtPS. This prove that, in WIMAX network, UGS traffic get first priority for servicing. For ertPS, average throughput nearly same with UGS, but the queue delays have more variation. While for rtPS the graphic shown, if node number larger than 40, the throughput become stagnant, tend to fixed in 2.500.000 byte/second.

About delay, each node has the same queue delay, all about 3 x 10^{-5} second. But fig. 20 shown that UGS traffic has smallest deviation standard queue delay, followed by ertPS and rtPS.

All of this data proving that UGS is most stable service flow and has first priority order.

While for transmission delay, all of service types have the same transmission delay, about 8 x 10^{-7} second. It is happened because, shown in fig. 16, all nodes have the same distance from BS.

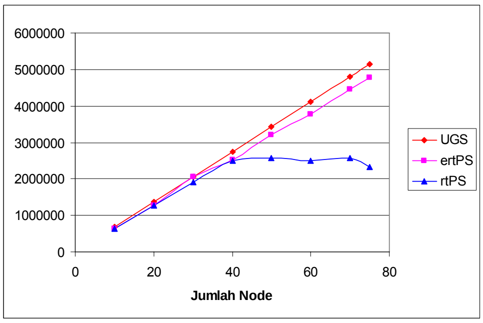


Figure 15. The Comparison of Average Throughput

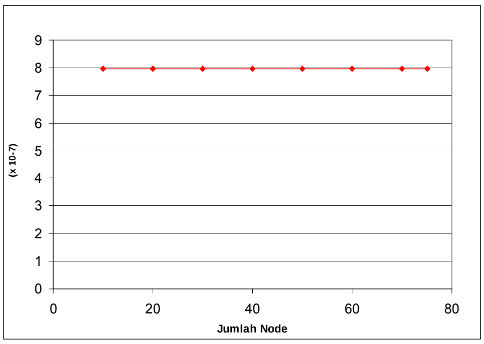


Figure 16. Transmission Delay for UGS, ertPS, and rtPS

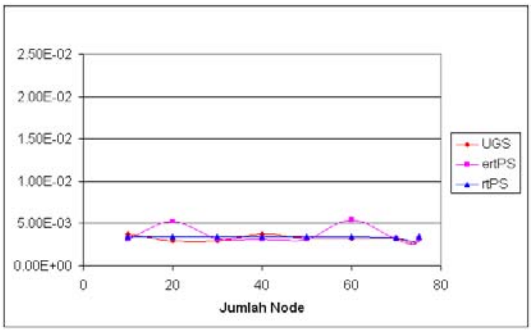


Figure 17. The Comparison of Average Queue Delay

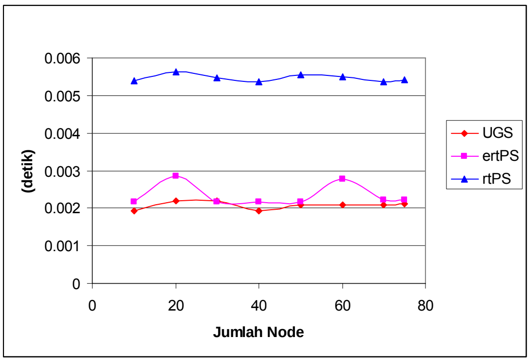


Figure 18. The Comparison of Deviation Standard of Queue Delay

The rtPS traffic, fig 20 shown, has largest deviation standard for queue delay. It can happen because rtPS has variable packet size and have third priority order.

B. Performance analysis of RR, WRR, and DRR

In this section will be using simulator to know the performance of Round Robin (RR), Weighted Round Robin (WRR), dan Deficit Round Robin (DRR) scheduling on WIMAX network, whether those three scheduling algorithm can allocate bandwidth fairly among SSs and the connections. The indicator of the goodness of a bandwidth allocation is the assuredness QoS of each service flow.

Scheduling algorithm will be tested through observing average throughput to increased SS/node number. In this scenario, each node serving three service type: UGS, ertPS, and rtPS. The simulations are done with three weighted scheme.

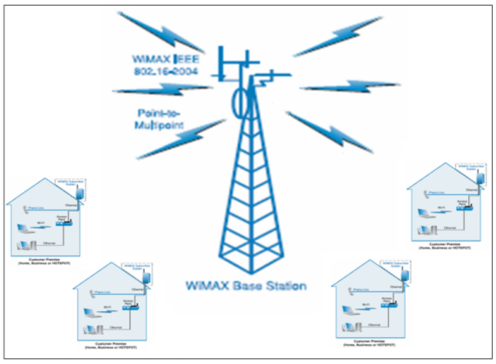


Figure 19. WIMAX Network Configuration with Multiple SS

The graphic below shown simulation result that related average throughput to node number, for WRR and DRR.

B.1 Round Robin (RR) and Weighted Round Robin (WRR)

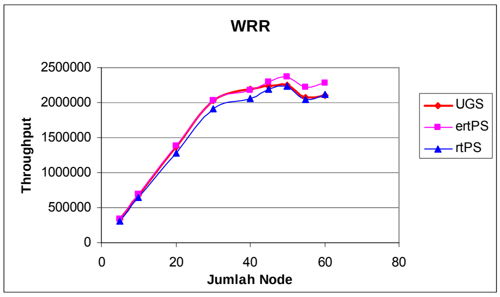


Figure 20. WRR with Weighted Scheme: 30-30-30

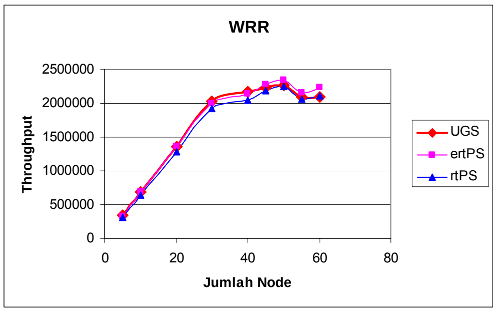


Figure 21. WRR with Weighted Scheme: 20-20-50

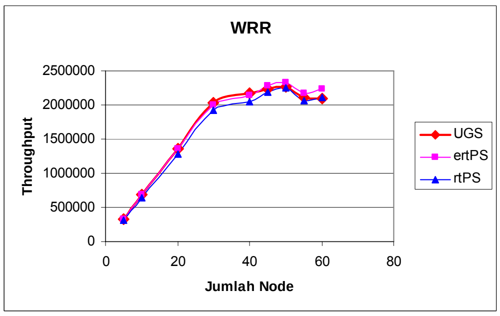


Figure 24. WRR with Weighted Scheme: 30-20-40

Average throughput for each weighted scheme can be seen in Table 2. It is obvious that through weighted control, certain average throughput can be reached. By giving larger weight for certain service type will be get larger average throughput for that service.

Table 4. Aggregat Average Throughput for RR and WRR

Bobot	UGS	ERTPS	RTPS
30-30-30	15262756	15812109	14807177
20-20-50	15278197	15524050	14840915
30-20-40	15282473	15524272	14832406

B.2 Deficit Round Robin (DRR)

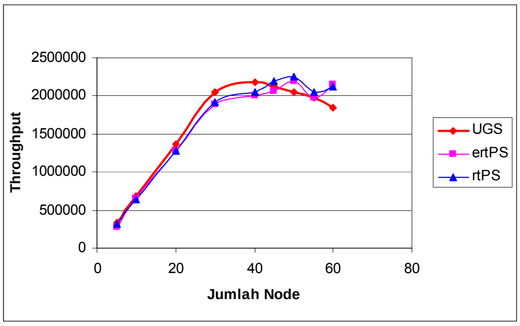


Figure 25. DRR with Quantum Scheme: 250-250-980

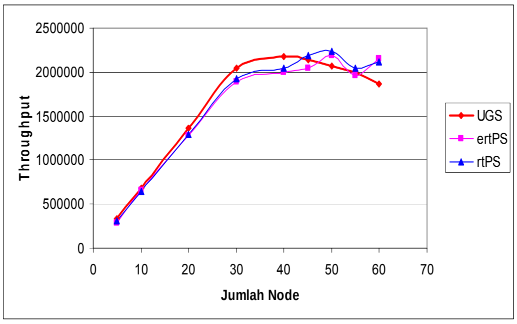


Figure 25. DRR with quantum scheme: 500-250-980

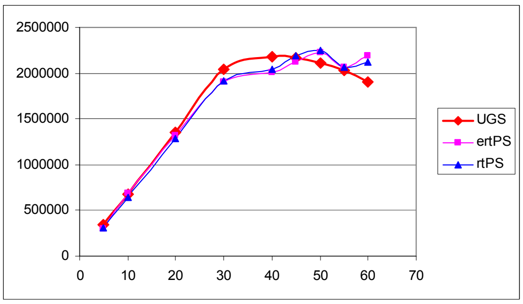


Figure 26. DRR with quantum scheme: 980-980-980

Table 5 Aggregat Average Throughput for DRR

Quantum			UGS	ertPS	rtPS
UGS	ertPS	rtPS			
980	980	980	14800308	14798198	14818691
250	250	980	14599392	14488870	14830510
500	250	980	14671887	14462965	14802097

Table 3 show that the effect of quantum to throughput nearly equal to the effect of WRR to that. But quantum quantity not represent queue prosentage to bandwidth system. Quantum quantity determined from maximum packet size in a queue. From the theory previously presented, known that, for each queue: **Q ≥ maximum packet-size**. Packet size for UGS is 210 byte, maximum packet size for ertPS and rtPS is 210 and 980 respectively. Therefore, quantum scheme give quantum quantity as big as 980 for all services. On row 2 and 3 of those table, we set the quantum Q < 980 for UGS and ertPS, we get smaller throughput for those two services

6. Conclusion

By observing throughput and delay, it is proven that, scheduling algorithm WRR and DRR may be considered to be used on WIMAX network. However, WRR scheduling more appropriate than DRR, it is shown by its larger aggregate throughput of WRR than DRR. Meanwhile, it's better not to use RR in WIMAX network because RR scheduling cannot allocate resource/bandwidth flexibly.

Based-on WIMAX characteristics observation, ertPS service flow is proven to be most flexible in adapting the circumstance. Delay and throughput of ertPS seem are not easily drop by scheduling scheme alteration and the increasing number of node and connection.

For future work, it is important to compare performance RR, WRR, and DRR with another scheduling algorithm so that the superiority and shortcoming can be seen more clearly . The limitation problem, namely error free channel, have to be dissapeared in order to get simulation result that resemble the reality.

If the limitation problem, viz error-free channel, have been disapeared, then:

1. Scheduling algorithm that take into account multiuser diversity factor, i.e. *Opportunistic Deficit Round Robin* and *Opportunistic Weighted Round Robin*, could be investigated.
2. Two principle that enable high performance in OFDMA: multiuser diversity dan adaptive modulation, will be an interesting object to be observed and analyzed.

References:

[1] IEEE, "IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems", IEEE Standard, October 2004.

[2] IEEE 802.16e-2005. "IEEE Standard for Local and Metropolitan Area Networks - Part 16: Air Interface for Fixed Broadband Wireless Access Systems"—Approved 7 December 2005

[3] **Chen, J. (2007),** ***The Design and Implementation of WiMAX Module for ns-2 Simulator*****.** http://ndsl.csie.cgu.edu.tw/WIMAX for NS-2.htm

[4] Chen, Jenhui, "The Design and Implementation of WIMAX Module for NS-2 Simulator", Department of Computer Science and Information Engineering, Chang Gung University, Taiwan.

[5] Borin, J. Freitaq, "WIMAX module for the NS-2 Simulator", http://www.lrc.ic.unicamp.br/wimax_ns2/, 2008.

[6] Wongthavarawat, K., and Ganz, A. (2003), "Packet *scheduling* for QoS suport in IEEE 802.16 broadband wireless access systems", *International Journal of Communication Systems,* vol.16, issue 1, February.

[7] Claudio Cicconetti, Luciano Lenzini, and Enzo Mingozzi,"Quality of Service Support in IEEE 802.16 Networks", *IEEE Network*, April 2006.

[8] Hemant Kumar, Abhijeet Bhorkar, Vishal Sharma,"An Opportunistic Uplink *Scheduling* Scheme to Achieve Bandwidth Fairness and Delay For Multiclass Traffic in WIMAX (IEEE 802.16) Broadband Wireless Network", Department of Electrical Engineering, IIT Bombay, Mumbay, India.

[9] Keshav, S. (2003), An Engineering Approach to Computer Networking. Pearson Education, Sixth edition, 2003.

[10] A.K. Parekh and R. G. Gallager. A generalized processor sharing approach to flow control in integrated services networks: The single node case. *IEEE/ACM Transactions on Networking,* 1(3):344-357, June 1993.

[11] J. C. R. Bennett and H. Zhang. W^2FQ: Worst-cse fair weighted fair queueing. In *INFOCOM*, pages 120-128, March 1996.

[12] E. L. Hahne, R. G. Gallager, "Round Robin *Scheduling* for Fair Flow Control in Data Communication Networks," International Conference on Communications, pp. 103–107, June 1986.

[13] Semeria, Chuck. Supporting differentiated service classes: Queue *scheduling* disciplines. Juniper Networks, 2001.

[14] Tanenbaum, A. S. (2003), *Computer Networks*, Prentice-Hall India, Fourth edition.

[15] Demers, A., Keshav, S. , Shenker, S.(1989), *Analysis and simulation of a fair queueing algorithm*. *Xerox PARC*.

[16] Sayenko, A., Alanen, O., Karhula, J., Hamalainen, T. (2006), *Ensuring the QoS Requirement in 802.16 Scheduling*. ACM 1-59593-477.

[17] Shreedhar, M., Varghese, G. (1996), Efficient Fair Queueing Using Deficit Round Robin, *IEEE/ACM Transaction on Networking*, Vol. 4.

[18] Chen, J., Jiao, W., and Wang, H., A Service Flow Management Strategy for IEEE 802.16 Broadband Wireless Access in TDD Mode. In *IEEE International Conference on Communications*, volume 5, pages 3422-3426, May 2005.

ANALYSIS PERFORMANCE CONGESTION CONTROL ALGORITHM ON MOBILE ADHOC NETWORK (MANET)

Istikmal[1], Gelar Budiman[2], Yudha Purwanto[3],Iman Hedi S [4]

[1,2,3,4] Institut teknologi Telkom
[1]itm@ittelkom.ac.id, [2]glb@telkom.co.id, [3]ydp@stttelkom.ac.id, [4]ihs@stttelkom.ac.id

Abstract

Adhoc network are multihop wireless network contain a collection of wireless mobile nodes dynamically forming a temporary network without the use of any existing network infrastructure. TCP (Transmission Control Protocol), a widely used end-to-end reliable transport protocol designed for wired network relies on packet loss as an indication of network congestion then triggers congestion control algorithm and decrease the throughput. Due to mobility, packet loss in adhoc network occur as result of channel error, route change, path break which disconnection and reconnection . In this paper we compare and evaluated the Congestion control algorithm in different environments. We analysis and evaluated congestion control algorithm performance in TCP over wired and then we examine congestion control algorithm in adhoc network. We use two different congestion control algorithm. We use standard congestion control algorithm and congestion control that has been modified for adhoc network and proposed with multi metric concept based on end-to-end measurement. TCP with standard congestion control algorithm performance degrades significantly in wireless adhoc due to the misinterpretation and an unable to detect appropriate condition in adhoc network. TCP with congestion control that has been modified has better performance over standard TCP with higher throughput and packet delivery ratio.

Key word : Adhoc, congestion control algorithm. TCP.

1. Introduction

Adhoc network is type of network are useful in any situation where temporary network connectivity is needed, these network are mutihop wireless network consisting of large number of radio equipped nodes that may be as simple as autonomous (mobile or stationary) sensors to laptops mounted on vehicles or carried by people.

TCP (Transmission Control Protocol) was designed to provide reliable end-to-end delivery of data over unreliable networks. TCP is widely used by many Internet services. Thus, even if the network infrastructure may change in the future Internet, TCP and its applications would be likely to be continuously used. In theory, TCP should be independent of the technology of the underlying infrastructure. In contrast, the end-to-end approach is easy to implement and deploy, requires no network support, and provides the flexibility for backward compatibility.

In this paper we focus on congestion control mechanisms of TCP. An essence of the congestion avoidance mechanism of TCP is to dynamically control the window size according to the congestion level of the network. In Figure 1 show the congestion control mechanisms model. It consists of N sender with *Xi* data, so the total data at the network is $\sum_{i=0}^{n} x_i$, if the amount of the total data at the network more then capacity at the network *Xgoal*. The network will use feedback ACK to inform the sender for decrease the throughput

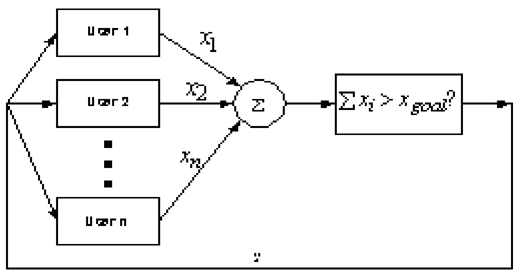


Figure 1.Congestion Control Mechanisms Model.

TCP performance degrades significantly in mobile ad hoc networks [12, 14]. Since TCP assumes that packet losses occur because of congestion in Ad hoc network, it will invoke congestion control mechanisms for packet losses caused by route failures, channel error, and path break resulting in the reduction in throughput.

Holland and Vaidya [12] showed that turning off replying from caches improves TCP performance for a network with a single TCP connection. But this approach will degrade TCP performance when multiple traffic sources exist because of increased routing overhead. TCP (Transmission Control Protocol) was designed to provide reliable end-to-end delivery of data over unreliable networks. In theory, TCP should be independent of the technology of the underlying infrastructure . In contrast, the end-to-end approach is easy to implement and deploy, requires no network support, and provides the flexibility for backward compatibility. In this paper we first analysis and evaluated congestion control algorithm performance in TCP over wired and then we examine congestion control algorithm in adhoc network. We use two different congestion control algorithm. We use standard congestion control algorithm and congestion control use an end-to-end approach to improve TCP performance in mobile ad hoc networks, End-to-end measurements are used to detect congestion, disconnection, route change, and channel error, and each detection result triggers corresponding control actions. This technique is one of the promising approaches to improves TCP performance in MANET

2. Congestion Control Algorithm and challenges in MANET

2.1 Standard Congestion Control Algorithm

In TCP Reno, the window size is cyclically changed in a typical situation. The window size continues to be increased until packet loss occurs. TCP Reno has two phases in increasing its window size; slow start phase and congestion avoidance phase. When an ACK packet is received by TCP at the sender side at time $t+t_A$ [sec], the current window size $cwnd(t+t_A)$ is updated from $cwnd(t)$ as follows [13] :

$$cwnd(t+t_A) = \begin{cases} \text{slow start phase :} \\ cwnd(t)+1, \text{ if } cwnd(t) < ssth(t); \\ \text{congestion avoidance phase :} \\ cwnd(t) + 1/(cwnd(t)), \text{ if } cwnd(t) \geq ssth(t); \end{cases}$$

where $ssth(t)$ [packets] is a threshold value at which TCP changes its phase from slow start phase to congestion avoidance phase. When packet loss is detected by retransmission timeout expiration, $cwnd(t)$ and $ssth(t)$ are updated as [13] :

$$cwnd(t) = 1; \qquad ssth(t) = cwnd(t) / 2$$

On the other hand, when TCP detects packet loss by a fast retransmit algorithm [3], it changes $cwnd(t)$ and $ssth(t)$ as :

$$ssth(t) = cwnd(t) / 2; \quad cwnd(t) = ssth(t)$$

TCP Reno then enters a fast recovery phase [13] if the packet loss is found by the fast retransmit algorithm. In this phase, the window size is increased by one packet when a duplicate ACK packet is received. On the other hand, $cwnd(t)$ is restored to $ssth(t)$ when the non–duplicate ACK packet corresponding to the retransmitted packet is received.

TCP Vegas controls its window size by observing RTTs (Round Trip Time) of packets that the sender host has sent before. If observed RTTs become large, TCP Vegas recognizes that the network begins to be congested, and throttles the window size. If RTTs become small, on the other hand, the sender host of TCP Vegas determines that the network is relieved from the congestion, and increases the window size again. Hence, the window size in an ideal situation is expected to be converged to an appropriate value. More specifically, in congestion avoidance phase, the window size is updated as [13]:

$$cwnd(t+t_A) = \begin{cases} cwnd(t) + 1, \text{ if } diff < \alpha / (base\_rtt) \\ cwnd(t), \text{if } \alpha / (base\_rtt) \leq diff \leq \beta / (base\_rtt) \\ cwnd(t) - 1, \text{ if } \beta / (base\_rtt) < diff \\ diff = (cwnd(t) / base\_rtt) - (cwnd(t) / rtt) \end{cases}$$

where rtt [sec] is an observed round trip time, $base\ rtt$ [sec] is the smallest value of observed RTTs, and α and β are some constant values. TCP Vegas has another feature in its congestion control algorithm; a *slow* slow start mechanism. The rate of increasing its window size in slow start phase is a half of that in TCP Tahoe and TCP Reno. Figure 2 show Window control of TCP vegas, the details of the algorithm are as follows:

1. First, the source computes the expected flow rate *Expected = CWND / BaseRTT* , where *CWND* is the current window size and *BaseRTT* is the minimum *round trip time*.
2. Second, the source estimates the current flow rate by using the actual *round trip time* according to *Actual = CWND / RTT* , where *RTT* is the actual *round trip time* of the packet.
3. The source, using the expected and actual flow rate, computes the estimated backlog in the queue from *Diff = (Expected – Actual) BaseRTT* .
4. Based on *Diff*, the source updates its window size .

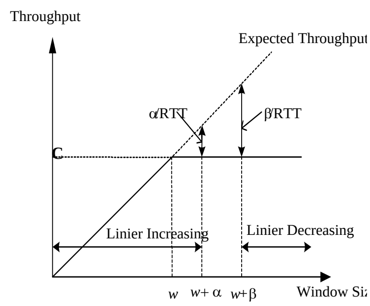


Figure 2. Window control of TCP Vegas.

2.2 Congestion control TCP Multi metric

TCP multi meric use end-to-end measurements to identify the presence of various ad hoc network conditions that if left unchecked, degrade throughput. This technique, first determine the network states that TCP must monitor then determine what available end-to-end metrics can be used to accurately identify network state so the TCP can mapping from metric measurements to the target states.

2.3 Congestion control algorithm challenges in MANET

The performance of TCP degrades in Ad hoc networks. This is because TCP has to face new challenges due to several reasons specific to Ad hoc networks: lossy channels (BER), route change, network partition, reconnection and disconnection (path break).

Effect of a High BER*:* Bit errors cause packets to get corrupted which result in lost TCP data segments or acknowledgment. When acknowledgment do not arrive at the TCP sender within a short amount of time [the retransmit timeout (RTO)], the sender retransmits the segment, exponentially backs off its retransmit timer for the next retransmission, reduces its congestion control window threshold, and *closes its congestion window to one segment*. Repeated errors will ensure that the congestion window at the sender remains small resulting in low throughput [10].

Effect of Route Recomputations*:* When an old route is no longer available, the network layer at the sender attempts to find a new route to the destination. It is possible that discovering a new route may take significantly longer than the RTO at the sender. As a result, the TCP sender times out, retransmits a packet, and invokes congestion control. Thus, when a new route is discovered, the throughput will continue to be small for some time because TCP at the sender grows its congestion window using the slow start and congestion avoidance algorithm. This is clearly undesirable behavior because the TCP connection will be very inefficient. If we imagine a network in which route computations are done frequently (due to high node mobility), the TCP connection will never get an opportunity to transmit at the maximum negotiated rate.

Effect of Network Partitions: It is likely that the *ad hoc* network may periodically get partitioned for several seconds at a time. If the sender and the receiver of a TCP connection lie in different partitions, all the sender’s packets get dropped by the network resulting in the sender invoking congestion control. If the partition lasts for a significant amount of time (say, several times longer than the RTO), the situation gets even worse be cause of a phenomena called *serial timeouts*. A serial timeout is a condition wherein multiple consecutive retransmissions of the same segment are transmitted to the receiver while it is disconnected from the sender. All these retransmissions are, thus, lost. Since the retransmission timer at the sender is doubled with each unsuccessful retransmission attempt (until it reaches 64 s), several consecutive failures can lead to inactivity lasting one or two *minutes* even when the sender and receiver get reconnected.

Effect of Multipath Routing: Some routing protocols maintain multiple routes between source destination pairs, the purpose of which is to minimize the frequency of route recomputation. Unfortunately, this sometimes results in a significant number of out-of-sequence packets arriving at the receiver. The effect of this is that the receiver generates duplicate acknowledgments (ACKs) which cause the sender (on receipt of three duplicate ACKs) to invoke congestion control.

3. Evaluation and simulation design.

3.1 Congestion control algorithm in wired

In this section we investigated congestion control in TCP Reno and TCP Vegas performance. We create bottleneck link and each TCP share the bandwidth with the same TCP.

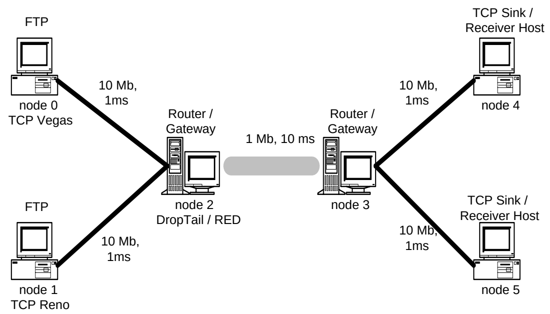


Figure 3. TCP Performance topology

We consider drop-tail and RED (Random Early Detection) as scheduling disciplines at the router buffer. Figure 3 show the topology N TCP Reno or TCP Vegas at bottleneck link.

Second, According to the past researches, a TCP Vegas version is able to achieve higher throughput than TCP Reno versions. By focusing on the situation where TCP Reno and Vegas connections share the bottleneck link, in this section we investigate the compatibility between TCP Vegas algorithm and Reno. We compare throughput performances of two versions.

3.2 Congestion Control in MANET

We analysis and evaluated TCP performance over adhoc network and then we examine TCP Multi metric. First scenario we use multihop to indicate delay, route change, channel error.

We use NS-2 simulator, 4 nodes with 3 hop in 1000 m x 1000 m, the sender send the packet after 10 second. The packet sizes is 512 bytes, each simulation lasts 100 second. We use Newreno for the TCP standard and the wireless link bandwidth is 2 Mbps with maximum window size eight packets.

In second scenario 20 wireless nodes roam freely in a 500 m x 500 m topology following a *random waypoint* mobility pattern, in which the pause time is zero so that each node is constantly moving. The wireless link bandwidth is 2M bps, and IEEE 802.11 and DSR are used as MAC and routing layer protocols. Figure 4 show the illustration of the second scenario. TCP NewReno[6] is used with a maximum window size of eight packets.

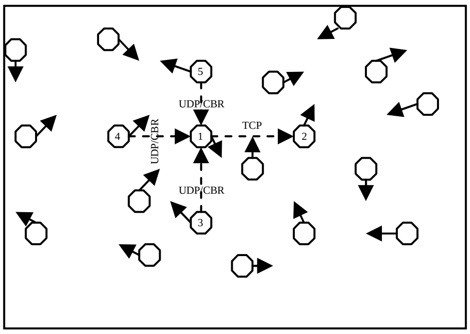


Figure 4. illustration of the second scenario

The packet size is 1460 bytes. Each simulation lasts 100 seconds. To introduce congestion, three competing UDP/CBR flows are run within the time intervals of [50,250],[100,200] and [130,170], respectively. Each UDP flow transmits at 180 Kbps.

3.3 TCP Multi metric

Knowing whether the current network is congested or not is important. As it turns out, proper congestion identification proves to be the biggest improvement to TCP in ad hoc networks. Previous research has indicated that identifying the following network states is necessary to improve TCP performance over ad hoc networks. Congestion (CONG): Here, we define congestion as queue build-up and packets being dropped due to buffer overflow at some nodes. Channel Error (CHERRes in MAN When random packet loss occurs, without slowing down, the sender should re-transmit the lost packet [3][4]. Route Change (RTCHG): The delivery path between the two end hosts can change from time to time, with disconnections that are too short-term to result in a TCP timeout. The receiver may experience a short burst of out-of-order packet delivery or packet losses. In both cases, the sender should estimate the bandwidth along the new route by setting its current sending window to the current slow start threshold, and initiating the congestion avoidance phase [2]. Disconnection (DISC): When the delivery path is disconnected for long enough to cause a TCP retransmission timeout.

3.3.2 Multi metric based on end-to-end measurement.

End-to-end measurement is widely used in TCP. The round trip time (RTT) is maintained by the TCP sender to calculate the retransmission timeout. Previous work uses delay related metrics to measure the congestion level of the network. A challenge in ad hoc networks is that packet delay is no longer only influenced by network queue length, but also is susceptible to other conditions such as random packet loss, routing path oscillations, MAC layer contention, etc. These conditions make such measurement highly noisy. Rather than pursue any single metric that is robust to all dynamics of the network, we devise four end-to-end metrics that tend to be influenced by different conditions so that the noise independence among them can be exploited by multi-metric joint identification.

Inter-packet delay difference *IDD* Metric *IDD* measures the delay difference between consecutive packets (Table 1). It reflects the congestion level along the forwarding delivery path by directly sampling the transient queue size variations among the intermediate nodes. Upon each packet arrival, the receiver calculates the IDD value. In addition, unlike the conventional inter-packet arrival delay (IAD), IDD is unaffected by random channel errors and packet sending behaviors. However, in an ad hoc network, there are still a number of situations in which *IDD* values might give an incorrect estimation of congestion. For example, *IDD* can be influenced by non-congestion conditions like mobility induced out-of-order packet delivery.

Short-term throughput *STT* Compared with IDD, STT is also intended for network congestion identification. However, it provides observation over a time interval *T* (half of the RTT), and is less sensitive to short term out-of-order packet delivery than IDD. Therefore, STT is more robust to transient route changes, which can be very frequent in a mobile ad hoc network. However, using STT alone to detect netework congestion can be susceptible to measurement noise introduced by bursty channel error, network disconnections or altering TCP source rates.

Packet out-of-order delivery ratio *POR* A packet is counted as being out-of-order if it arrives after a packet that was sent later than it (by the same TCP sender). The receiver records a maximum sending time for all the received packets from the TCP connection, denoted by T_{maxtz} . Every received packet that has a sending time-stamp more than T_{maxtz} is added into POR . POR is intended to indicate a route change event. During the route switching period, multiple delivery paths exist. Packets along the new path may catch up, and those along the old path are then delivered out-of-order.

Packet loss ratio *PLR* At each time interval [t, t + T], we compute this metric as the number of missing packets in the current receiving window. POR can be used to measure the intensity of channel error.

Metric	Definition
IDD	$A^{i+1} - A^{i} - (S^{i+1} - S^{i})$, where A^{i} is the arrival time of packet i and S^{i} is its sending time from the sender
STT	$N_p(T)/T$, where $N_p(T)$ is the # of received packets during intrval T
POR	$N_{po}(T)/N_p(T)$ where $N_{po}(T)$ is # of out-of-order packets during T
PLR	$N_l(T)/N_p(T)$ where $N_l(T)$ is # of lost packets during intrval T

(a)

Metric	IDD an STT	POR	PLR
CONG	(High,Low)	*	*
RTCHG	Not (High,Low)	HIGH	*
CHERR	Not (High,Low)	*	HIGH
DISC	(*, ≈ 0)	*	*
NORMAL	Default		

(b)

Table 1 (a) Multi Metric (b) Metrics pattern in five network states

From the last research [8] when the maximum network queue size exceeds half of the buffer capacity (25 packets), IDD is clearly *high* and STT clearly *low*. We formalize this observation by defining a value to be HIGH or LOW if respectively it is within the top or bottom 30% of all samples. However, when the network queue size is small (non-congestion case), both IDD and STT vary from LOW to HIGH Specifically, we identify a congestion state when both IDD is HIGH and STT is LOW, and non-congestion state if otherwise.

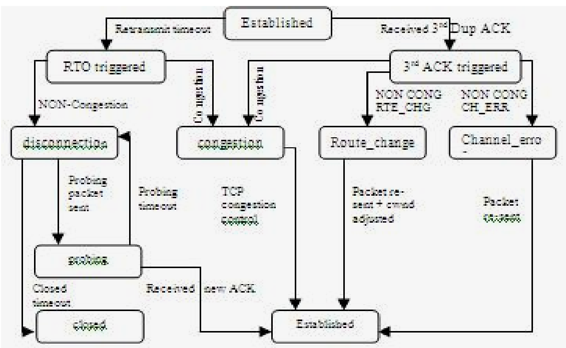


Figure 5.TCP Multi metric state mechanism

Combining these two cases, multi-metric identification improves the accuracy in non-congestion states while maintaining a comparable level of accuracy in the congestion state. Therefore, it achieves better identification performance over a variety of network conditions. The key insight here is that in the non-congestion state, IDD and STT are influenced differently by various network conditions, such as route change and channel error; while in congestion state, they are both dominated by prolonged queueing delay. Table 1 show the multi metric and metrics pattern in five network states. Figure 5 show the TCP Multi metric state mechanism.

4. Analysis of Standard Congestion Control and Multi Metric Algorithm.

4.1 Congestion control algorithm in wired

Based on the simulation of two of equal algorithm and N equal algorithm of TCP Reno and Vegas which implements drop tail and RED (Figure 7 and 8) can be recognized that the performance of Vegas TCP algorithm is better than Reno TCP algorithm. It is verified by whether the Throughput of TCP Vegas are higher than TCP Reno (Figure 6 – 8).

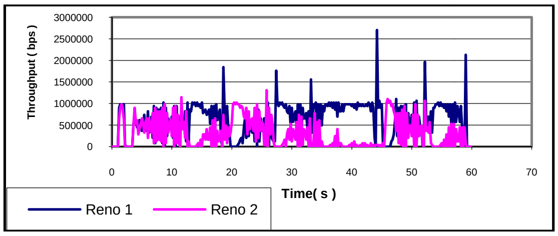


Figure 6. TCP Reno Performance with Drop tail gateway.

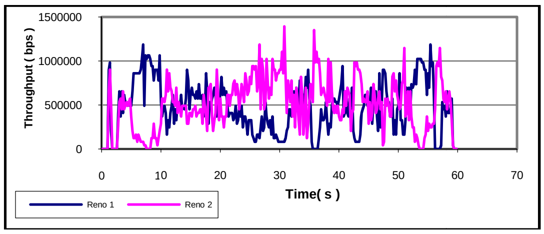


Figure 7. TCP Reno Performance with RED gateway.

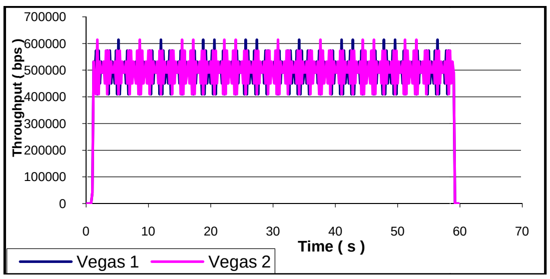


Figure 8. TCP Vegas Performance with RED and Drop tail gateway.

From simulation result and interaction analysis of N equal algorithm both Vegas and Reno, it is obtained that as the number of users' increase, the more data flow number, then lessen the index.

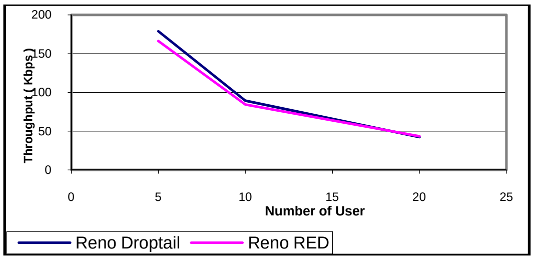


Figure 9. TCP Reno Throughput vs N User with Drop tail and RED gateway.

As the increment number of user that attempt to use bandwidth, then the allocation of each is reduced consequently lessen the throughput. Then the cwnd get lower.

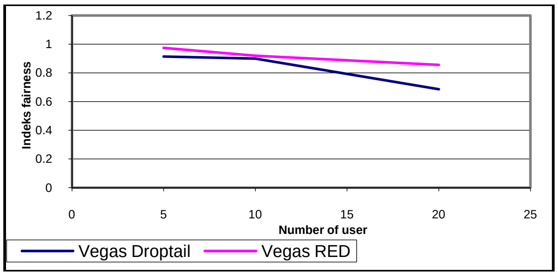


Figure 10. TCP Vegas Throughput vs N User with Drop tail and RED gateway.

Vegas TCP experience unfairness and incompatibility in bandwidth dividing when interacted in link bottleneck comparing to Reno. The RED implementation will improve the fairness index opposing decreasing total throughput.

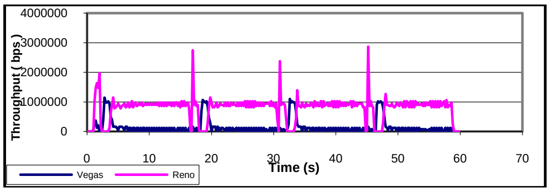


Figure 12. TCP Vegas and Reno Compatibility with Drop tail gateway.

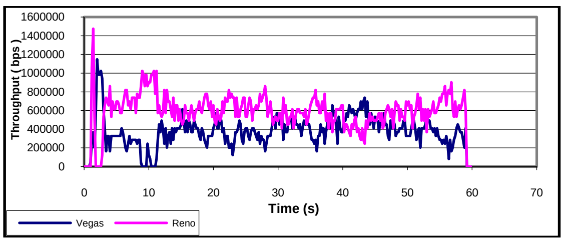
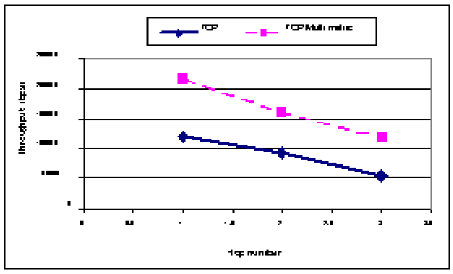


Figure 13. TCP Vegas and Reno Compatibility with RED gateway

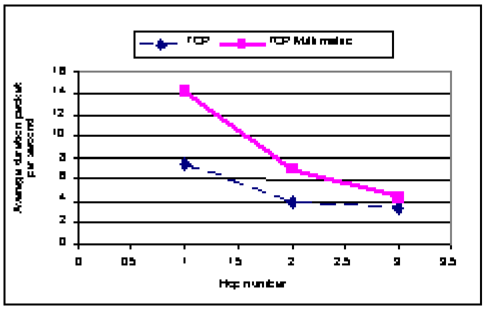
Although RED performance is better than drop tail, RED still has weakness which is each relation experience equal loss rate even though using lesser value that is should be, but experience packet loss. This shows incapability of RED to service a fair share because of adaptive relation which is aggressive user even though the congestion is not to worst.

4.2 Congestion Control Algorithm In MANET

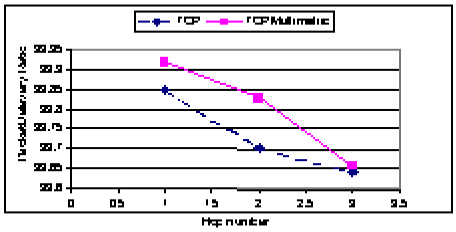
We analysis and evaluated TCP performance over adhoc network and then we examine TCP Multi metric. We use five scenario to evaluated TCP and TCP multi metric performance. First scenario we use multihop to indicate delay, route change, channel error.



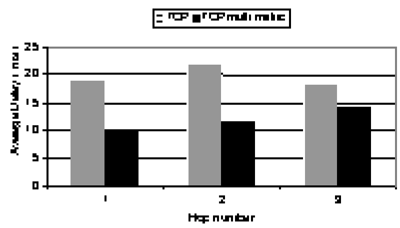
(a)



(b)



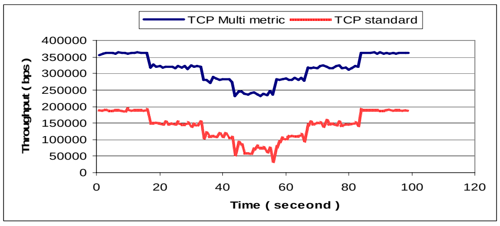
(c)



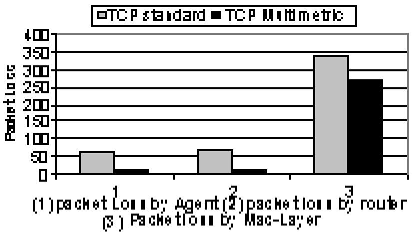
(d)

Figure 11. TCP and TCP Multi metric performance at mutihop adhoc network of varying length (in hops); (a) throughput, (b) average duration packet per second, (c) packet delivery ratio, (d) average delay

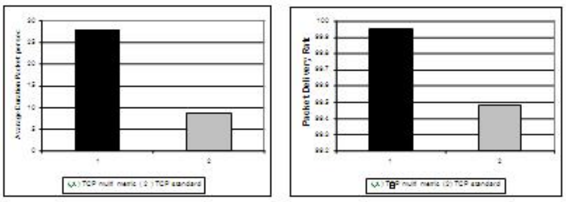
We use newreno for the TCP standard and the wireless link bandwidth is 2M bps with maximum window size eight packets. In figure 11 we can see the TCP multi metric has better performance than TCP standard. TCP multi metric get higher throughput, average duration packet, and packet delivery ratio but lower in average delay. In second scenario 20 wireless nodes roam freely in a 500 m x 500 m topology following a *random waypoint* mobility pattern, in which the pause time is zero so that each node is constantly moving. The wireless link bandwidth is 2M bps, and IEEE 802.11 and DSR are used as MAC and routing layer protocols. TCP NewReno[12] is used with a maximum window size of eight packets. The packet size is 1460 bytes. Each simulation lasts 100 seconds. To introduce congestion, three competing UDP/CBR flows are run within the time intervals of [50,250],[100,200] and [130,170], respectively. Each UDP flow transmits at 180 Kbps.



(a)



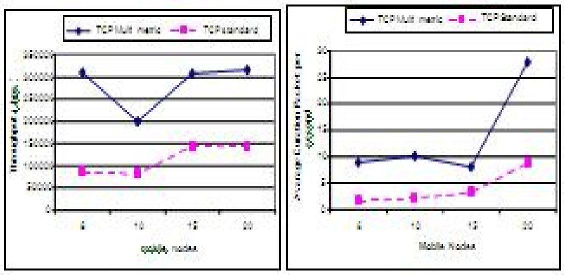
(b)



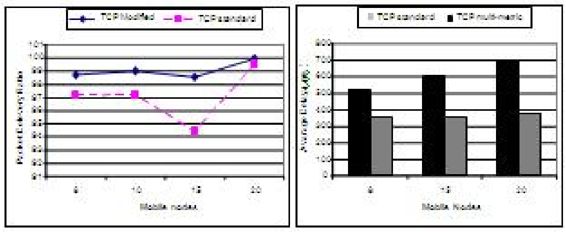
(c) (d)

Figure 12. TCP standard and TCP Multi metric performance at competitive adhoc network; (a) throughput, (b) packet loss, (c) average duration packet per second, (d) packet delivery ratio.

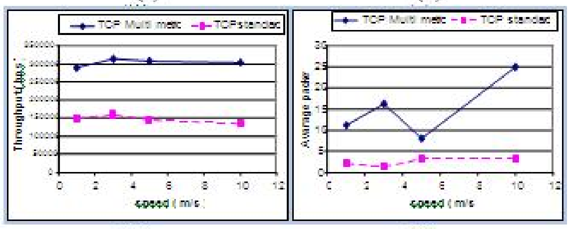
From figure 12 we can see that TCP multi metric has higher throughput then TCP standard. Packet loss by router is high that indicated packet loss is cause by route change. TCP standard relies on packet loss as an indication of network congestion then triggers congestion control algorithm and decrease the throughput. TCP performance degrades significantly in wireless adhoc due to the misinterpretation and an unable to detect appropriate condition in adhoc network. In other scenario we use the same competitive adhoc network then we change the number of mobile nodes, speed, and traffic.



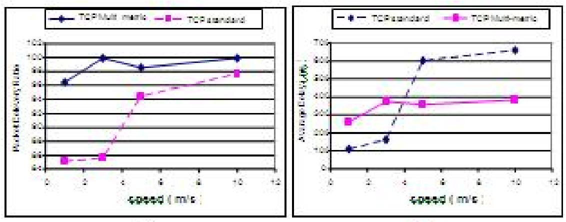
(a) (b)



(c) (d)



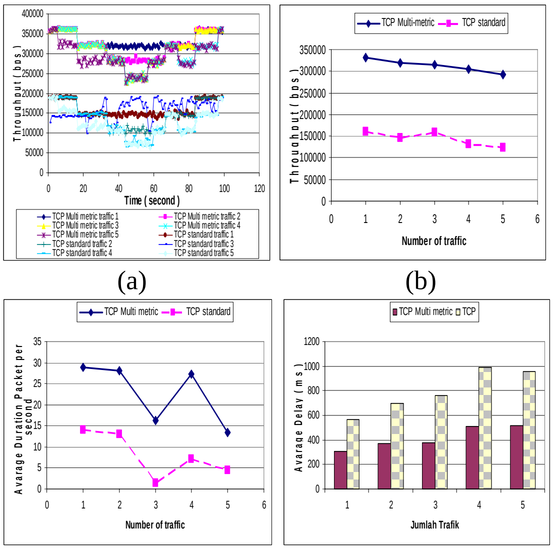
(e) (f)



(g) (h)

Figure 13. TCP Multi metric and TCP standard performance at competitive adhoc network : a-d number of mobile nodes increase; e-f speed of mobile nodes increase; (a & e) Througput, (b & f) Avarage duration packet per second, (c & g) packet delivery ratio, (d & h) average delay

When acknowledgment do not arrive at the TCP standard sender within a short amount of time [the retransmit timeout (RTO)] cause by route change and channel error due to number of mobile nodes and speed increase, the sender retransmits the segment, exponentially backs off its retransmit timer for the next retransmission, reduces its congestion control window threshold, and *closes its congestion window to one segment*. Repeated errors will ensure that the congestion window at the sender remains small resulting in low throughput. On the other hand TCP multi metric use end-to-end measurements to identify the presence of various ad hoc network conditions and get higher throughput and lower average delay.



(a) (b)

(c) (d)

Figure 14. TCP multi metric and TCP standard performance at competitive adhoc network with the number of traffic increase : (a) throughput (b) Throughput each connection (c) packet delivery ratio (d) average delay

When an old route is no longer available, the network layer at the sender attempts to find a new route to the destination. It is possible that discovering a new route may take significantly longer than the RTO at the sender. As a result, the TCP standard sender times out, retransmits a packet, and invokes congestion control. Thus, when a new route is discovered, the throughput will continue to be small for some time because TCP standard at the sender grows its congestion window using the slow start and congestion avoidance algorithm. End-to-end measurements are used to detect congestion, disconnection, route change, and channel error, and each detection result triggers corresponding control actions, with this mechanism TCP multi metric get better throughput, average duration packet per second and lower avarge delay in competitive adhoc network with the number of traffic increase.

5. Conclusion

We analysis and evaluated congestion control algorithm performance in TCP over wired and adhoc network we analysis and evaluated TCP standard performance over adhoc network and then we examine and explores TCP that has been modified for adhoc network and proposed with multi metric concept based on end-to-end measurement. The fundamental problem of transport protocol design in mobile ad hoc networks is that such networks exhibit a richer set of behaviors, including congestion, channel error, route change and disconnection, that must be reliably detected and addressed. Detection is challenging because measurement data is noisy. Our simulations show TCP with standard congestion control algorithm performance degrades significantly in wireless adhoc due to the misinterpretation and an unable to detect appropriate condition in adhoc network, in wired we can see that TCP standard detect the congestion by packet loss. The congestion control TCP multi metric is able to significantly reduce the probability of false detection and greatly improves the performance with get higher throughput and lower delay average then congestion control in TCP standard.

References:

[1] G. Holland and N. Vaidya, “Analysis of TCP performance over mobile ad hoc networks,”.

[2] P. Sinha, N. Venkitaraman, R. Sivakumar and V. Bharghavan, “WTCP: A reliable transport protocol for wireless wide-area networks,” .

[3] H. Balakrishnan, S. Seshan, E. Amir, and R. Katz, “Improving TCP/IP performance over wireless networks,”.

[4] H. Balakrishnan, and R. Katz, “Explicit loss notification and wireless web performances,” *Globecom’98*.

[5] S. Floyd, “TCP and explicit congestion notification,” *ACM CCR*, 1994.

[6] M.Allman, V.Paxson, and W. Stevens, ”TCP Congestion Control” *RFC 2581* April 1999.

[7] S. Biaz, N. Vaidya, ”Discriminating congestion losses from wireless losses using inter-arrival times at the receiver” *IEEE Symp. ASSET’99* March 1999

[8] Zhenghua Fu, Ben Greenstein, Songwu Lu, ”Design and implementation of a TCP-Friendly transport protocol for ad hoc wireless networks” UCLA Technical Report, 2002.

[9] H. Balakrishnan, V. Padmanabhan, S. Seshan, and R. Katz, “A comparison of mechanisms for improving TCP performance over wireless links,” *IEEE Transactions on Networking*, vol. 5, no. 6, pp. 756–769, Dec. 1997.

[10] K. Brown and S. Singh, “M-TCP: TCP for mobile cellular networks,” *ACM SIGCOMM Computer Communication Review*, vol. 27, no. 5, pp. 19–43, Oct. 1997.

[11] M. Gerla, K. Tang, and R. Bagrodia, “TCP performance in wireless multihop networks,” *WMCSA’99*.

[12] L.S. Brakmo, S. W.O’Malley, and L. L. Peterson,” TCP Vegas: New techniques for congestion detection and avoidance,” . Departement of Computer Science University of Arizona.

[13] Kenji Kurata, Go Hasegawa, Masayuki Murata, 2000, “Fairness Comparison between TCP Reno and TCP Vegas for Future Deployment of TCP Vegas.

[14] S. Floyd, V. Jacobson, 1993, Random early detection gateways for congestion avoidance, University of California, IEEE/ACM.

[15] L.S Brakmo, S. W. O’Malley, and L. L. Peterson, “TCP Vegas : New techniques for congestion detection and avoidance”.

PERFORMANCE EVALUATION OF MULTI-RADIO AOVD IN HYBRID WIRELESS MESH NETWORKS BASED ON MANHATTAN MOBILITY MODEL

Kassahun Endris Adem[1], Mohamed Othman[2]

[1,2]Dept of Communication Tech and Network, Universiti Putra Malaysia
[1]kasspro@gmail.com, [2]mothman@fsktm.upm.edu.my

Abstract

Wireless Mesh Networks (WMNs) have recently gained increasing attention and have emerged as a technology with great potential for a wide range of applications due to their self-configuring and self-healing capabilities, as well as their low equipment and deployment coasts. WMNs is a super-set of traditional mobile ad-hoc networks (MANETs), where it's comprised of mobile client devices. A WMN can contain relatively static devices, mesh routers. In hybrid WMNs is characterized by high level of heterogeneity, since static routers are typically much less resource constrained than mobile client's devices, and are also often equipped with multiple radio interfaces. One of the key features of wireless mesh networks is their limited capacity and scalability due to co-channel interference. A simple and relatively low-cost approach to address this problem is the use of multiple wireless network interfaces (radios) per node. Operating the radios on distinct orthogonal channels permits effective use of the frequency spectrum, thereby, reducing interference and contention. In this paper, we evaluate the performance of the multi-radio Ad-hoc On-demand Distance Vector (AODV) routing protocol under varying mobility and different node configurations, in Manhattan traffic mobility model with a specific focus on hybrid WMN. Our simulation results show that multi-radio AODV gains significant performance over the standard AODV protocol.

Keywords: Routing; Multi-radio; wireless mesh network; Manhattan; AODV

1. Introduction

As various wireless networks evolve into the next generation to provide better services, a key technology, wireless mesh networks (WMNs), has emerged recently and gained considerable popularity due to their low cost, self-configuring and rapid deployment capabilities. A WMN is dynamically self-organized and self-configured, with the nodes in the network automatically establishing and maintaining mesh connectivity among themselves. This feature brings many advantages to WMNs such as low up-front cost, easy network maintenance, robustness, and reliable service coverage. There are two types of nodes in WMN [1]: MESH_ROUTERs and MESH_CLIENTs. MESH_ROUTERs form the backbone infrastructure and provide connectivity between MESH_CLIENTS. They may also provide access to a wired network. MESH_ROUTERs are also typically equipped with multiple wireless interfaces (radios) (e.g., IEEE 802.11b/g and 802.11a NICs providing 3 and 12 non-overlapping frequency channels respectively which could be used simultaneously within a neighbourhood) are, therefore, able to establish high capacity connections by using multiple orthogonal channels. MESH_CLIENTs are mobile devices, which take advantage of the existing communication infrastructure provided by the MESH_ROUTERs. Due to their mobility, MESH_CLIENTS are typically more resource constrained than MESH_ROUTERS. WMNs technology have been proposed in numerous application scenarios such as broadband home networking, community and neighbourhood networking, building automation, Intelligent transport system, high speed metropolitan area networks, local area networks (LANs) for Hotels, Malls, Parks, Trains, etc and Ad hoc deployment of LANs such as public safety, rescue and recovery operations, high speed mobile video applications on board public or real time racing car telemetry.

Based on the architectures, deployment configuration and functionality of nodes, WMNs can be broadly categorized into three main types [1]: Infrastructure mesh, client mesh and hybrid mesh networks. In Infrastructure architecture type [2], the MESH_ROUTERs forming an infrastructure for MESH_CLIENTs that connected to them. The WMN infrastructure/backbone can be built using various types of radio technologies, in addition to the mostly used IEEE 802.11 technologies. The MESH_ROUTERs form a mesh of self-configuring, self-healing links among themselves. With gateway functionality, MESH_ROUTERs can be connected to the Internet. This approach, also referred to as infrastructure meshing, provides backbone for conventional MESH_CLIENTs and enables integration of WMNs with existing wireless networks, through gateway/bridge functionalities in MESH_ROUTERs. Conventional MESH_CLIENTs with the same radio technologies as MESH_ROUTERs, they can directly communicate with MESH_ROUTERs. In this architecture, clients have a passive role and do not contribute to the mesh network infrastructure. This is similar to a traditional WLAN, with the key difference that the wired backbone is replaced with a wireless multi-hop network.

In client mesh architecture, client meshing provides peer-to-peer networks among MESH_CLIENT devices wherein client nodes constitute the actual network to perform routing and configuration functionalities as well as providing end user applications to customers. Thus, a MESH_ROUTER is not required for these types of networks. A client WMN is basically the same to a traditional pure ad-hoc network [10].

The architecture of hybrid WMNs [2] as shown in Figure 1 is the most basic type of WMN, combining the concepts of infrastructure and client mesh networks. It consists of relatively static MESH_ROUTERs which form the backbone of the network and mobile clients that can act as a dynamic extension of the static infrastructure part of the network, by implementing routing and packet forwarding functionality. Mesh_clients can access the network through Mesh_ROUTERs as well as directly meshing with other Mesh_clients. While the infrastructure provides connectivity to other networks such as the Internet, Wi-Fi, WiMAX, cellular, and sensor networks; the routing capabilities of clients provide improved connectivity and coverage inside the WMN. The hybrid mesh architecture is very flexible and allows combining the benefits of both infrastructure and client meshes [11]. In this paper, we mainly focus our performance evaluation on this architecture type.

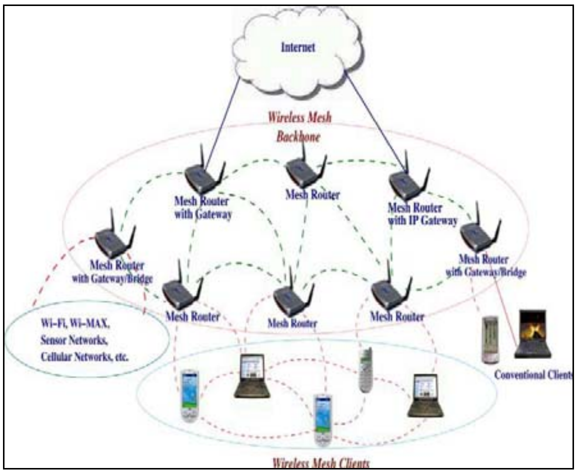


Figure 1. Hybrid WMNs [2] explain

The rest of this paper is organized as follows: section 2 provided an overview of the most relevance related work. Section 3 described the multi-radio extensions to AODV. Section 4 discussed simulation results and their discussion. Section 5 concludes this paper.

2. Related Works

There are several researches that related to multi-radio routing protocols for WMNs have been proposed over these years. Those works had tried to address the problem of limited capability and scalability of multi-hop wireless mesh networks. Here we categorised some of the most relevant related works in proactive, reactive, and hybrid routing protocols and discussed briefly as follow.

2.1 Proactive Routing Protocols

Proactive routing protocols computed and maintained routes between nodes pairs on a periodic basis by sharing tables or distance vectors. Destination-Sequenced Distance-Vector for Multi-Channel (DSDV-MC) [7] and Hyacinth [14] are proactive routing protocols, that proposed to address the limited capability and scalability of single shared interface in wireless mesh networks.

DSDV-MC routing protocol is a multi-channel extension of DSDV routing protocols. The routing table at each node lists all available destinations, the number of hops to each destination, and the channel indices of neighbouring nodes. To maintain the consistency of routing tables in a dynamically varying topology, each node periodically transmits updates. Nodes also transmit updates immediately when significant new information is available, such as a topology change or a channel switch.

Hyacinth is a multi-channel static wireless mesh network protocol that uses multiple radios and channels to improve the network performance. It supports a fully distributed channel assignment algorithm, which can dynamically adapt to varying traffic loads. Hyacinth's channel assignment algorithm breaks a single-channel collision domain into multiple collision domains, each operating on a different frequency. Nodes control and coordinate allocation of channels to interfaces via periodic exchange of messages containing channel usage status. Using the per-channel total load information a node can issue a change channel message to its neighbour in order to switch to a less utilised channel.

2.2 Reactive Routing Protocols

In these routing protocols, route between nodes are maintained only when required. Single-radio multi-channel routing (MCRP) [15] and Multi-radio multi-channel routing (MCR) [5] are reactive routing protocols that proposed to address problem of single shared interfaces.

The MCRP [15] is a routing protocol specifically designed for networks with single-radio nodes, which support a channel switching delay of 80 μs or less [6]. The protocol assigns channels to data flows rather than assigning channels to interfaces. This implies that all nodes supporting a flow have to be on a common channel. The advantage of this mechanism is that once the route is established, nodes are not required to switch channels for the duration of the flow.

The MCR protocol [5] has been developed for dynamic, multi-radio WMNs. The protocol makes use of an interface switching mechanism to assign interfaces to channels. Two types of interfaces are assumed: fixed and switchable. Switching is carried out depending upon the maximum number of data packets queued for a single channel. The switching mechanism assists the MCR protocol in finding routes over multiple channels. MCR uses a new routing metric, which is computed as a function of channel diversity, interface switching cost and hop-counts. The diversity cost is assigned according to the least number of channels used in a route. Thus, a route with a larger number of distinct channels in a route is considered to be having a lower diversity cost.

2.3 Hybrid Routing Protocols

The hybrid routing protocols are generally employs proactive routing in the static portion of the network and reactive routing in the mobile portion of the network. AODV-Spanning Tree (AODV-ST) [12] and Multi-Radio Link Quality Source Routing (MRLQSR) [4] protocols are hybrid routing protocols that address the issues of limited capability and scalability of wireless mesh networks.

AODV-ST [12] is a hybrid routing protocol developed specifically for infrastructure mesh networks. The protocol uses the Expected Transmission Time (ETT) routing metric to select routes. AODV-ST has been designed with the aim of providing Internet access to MESH_CLIENTs with the help of one or more gateways. AODV-ST uses a proactive strategy to discover routes between the MESH_ROUTERs and the gateways, and a reactive strategy to find routes between MESH_ROUTERs. In the proactive case, the gateways periodically broadcast special ROUTE_REQUEST packets to initiate the creation of spanning trees. All subsequent ROUTE_ REQUEST packets with a better routing metric are used to update the existing reverse route to the gateway.

The MRLQSR [4] protocol has been developed for static community wireless networks. MRLQSR adapts Link Quality Source Routing (LQSR) to operate over multiple channels and multiple interfaces, using the Weighted Cumulative Expected Transmission Time (WCETT) metric. LQSR protocol combines link-state proactive routing with the reactive strategy from ad hoc networks.

MRLQSR protocol identifies all nodes in the wireless mesh network and assigns weights to all possible links. The link information including channel assignment, bandwidth and loss rates are propagated to all nodes in the network. The ETT on each link is computed using the Expected Transmission Count (ETX), bandwidth and packet loss. The ETT metric is further used to compute the WCETT, which defines the path metric designed for multi-radio WMNs.

The related work discussed various switching mechanisms (for channel or/and interface), which can be used to improve the routing performance of a network. However the accurate execution of these mechanisms in a mobile network entails the availability of a virtual switching protocol and incurs switching delays [3, 4]. The switching protocol may be integrated with the MAC layer or may be implemented in a higher layer protocol. In either case, it requires accurate synchronisation between the nodes negotiating an interface or channel switch [11]. In contrast, AODV-MR [11] considers a network where channels are statically allocated to interfaces, thus, multi-radio AODV protocols does not require exact synchronisation between nodes but is still able to make of multiple radios per node by establishing channel diverse paths. In this paper, we evaluate the performance of this AODV-MR in hybrid wireless mesh network based on Manhattan mobility model.

3. Multi-Radio AODV Protocol

The multi-Radio AODV protocol is a multi-homing extension to AODV protocol [11]. In AODV-MR, each node is equipped with multiple radios and each radio can operate on one of the multiple non-overlapping channels. Thus, the route discovery and maintenance process of standard AODV is needed to extend, in order to support multiple interfaces capability of multi-radio AODV protocols.

3.1 Route Discovery

When a route is required in AODV-MR, RREQ is broadcast on all interfaces. Intermediate nodes with one or more interfaces operating on a common channel receive the RREQ and create a reverse route that points towards the source node. If the RREQ is a duplicate, it is simply dropped. The first RREQ received by the destination or any intermediary node is selected and all other RREQs belonging to the same route discovery are discarded. The RREP is generated in response to the selected RREQ and is sent back to the source node on the established reverse route.

Each RREQ packet in AODV-MR contains an ID, source and destination IP addresses, sequence numbers, hop-count and control flags as standard AODV and additional interface number, as shown in Table 1. The ID field uniquely identifies the RREQ packet and the sequence numbers indicate the freshness of control packets. The hop-count represents the path length between the source and destination. The additional in AODV-MR maintains an interface

Table 1. ADOV-MR Routing Table

Destination IP	Destination Seq. No.	Destination Valid Flag	Flags	Network Interface	Hop Count	Next Hop	Precursors List	Lifetime

number in the routing table. The interface number indicates the network interface via which a next hop node, for a particular path, can be reached. This information permits sending of the RREP packet and future data packets on the correct interface.

The intermediate nodes, after updating their routing tables, broadcast the RREQ packet on all interfaces except the one on which the RREQ packet was initially received. If the number of nodes in the network is n and the number of interfaces per node is i, the maximum number of RREQ packets propagated in the network will be i. (n – 1).

The propagation of RREQ packets continues until the IP time to live (TTL) expires. If the RREQ packet is received by the destination itself or any node with a fresh route to the destination, a RREP packet is sent to the source node using a loop free path. When the RREP packet is received by intermediate nodes, a forward route to the destination is established by creating the corresponding routing table entries, including the interface number. The overall algorithms of multi-radio AODV route discovery process show in the Figure 2, flow chart.

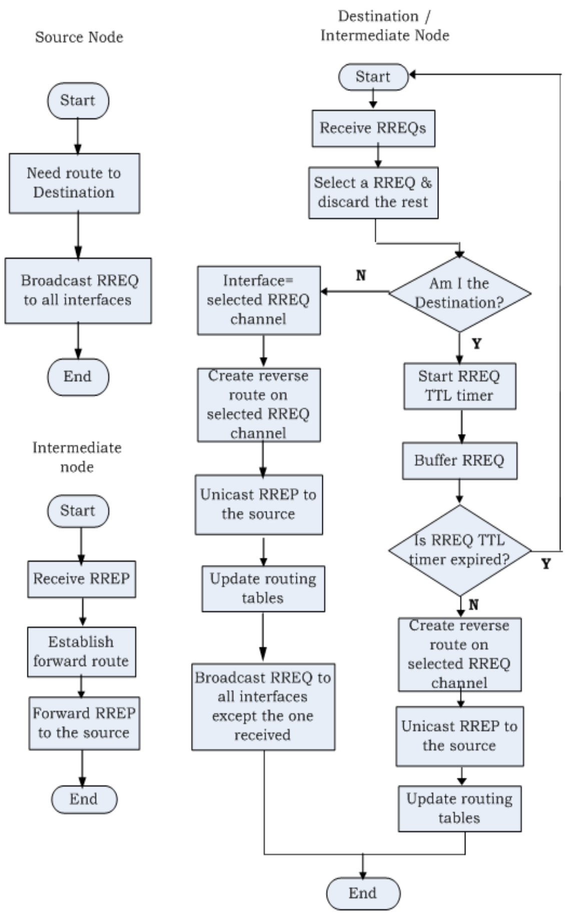


Figure 2. Multi-radio AODV route discovery flow chart

3.2 Route Maintenance

When a link in a route breaks, the node upstream of the break invalidates all its routes that use the broken link. After that, the node broadcasts a route error (RERR) message to its neighbours in all interfaces. The RERR message contains the IP address and interfaces number of each destination which has become unreachable due to the link break. Upon reception of a RERR message, a node searches its routing table to see if it has any route(s) to the unreachable destination(s) (listed in the RERR message) which use the originator of the RERR as the next hop. If such routes exist, they are invalidated and the node broadcasts a new RERR message to its neighbours. This process continues until the source receives a RERR message. The source invalidates the listed routes as previously described and reinitiates the route discovery process if needed.

4. Simulation Results and Discussions

4.1 Simulation Environment

To evaluate the performance of AODV-MR in hybrid wireless mesh network, we compare the performance of AODV-MR with the standard single-radio AODV through extensive simulation in NS2 [9]. To support multi-radio in NS2 simulation, we extend NS2 by using “Adding Multiple Interface Support in NS2” [13] document.

4.2 Network Topology

The simulation carried out with Manhattan mobility models, based on regular grid mesh network topology. In these topologies, the mesh network consists of static MESH_ROUTERs and mobile MESH_CLIENTs. A dense WMN covering an area of 1 square km is established using 25 static MESH_ROUTERs arranged in a regular 5X5 grid, as shown in Figure 3.

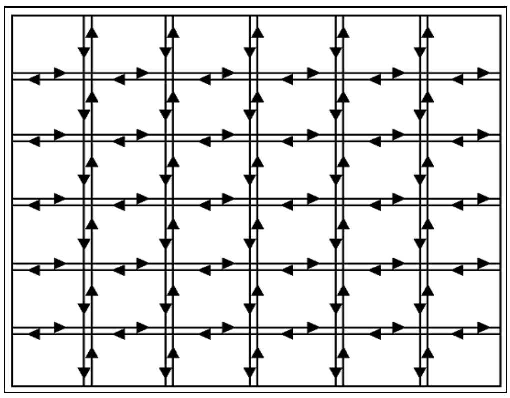


Figure 3. Manhattan Mesh Network Topology

MESH_ROUTERs is placed on cross points of each street. The distance between adjacent MESH_ROUTERs is set to 176m to ensure diagonal connectivity. The network further consists of 50 randomly placed mobile MESH_CLIENTs on street. The mobile node is allowed to move along the grid of horizontal and vertical lanes of the streets on the map. At an intersection of a horizontal and a vertical street, the mobile node can turn left, right or go straight with certain probability [16]. Concurrent UDP flows are established between randomly selected source and destination MESH_CLIENT pairs.

4.3 Simulation Setup

To evaluate the performance of Multi-Radio AODV protocol, intense simulations were conducted under varying MESH_CLIENT speeds based on Manhattan grid mobility model. Both standard AODV and AODV-MR routing protocols have been evaluated in this simulation.

The maximum speed of MESH_CLIENTs in this simulation varied from 0 to 20 mps. The packet size is set to 512 bytes and 30 number of concurrent UDP flows are set. The reset simulation parameters are listed in Table 2.

Table 2. Simulation Parameters Setup

Examined Protocols	AODV & AODV-MR
Simulation time	900s
Simulation area	1000X1000 m
Propagation model	Two-ray ground reflection
Maximum speed of MESH_CLIENTs	20 m/s
Transmission range	250 m
Traffic type	CBR(UDP)
Packet rate	32 pkts/s
Number of MESH_ROUTERs	25
Number of MESH_CLIENTs	50
Number of 802.11b radios per MESH_CLIENT	1
Number of 802.11b radios per MESH_ROUTER	6

4.4 Mobility Model

We use Manhattan Grid Mobility model for MESH_CLIENTs in our simulation. Manhattan mobility model use defined maps to emulate the movement pattern of mobile nodes on streets. As shown in Figure 3 above, the map is composed of a number of horizontal and vertical streets. Each street has 2 lanes for each direction (north and south direction for vertical streets, east and west for horizontal streets).

The mobile node (MESH_CLIENTs) is allowed to move along the grid of horizontal and vertical streets on the map. At an intersection of a horizontal and a vertical street, the mobile node can turn left, right or go straight. This choice is probabilistic: the probability of moving on the same street is 0.50, turning left is 0.25 and turning right is 0.25. The velocity of a mobile node at a time slot is dependent on its velocity at the previous time slot. Also, a nodes velocity is restricted by the velocity of the node preceding it on the same lane of the street. Thus, the Manhattan mobility model is also expected to have high spatial dependence and high temporal dependence. It too imposes geographic restrictions on node mobility.

4.5 Performance Metrics

In our simulation, we consider the following four typical performance metrics to evaluate the performance of multi-radio AODV routing protocols.

1) **Packet delivery ratio:** The ratio between the number of data packets successfully received by destination nodes and the total number of data packets sent by source nodes. This performance metric evaluates the best-effort traffic.
2) **Routing packet overhead:** The ratio between the total numbers of control packets generated and the total numbers of data packets that are success fully received. Routing overhead is important to compare the scalability of the routing protocols, the adoption to low-bandwidth environments and its efficiency.
3) **Average latency:** The mean time (in seconds) taken by data packets to reach their respective destinations. This includes all the possible time taken by buffering during route discovery latency, queuing at the interface queue, retransmissions at the MAC, and propagation and transfer times through channel.
4) **Path optimality:** The ratio between the length (number of hops) of the shortest possible path and the actual path taken by the data packets. This performance metrics depict the deviation of paths taken by data packets from the shortest possible paths.

4.6 Simulation: Based on Manhattan Mobility Model

This simulation is intended to evaluate the performance of multi-radio AODV in more realistic mobility model. We used Manhattan mobility model. This mobility model is useful in modelling movement in an urban area where a pervasive computing service between portable devices is provided and more realistic than random way mobility model. In this simulation, we have varied the maximum speed of the MESH_CLIENTs from 0 to 20mps in Manhattan mobility traffic. The results indicate that multi-radio AODV sustaining multiple ongoing flows, resulting high packet delivery ratio.

As shown in Figure 4, multi-radios AODV achieved higher PDR for all mobility speeds than standard AODV. Standard AODV achieves the PDR around 27% at Zero MESH_CLIENT mobility, then the PDR gradually decreases to 22% when the MESH_CLIENTs speed reach 20mps, whereas the

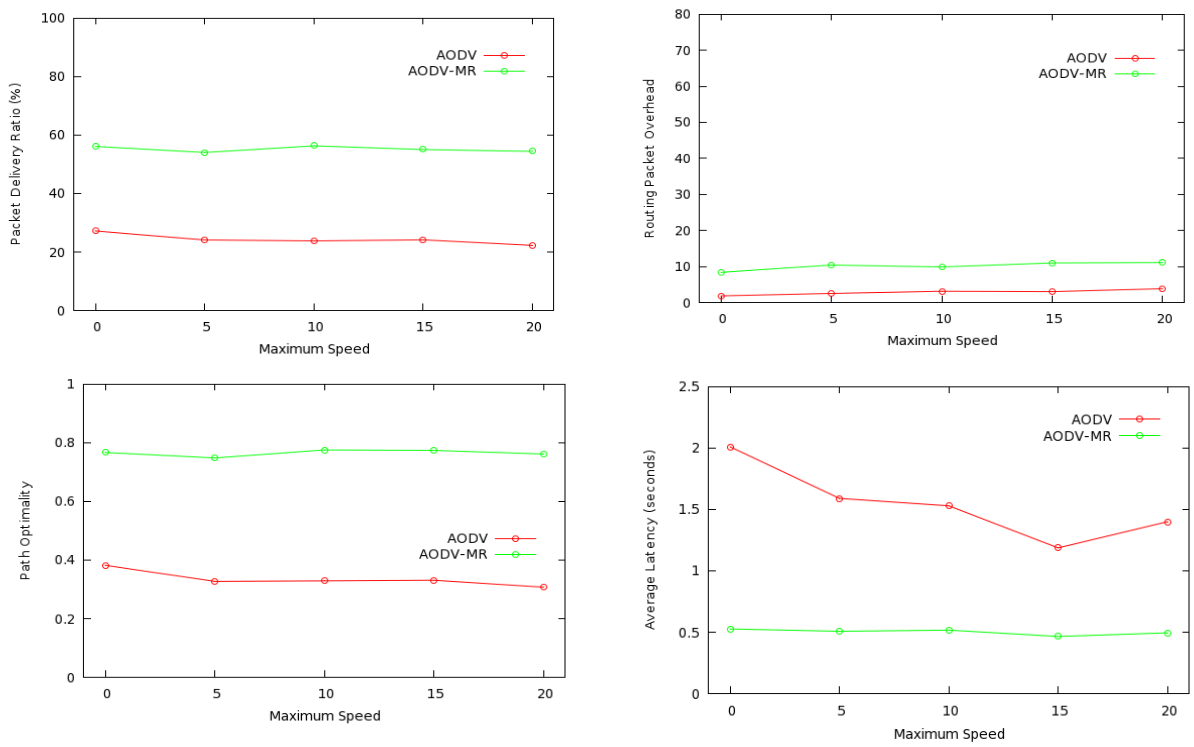


Figure 4. Result: Varying MESH_CLIENTs Speed Performed in Terms of Packet Delivery Ratio, Average Latency, and Path Optimality than that of the Standard AODV.

PDR of multi-radio AODV shown almost constant in all MESH_CLIENT speed. As the simulation results indicate that standard AODV loses more than 75% of the transmitted packets under varying node speeds. This loss is primarily attributable to the large traffic volume traversing nodes with single radios.

Nodes in standard AODV have to share a single radio among multiple flows observe congestion at the physical layer and consequently experience frequent packet drops. In contrast with AODV, the MESH_ROUTERs used by multi-radio AODV have six radios each, which facilitate multiple flows to be distributed over distinct physical layers. This is turn helps in Routing packet overhead of multi-radio AODV is considerably higher than the standard AODV. This fundamentally occurs due to the increased number of ROUTE REQUEST packets being transmitted on multiple radios during the route discovery process.

Multi-radio AODV in Manhattan mobility model outperforms standard AODV in terms of average latency and path optimality. As the result shown in the above Figure 4, the average latency of the packets using multi-radio AODV remains around 5 ms over all MESH_CLIENT mobility speed. Multi-radio AODV forms channel diverse paths with the help of multiple radios tuned to orthogonal channels. These paths help in minimising the latency and to discover shortest possible path in route discovery process. Standard AODV offer no option for channel diverse paths and, hence, exhibits higher latency with increasing traffic load.

5. Conclusion

Hybrid wireless mesh networks are composed of a combination of static MESH_ROUTERs and mobile MESH_CLIENTs. These two types of node differ considerably in terms of their capability to forward packets. MESH_ROUTERs are typically much less resource constrained than mobile MESH_CLIENTs, and can be equipped with multiple radio interfaces. These networks have a great potential for a wide range of applications owing to their self-configuring, self-optimising and self-healing capabilities, coupled with their low cost of equipment and deployment. One of the key shortcomings of wireless mesh networks is their limited capacity and scalability, which is typical for multi-hop wireless networks. It has been shown that using multiple radios per node, operating on orthogonal channels, can greatly increase the capacity of wireless mesh networks.

In this paper, we have presented a performance evaluation of a multi-radio version of the AODV routing protocol based on Manhattan mobility Model. Our simulation results had shown a significant improvement in terms of packet delivery ratio and latency over the single-radio AODV, especially under high load conditions in more realistic Manhattan mobility model. We believe there is a great potential to reduce the router overhead and further improve the performance multi-radio AODV, by developing adaptive routing algorithm, combination of both reactive and proactive routing algorithm. Since the routing algorithm adjusts itself to the network mobility, we are sure that it will reduce the routing information traffic overhead, and which leaves more resource for user traffic, this improved the overall performance of the network by making optimal use of the channel.

References

[1] Akyildiz, I.F., & Wang, X. (2005), A survey on wireless mesh networks, IEEE Communications Magazine, 43(9), S23–S30.

[2] Akyildiz, I.F., Wang, X., & Wang, W. (2005), Wireless Mesh Networks: a survey, Computer Networks journal (Elsevier), 47(4), pp.445-487.

[3] Chandra, R., & Bahl, P. (2004), MultiNet: Connecting to Multiple IEEE 802.11 Networks using a Single Wireless Card, In: Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM), 2, pp. 882 - 893.

[4] Draves, R., Padhye, J., & Zill, B. (2004), Routing in Multi Radio, Multi-Hop Wireless Mesh Networks, In: Proceedings of the 10th Annual International Conference on Mobile Computing and Networking, pp. 114-128.

[5] Kyasanur, P., & Vaidya, N.H. (2006), Routing and Linklayer Protocols for Multi-Channel Multi-Interface Ad Hoc Wireless Networks, SIGMOBILE Mobile Computing and Communications Review, 10, pp. 3-43.

[6] Kyasanur, P., & Vaidya, N.H. (2005), Routing and Interface Assignment in Multi-Channel Multi Interface Wireless Network, In: Proceedings of the IEEE Wireless Communications and Networking Conference, 4, pp. 2051- 2056.

[7] Lee, U, Midkiff, S.F, & Park, J.S. (2005), A proactive routing protocol for multi-channel wireless ad-hoc networks (DSDV-MC), In: International Conference on Information Technology: Coding and Computing, 2, pp. 710–715.

[8] Minseok. Oh, (2008), An Adaptive Routing Algorithm for Wireless Mesh Networks, In: The 10th International Conference on Advanced Communication Technology, Phoenix Park, Korea, pp. 2087-2091.

[9] NS. The Network Simulator, (1989), Available from: http://www.isi.edu/nsnam /ns/ [Accessed October 2008].

[10] Pirzada, A.A., & McDonald, C. (2004), Establishing trust in pure ad-hoc networks, In: Proceedings of the 27th Australasian Computer Science Conference (ACSC), Australian Computer Society, 26, pp. 47–54.

[11] Pirzada, A.A., Portmann, M., & Indulska, J. (2008), Performance analysis of multi-radio AODV in hybrid wireless mesh networks, Computer Communications, 31, pp. 885-895.

[12] Ramachandran, K., Buddhikot, M., Chandranmenon, G., Miller, S., Belding-Royer, E., & Almeroth, K. (2005), On the design and implementation of infrastructure mesh networks, In: Proceedings of the IEEE Workshop on Wireless Mesh Networks (WiMesh), IEEE Press, pp. 4–15.

[13] Ramon, A.C., & Jesus, P.C. (2007), Adding multiple interface support in NS-2, Available from: http://personales.unican.es /aguerocr/ [Accessed January 2009].

[14] Raniwala, A., & Chiueh, T.C. (2005), Architecture and Algorithms for an IEEE 802.11-based Multi-Channel Wireless Mesh Network, In: Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM).

[15] So, J., & Vaidya, N.H. (2004), A Routing Protocol for Utilizing Multiple Channels in Multi-Hop Wireless Networks with a Single Transceiver, Dept. of Computer Science and Coordinated Science Laboratory, University of Illinois at Urbana Champaign, Technical Report, Avaliabe from: http://www.crhc.Illinois.edu/wireless/papers/multichannel_routing.pdf/ [Accessed October 2008].

[16] User Manual for IMPORTANT Mobility Tool Generator in NS-2 Simulator, (2004), Available from: http://nile.usc.edu/important/software.html/ [Accessed April 2009].

ANALYSIS OF DELAY BOUND IN IEEE 802.11g WLAN OVER FIBER NETWORKS

Erna Sri Sugesti[1], Purnomo Sidi Priambodo[2], Kalamullah Ramli[3], Bagio Budiardjo[4]

[1]Faculty of Electrical Engineering Department, Institut Teknologi Telkom
[2,3,4]Faculty of Engineering, University of Indonesia
[1]ern@ittelkom.ac.id, erna.sri@ui.ac.id, {[2]pspriambodo, [3]bbudi, [4]k.ramli}@ee.ui.ac.id

Abstract

In this paper, we elaborate a new delay calculation on IEEE 802.11g-over-Fiber networks theoretically. It is found that delays generated by optical fiber and optical transceiver are more dominant than the air propagation. In this case, by inserting optical transceiver delay made the fiber length difference of 100 m.

Keywords: WLAN, IEEE 802.11, optical fiber, delay,

1. Introduction

The technology of Radio over Fiber (RoF) has been developed since in the mid of '90s. In the beginning, the development of this technology is mainly to support wireless cellular networks as a backhaul or an access network. This technology is accepted by the market rapidly, it can be by the following factors: the migration of cell coverage from macrocells into microcells, the enabling unified operation of multicarrier frequency system from multioperator, the simplicity of the RoF realization with better performance than coaxial networks, and the ease of increasing capacity by using wavelength division multiplexing (WDM) techniques [8][3][1].

Recently the tremendous rapid growth of internet users brings onto last-mile network traffic problems, especially for wireless-LAN (WLAN) applications. The success implementation of RoF networks has motivated to explore another hybrid network between IEEE 802.11 Wireless-LAN (WLAN) and the optical fiber, which is called Fi-Wi (Fiber-Wireless) network. Intuitively there is no doubt that the hybrid network will perform better to some extent. In fact, however, the throughput of the hybrid network demonstrates worse than the WLAN [5]. One of the reasons is the make use of a carrier sense multiple access with collision avoidance (CSMA/CA) without any modification.

There are some kinds of WLAN technologies; in this case we decide to employ the utmost technology applied in the world is IEEE 802.11g [1]. The main characteristics of IEEE 802.11g are shown on Table 1 [11].

Table 1. Resume of IEEE 802.11g Technology [11]

Freq.	Physical	Access	Max range	Power
2,4 GHz	OFDM	CSMA/CA	50-100 m	Medium
Throughput				
Physical	Effective			Region
54 Mbps	=< 22 Mbps			Worlwide

This technology applies adaptive modulations which depend on the traffic situation and the distance between an Access Point (AP) and Mobile Units (MU), Table 2. From both tables, it can be seen that if the MU location closer to the AP it will obtain higher bit rate and vice versa. However, the maximum range is only 100 m. It means that other APs should be installed in order to extend the coverage area.

Tabel 2. Modulation Systems in IEEE 802.11g[9]

User Rate (Mbps)	*FEC Coding*	*Line Rate (Mbps)*	*∑Bits per SC*	*Mod. BW*	*Eff. (bps/Hz)*
54	3/4	72	1.5 Mbps	64-QAM	2.7
24	1/2	48	1 Mbps	16-QAM	1.2
12	1/2	24	500 kbps	QPSK	0.6
6	1/2	12	250 kbps	BPSK	0.3

The standard of IEEE 802.11g WLAN deploys a license-free frequency 2.4 GHz. It definitely becomes a crowded-usage bandwidth and high-density traffic where anyone can use it as a propagation medium. With a very short range of 100 m, overlapping cells or co-channel interference may occur easily if the wireless network design unconsiders a site survey in advance. Moreover, some products of this standard are equipped with Power Level (PoE) management. This denotes that the coverage areas may vary that depend on power setting of the AP.

To avoid such complex problems, the Fi-Wi network seem to give a proper solution. By using the optical fiber for extending the wireless coverage area to make a new WLAN cluster, we can penetrate that overlapping coverage area without doubting that the wireless electromagnetic disturbance may deteriorate the information signals.

In this paper, we investigate the delay mechanism in propagation media, especially on optical fiber and optoelectronic transceiver delay factors which can affect the network performance. The main purpose is to obtain the propagation delay of IEEE 802.11g WLAN over fibers numerically, which then we could have the maximum length of fiber in order to enlarge the coverage area and to increase the network throughput, as well. This preliminary study is based on theoretical approach.

2. Theoretical Background

2.1. Medium Access Control Structure

The family of IEEE 802.11 standard applies a carrier sense multiple access with collision avoidance (CSMA/CA) medium access protocol (MAC) with binary exponential, which called Distributed Coordination Function (DCF) and Point Coordination Function (PCF). DCF is a class of coordination function where the same coordination function logic is active in *every station* (STA) in the basic service set (BSS) whenever the network is in operation [13]. Meanwhile, PCF is similar to DCF but *only one STA* in a BSS at any given time. The relationship of both in MAC is shown in Fig. 1.

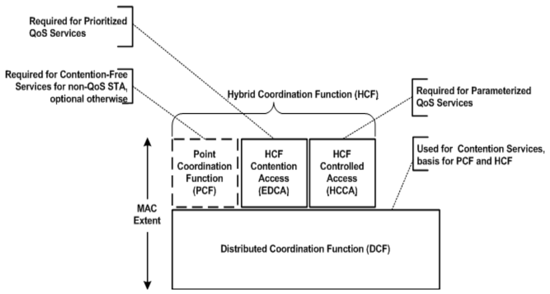


Figure 1. MAC Architecture [13]

The DCF creates two mechanisms, basic access and Request-to-Send/Clear-to-Send (RTS/CTS). Shortly, both mechanisms are depicted in Fig 2 and 3, respectively. Among those timing procedures, we concern to Short Interframe Space (SIFS) one since this is a basic timing and it plays key role in frame exchanges. The SIFS is constructed by several different timings as stated in Eq. (1) [13]. For this time being, this is not our concern now. Every standard has its determination values. For example, the SIFS of 802.11g is 10 μs [13]. Subsequently, the total round trip delay between frames must be less than this value. Unless, the loss frames will take place and it should be retransmitted. Next, the effect of this occurrence is decreasing network throughput.

$$SIFSTime = aRxRFDelay + aRxPLCPDelay + aMACProcessingDelay + aRxTxTurnaroundTime \tag{1}$$

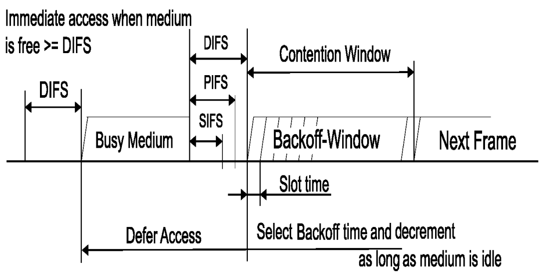


Figure 2. Basic Access Method [6]

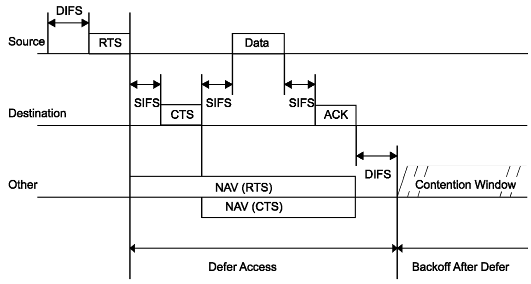


Figure 3. RTS/CTS/Data/ACK and NAV setting [6]

2.2. Optical Fiber and Optical Transceiver

The optical fiber is made of isolator materials, such as glass or plastics, that makes it immune to electromagnetic wave disturbance. Based on the index profile of the core, the optical fiber types are step index (SI) fiber and graded index (GI) fiber, meanwhile based on the number of modes propagate in the core are single mode fiber (SMF) and multimode (MM).

The optical fiber materials are classified into dispersive substance. Besides, optical sources such as a light emitting diode (LED) and a light amplification by stimulated emission of radiation (LASER) may produce multi optical frequencies. Therefore, the optical signals propagate along the fiber delivered into several refractive indices distributed inhomogeneously in the core. With this phenomenon, we may use the effective refractive index parameter (n_{eff}) in spite of an absolute value. Some methods to determine the effective refractive indices have been analyzed by Chiang [4]. In physics, it may be defined that refractive index describe as the comparison between velocity of light in vacuum (c) and velocity of light in a medium (v), Eq. (2). Meanwhile v is also define as the optical waves travel during a time of t_f (s) and at distance of L (m), Eq. (3). By substituting Eq. (3) into (2) and eliminating v, the one-trip delay of optical fiber can be derived as the result in Eq.(4).

$$n_{eff} = \frac{c}{v} \tag{2}$$

$$v = \frac{L}{t_f} \tag{3}$$

$$t_f = \frac{n_{eff}L}{c} \tag{4}$$

The MAC protocol requires one round trip delay.

An optical transmitter comprises of the optical source and a driver circuitry. The delays in the optical transmitter are due to the conversion process from electronic state into photonic state in the converter components. The accumulation transient time of the components are usually stated in rise/fall time characteristic of the driving circuit. The output light of LED is photons generated by spontaneous recombination. The modulation speeds of LED are restricted by two main factors, extrinsic and intrinsic [2]. The extrinsic factor is the junction capacitance of the diode. Together with resistance, it affects to a characteristic RC time constant. While the intrinsic factor stems from the charge storage and diffusion capacitance of a p-n junction under forward bias. The recombination lifetime in junction materials plays a key role in affecting the modulation bandwidth. In general, the modulation bandwidth of an LED relies on the device configuration, doping level in the active layer, the lifetime of the injected carriers, and parasitic capacitance and resistance in the circuit. The bandwidth more than 1 GHz can be achieved now.

The semiconductor laser type which uncooled, low cost, and popular used for application in short distance network such as local area network (LAN), storage area network (SAN) is the vertical-cavity surface-emitting laser (VCSEL). The VCSEL is characterized by a complex interplay between optical, electrical, and thermal effects [14]. The maximum modulation bandwidth of the laser is limited by damping effects. The capability of modulation speed is up to 20 GHz. The VCSEL is suitable for outdoor usage for its heat resistance environment. A demonstration has been carried out at wavelength of 850 nm and the temperature of 85^{o}C for 2-km MMF shows full WLAN operation at 3,85 Mbps, with no observable degradation [10].

The optoeletronic device at receiver which is assigned to reconverts the photons into the electrons is a photodetector. Commonly, two types of photodetector that suitable for optical fiber communication, *p-i-n* (positive-intrinsic negative) photodiode and avalanche photodiode (APD). They operate at a certain wavelength region, such as around 850 nm, or 1300-1600 nm. The pin photodiode is constructed by pn junction diode in which an undoped *i*-region is inserted between p$^+$ and n$^+$ regions. A trade off must be taken for certain purpose design, for example for high response speed the depletion layer should be small and for high quantum efficiency, or responsivity, the depletion layer width should be large. There is no internal optical gain. By careful choice material parameter and device design, very large bandwidth can be obtained. The response speed depends on the carrier diffusion time in which creates a long tail pulse for ~1-10 ns. Therefore, this type is fit for low speed and short-distance link. On the other hand, the APD is built by many different type layers with specific purposes to generate the internal gain. For example, the separate absorption and multiplication (SAM) APD [9]. These configurations combine low leakage with sensitivity at long wavelengths. Due to some weakness arised from this type, further development construction created into next generation, called separate-absorption graded-multiplication (SAGM)-APD. For high-speed operation, usually the rise time of 100-200 ps and fall time ~10-100 ns can be achieved. Therefore, the APD makes it more suitable applied for high-speed operation and long-distance link.

3. Analysis of Delay-Bound

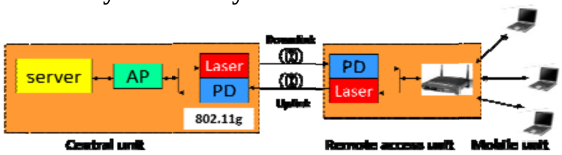


Figure 4. Model of Hybrid Link IEEE802.11g WLAN-over-Fiber [5].

Table 3. Parameters of Delays

Transceiver Delay (μs)		Air Propagation Delay (μs)	
OPT1[6]	1.6	A	1
OPT2[7]	1.25	B	0.1
		C	0.01

We use a hybrid link model as in Fig. 4 with MAC protocol of IEEE 802.11g, focusing on the optical segment that is a laser, an optical fiber and a photodiode (PD), and the direction of signal flow may downlink or uplink is similar. The total delays are the contribution of optical transceiver D_{opt}, optical fiber (t_f) and air propagation (D_{air}). The optical transceiver delay data can be found in Table 3. These data are based on the equipment data sheet. The optical fiber delay uses Eq. (4) with the setting of n_{eff} = 1.5, as an example. This equation may valid for SM or MM fiber type. The air propagation delays are varied, since the distance between AP and MU are varied as well. Next, the roundtrip total delay can be expressed as

(5)

Fig. 5 shows the results of the delay calculation.

From Fig. 5, it can be seen that the delay increase linearly as the distance longer. As stated that the delay limit based on 802.11g is 10 μs. To fulfill this requirement, we cut off the fiber length that resulting allowable delay. The shortest fiber length is 400 m meanwhile the longest fiber used is around 700 m only. For all the case, choosing shorter delay on optical transceiver shows a significant difference for the air delay, about 100 m. The closeness lines between (OPT1,B) and (OPT1,C) and between (OPT2,B) and (OPT2,C) indicate that the air propagation delay change of 0,1 μs give very small contribution to the total delay. Therefore, we can abandon the air propagation delay in this network.

4. Discussion

Our calculation on the delays are confirmed with the results of [5] but not for [12]. However, both papers did not concern with the optical transceiver. This made the length of fiber longer than our calculation. Those papers are published by an intensive research group specialized on Radio or WLAN over Fiber which has never consider the optical transceiver take into account. Therefore, this paper is the first attempt inserting the optical transceiver into the total delay calculation. For different transceiver specification, in this case the fiber length difference is around 100 m.

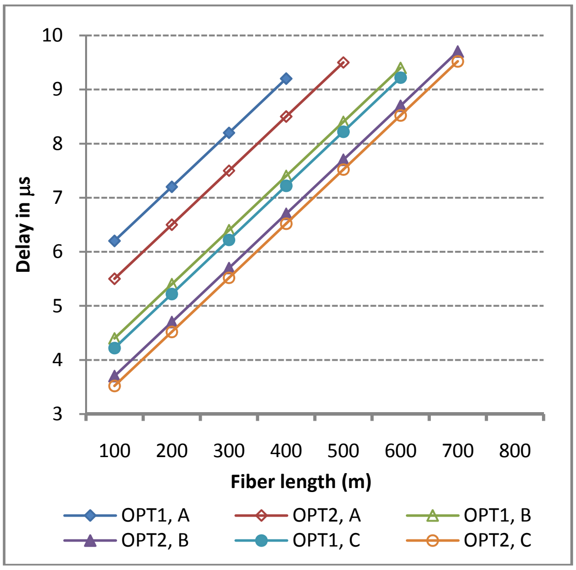


Figure 5. The length of Fiber Under SIFS

Comparing the hybrid network with an optical access network (OAN) in the term of fiber length, there will be the unbalance situation since the optical signal may travel tens of kilometer and can carry information in the order of hundreds Mbps in the later network. In my opinion, it seems due to the usage of different protocols. The OAN usually applies synchronous digital hierarchy (SDH) protocol, meanwhile the hybrid network utilizes CSMA/CA which the procedure is very tight constraint in timing. If we want to extend the coverage area of the hybrid network with maintaining the network throughput, we may not rely on the physical layer re-engineering any longer.

Some alternative solutions offered are re-engineering the data link logic layer MAC or to combine with a free space optical communication (FSOC) system which the refractive index of the air is smaller than the optical fiber to enable the longer distance coverage.

5. Conclusion

The insertion of optical transceiver, introducing new precision delay bound calculation, in this case the result show around 100 m of fiber length difference. The delay bounds are affected dominantly by the optical transceiver and the optical fiber.

References:

[1] Al-Raweshidy, H, Komaki, S. 2002, Radio over Fiber Technologies for Mobile Communications Networks, Artech House.

[2] Bhattacharya, P., 1994, Semiconductor Optoelectronic Devices, Prentice Hall, New Jersey, USA.

[3] Chia, M.Y.W., Luo, B., Yee, M.L., Hao, J.Z., 2004, Hybrid Radio over Fiber Wireless LANs and Ethernet Networks, Optical Society of America.

[4] Chiang, K.S., 1986, Analysis of Optical Fibers by the Effective-index Method, J. of Applied Optics, Vol. 25 No. 3.

[5] Das, A., Mjeku, M., Nkasah, A., Gomes, N. J,, 2007, Effects on IEEE 802.11 MAC Throughput in Wireless LAN over Fiber Systems, J. of Lightwave Technology, Vol. 25, No. 11.

[6] Data sheet, *Single Mode Fiber Optic Transceiver E-FRL-MC03 USER'S GUIDE*, Transition Network, Minneapolis, USA

[7] Data sheet, 2007.*Sun RoHS Compliant 10 Gb/s 850nm XFP Optical Transceiver*, Sun Microsystem,

[8] Fabbri, M., Faccin, P., 2007, Radio over Fiber Technologies and Systems: New Opportunities, ICTON

[9] Finneran, M.F., 2008, *Voice Over WLANs*, Elsevier Inc, USA.

[10] Hartmann, P., Qian, X., Penty, R. V., White, I. H., 2005, High Performance Uncooled 1.3 µm VCSELs for RF-over-Fiber Applications, Conference on Lasers & Electro-Optics (CLEO).

[11] Ilyas, M., Ahsan, S., 2005. *Handbook of Wireless Local Area Networks*, Taylor & Francis.

[12] Mjeku, M., Gomes, N. J., 2008. Analysis of the Request to Send/Clear to Send Exchange in WLAN over Fiber Network, J. of Lightwave Technology, Vol. 26, No 15.

[13] Part 11, 2007, Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, IEEE Standard 802.11.

[14] Piprek, J., 2005, Optoelectronic Devices

[15] Advanced: Simulation and Analysis, Springer Science+Business Media, Inc, USA.

MOBILE BANKING: SAFE, AT LEAST FOR NOW

Bhavna C.K. Nathani[1], Charles Lim, M.Sc. ECSA ECSP CEH CEI[2]

[1,2]Binus University International
[1]bhav_1989@hotmail.com, [2]clim@jwc.binus.ac.id

Abstract

Mobile banking is eye-catching because it is a very convenient approach to perform remote banking. But there are security problems with its current implementation. These security problems include feature like, not encrypted text messages, which are stored on phone, authentication of User ID / Password of bank's customer, etc. This paper discusses how mobile banking works, to which extent is mobile banking safe, what are the upcoming challenges and security problems that should be overcome and how to overcome it. This paper allows readers to fully understand the development of banking technologies on their daily transactions and allow readers to the other side of the wall about mobile banking.

Keywords: Security, Workflow, WAP, SMSC, Protection

1. Introduction

People love to use technology in every field they can. Every region is using this new trend to attract more users. In this technological era, banks are also in the same competition of drawing more customers to use their services. After the invention of e-banking or electronic banking, banks are pushing through the notion of Mobile banking to their clients. Figure 1 is a graphical representation of Asia Pacific mobile device subscribers [8]. The Asia Pacific region holds five of the world's ten major mobile markets. China and India are the number one and two respectively and Indonesia, Japan and Pakistan, the sixth, eighth and tenth largest.

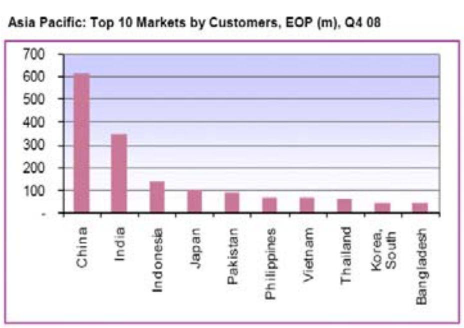


Figure 1

The background for choosing this subject is because there are many controversies regarding mobile banking, which is very interesting to discuss. Research methods include browsing certain websites and journals.

To begin with, Mobile Banking is a concept of making payments, checking account balance, and performing money transactions via mobile phones. This concept is widely accepted and bank customers are looking for banks that provide such facilities on their accounts. According to the German mobile operation Mobilcom, mobile banking will be the most successful application for the future of mobile technology. This success can be seen in Smartphone due to their capability of providing services anywhere and anytime [10]. As the era is now shifting towards mobile banking, there is a challenge for banks to expand this technology in terms of its safety and usefulness.

In this paper, we investigate the safety and security problems of current implementations of mobile banking, its upcoming challenges and how to overcome the problems. In section 2, we provide an overview on the workflow of mobile banking. In section 3, we describe the extent of safety of mobile banking. Section 4 and 5 discusses the current issues of mobile banking and how to achieve maximum security. Section 6, depict the future challenges of mobile banking and ways to overcome it. In section 7, we conclude our research paper.

2. Workflow of Mobile Banking

Mobile banking can be base on Short Message Service or WAP gateway. A mobile banking system basing on Short Message Service comprises the SMSC, the mobile banking Server, Message Processing Server and a compatible mobile phones, as shown in Figure 2 [7]. A Short Message Peer-to-Peer (SMPP) interface is use between the mobile banking server and the databases of the bank systems, through which the financial organizations talk with the SMSC.

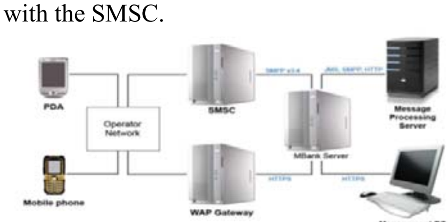


Figure 2

The SMSC is connected with the mobile banking modules via SMPP. To transmit SMS to mobile phone users, the operator network login to the SMSC by bind transmitter command and to receive short messages from the mobile phone, they login to the SMSC by bind receiver command. The messages include the accounts assigned to the network and related password.

After a mobile subscriber sent a short message to the SMSC, the SMSC routes it to the very mobile banking network operator. This operator authenticates the mobile subscriber, and decides if it has access rights. Short messages are encrypted and deencrypted on the STK card and the mobile banking network [11].

Enabling transactional capabilities, WAP is superseding earlier banking applications based on SMS, the two-way messaging protocol that catered for basic information-led mobile banking. WAP allows the use of Wireless Mark-up Language (WML), a stripped down version of HTML, to build platform-independent wireless applications. In a typical WAP solution, data packets from a wireless device pass along a wireless network in WML format to WAP server/gateway. The essential data is then reconfigured and passed to a standard HTML-capable web server. This works conversely if HTML packets must pass through a WAP server/gateway to reach a wireless device [9].

Similar to a PC requiring an internet browser installed in order to access content online, a mobile device requires a WAP browser installed in order to access information on WAP sites. WAP requires intermediary interconnects such as WAP proxies and gateways to bridge communication between the carrier's wireless network and the WWW internet [9].

3. Is mobile banking safe?

If you are not yet using your phone to pay your bills, check your account and doing monetary transaction account, you will soon. "There's little doubt that the era of mobile banking is coming," says Mark Schwanhausser, an analyst at Javelin Strategy and Research in Pleasanton, Calif [5].

That begs the question: How safe is it? With all the tech-savvy crooks and identity thieves lurking about, is it really a good idea to have our precious financial information floating around the airwaves or residing on a piece of gear that we could easily lose? [5].

3.1 Safety extent of mobile banking

1) **Banks offers three ways to connect our phone with them.** Each bank has its own way of entering into this new trend, and each phone is equipped with its own set of limitations. The most common way is through internet surfing options using the web browser on phones. It is the same as you log on to your bank account via computer [6]. Other banks allow you to access your account through special software you have downloaded onto your phone. In addition, a few banks offer balance information and other features through Short Message Service. However, these conveniences tend to make clients forget to keep their identity protection plan [6].
2) **Account numbers are not usually transmitted over wireless cell phone connections.** There is a major difference between accessing your account online and logging in from your phone [5]. In mobile banking, you will not be asked for your account information and after accessing your account your account number will never be visible making it less likely (although possible) to be compromised. This might reduce some of the problems of identity theft.
3) **Most banks use encrypted servers for all wireless transactions.** Banks are not taking the risk of using unencrypted servers. If handled correctly, wireless connections can be safer than wired connections [5]. Encrypted servers add to your safety but it still does not guarantee it.

According to researchers, all information transmitted between servers and mobile devices is encrypted as with regular online banking. As a result, there is less chances of fraud to happen [6]. Less chance does not mean that it is 100% safe. Some banking Institutions, agree to all of customer's mobile fraud losses. While other banks, claim that it is customers choice to use mobile banking, so any losses is not covered by the bank.

4. Current mobile banking issues

In this section, we will take a glance of the current security issues of mobile banking using SMS and WAP technology and other neglected problems

4.1 Security problems with SMS

The initial plan for SMS was intended for uses to send messages across the network. Text encryption, end-to-end security, non-repudiation and authentication were neglected during the design of SMS system. Below are some security problems of using SMS.

4.1.1 Modifying Originator's Address

SMS spoofing is an attack in which the sender of the message can be set by replacing the originating mobile number or Sender ID with alphanumeric text. Spoofing can be legal such as setting the company name from which the message is being sent and illegal such as impersonating another person or company [10].

4.1.2 SMS Encryption

End to end to encryption in SMS is not available at present. The default data format is in plain text. Only one encryption that occurs during the transmission, which is the encryption between the mobile station and base transceiver station. The encryption algorithm that is used is A5, which is verified to be susceptible [10].

4.2 Security problems with WAP technology

WAP technology that is used in mobile banking has been proven secure, but there also exist certain loopholes, which can cause security issues. These loopholes are:

- End to end encryption between customer and bank server, customer and the gateway and bank server and gateway, does not exist.
- Public key cryptosystems that is offered by Wireless Transport Layer Security (WTLS) are not powerful enough to meet the present security requirements of WAP technology [2]. Considering the low processing power of mobile phones, the key sizes have been restricted.
- Anonymous key exchange presented by WTLS handshake is not considered secure. This is because the clients and servers are not authenticated [2].

5. Achieving Maximum Security

Achieving maximum security in mobile banking should not and cannot only be done by banks. Customers or users should also be a part of it to maximize the safety.

5.1 From banking point of view

To resolve the issue of the absence of end-to-end encryption in WAP technology, bank server could have their own Access Point Name (APN) in any of the mobile phone networks. This APN would be the WAP gateway for the bank. As a result, customers would be directly connected to the bank without a third party interrupting the communication. Banks should also disallow the option of handshaking in key exchange.

Messages that are sent through short message service should be encrypted by a secret one-time password, which only shared between the user and the bank server [3]. The confidentiality of the password depends on how strong is the algorithm that is implemented. It is assumed that the list of passwords is only known by the authorized users.

For mobile phone users to be authenticated by the receivers, they must provide their authentication details to the receiver. This process involve the verification of the message PIN with the receiver stored PIN [3]. This PIN was selected by user when they register for mobile banking.

The bank server does not generate the same one time password more than once and only for one single user. As a result, user cannot deny not sending the message because only that specific password and sequence number is used to encrypt the message. If the bank server can use the same password pair to decrypt the message, then they are sure that the authorized user must have sent it [3].

5.2 From users point of view

According to Onoo W. Purbo, an internet expert, most mobile banking security issues come from the non-technical aspect. As users of mobile banking, we should expect much from technology. We as humans can also do our part in protecting our own account from theft. Below are some ways that it could be done:

1) Protect your password
- Choose a good password:
  - Password must consist of 6 to 8 numeric digits.
  - Do NOT use obvious passwords, such as your birthdays.
  - Do NOT use the same password for other Internet sites or financial services.
  - Wherever possible the customers should think of a password that means something to them, but not to others.
- Handle your password:
  - Do NOT disclose your password to anyone.
  - Do NOT allow anyone else to use your password.
  - Do NOT write down or record the password without disguising it, as this can easily lead to discovery and compromise.
  - Beware of common social engineering attacks, e.g. people pretending to be a member of the police or Bank staff to ask you to tell them your password.

2) Protect your mobile phone
- Set up password of your mobile phone
  - Usually mobile phones allow you to set up a PIN, which will be asked when you switch on your mobile phone. This can limit the use of your mobile phone from others in case it has been stolen or lost.
- Avoid sharing your mobile phone
  - Be careful in using your mobile phone to do Mobile Banking if it is likely to be shared with others.
  - Remove any temporary files that were stored in the phone memory during usage of mobile banking, as those files may contain information such as account numbers.
  - Clear the browsing history regularly.

3) Protect your personal information
- Maintain your confidential information
  - Do NOT share your password with anyone other than yourself.
  - Do NOT give away your personal information to anyone especially over the phone or to a web site unless you have verified the recipient.

- Keep documents safe
    - If possible, destroy or keep safe any receipts, statements and any bills that contain personal information.

6. Future challenges of mobile banking and ways to overcome it

Though there are few problems related to mobile banking and it is still on the border line of safe zone does not mean that it is not going to go away from its safety zone. Some future challenges of mobile banking are describe below:

1) One challenge of mobile banking is to avoid pharming. Pharming is a hacker's attack, which aims to redirect a particular website to another fake website. Pharming can be carried out by changing the host file on the victim's computer or by exploiting the vulnerability of DNS server software [1].
   One of the proposed solutions is to make sure that mobile phone users are using HTTPS, which is a secure web connection, to access sensitive websites such as banking and only accept valid public key certificates issued by trusted sources. Unknown certificate should not be accepted every time for crucial businesses [1].
2) According to experts study, certain upcoming malware are not recognizable by antivirus program. These malware waits for mobile phone users to login to their bank account and steal information. In 70% cases, this malware can successfully by passes anti virus detection [4]. Therefore, even antivirus programs with up-to-date signatures are mainly unable to detect these malwares. Perhaps, experts should raise the bar of malware research.

7. Conclusion

For Service providers, mobile banking offers the next surest way to achieve growth. They are also using the complexity of their supported mobile phones banking services to attract new customers and retain old ones because of the fact that, one day mobile phones will replace the use of fixed telephone lines. Mobile banking is not 100% safe. We should not only expect banks to provide security, as users, we should also know the safety measure that has to taken.

In all sincerity, it is too early to say that mobile banking will offer all the needed conveniences that it promises without increasing customer's risk of theft. So, if mobile banking sounds like something you are interested in, it might be worth trying. However, prior to signing up, ask both your bank and your cell phone provider specific questions about encryption and identity theft protection. It is your identity that will be at stake and it is your job to protect it.

References

[1] Buyung., *Bahaya Baru Online Banking*, Available from: http://buyung30.wordpress.com/2009/02/12/bahaya-baru-online-banking/; accessed 5th October 2009.

[2] Burak Bayoglu, 2004, *Performance evaluation of WTLS handshake protocol using RAS and elliptic curve cryptosystems*.

[3] Chikomo. K., 2008, *Security of Mobile Banking*.

[4] Ecommerce Journal, *Banking malware Zeus can successfully bypasses anti-virus detection*, Available from: Banking malware Zeus sucessfully bypasses anti-virus detection; accessed 6th October 2009.

[5] Howard, N., *Is mobile banking safe?,* Available from: http://www.bankrate.com/finance/personal-finance/is-mobile-banking-safe-1.aspx; accessed 3rd October 2009.

[6] N.A., *Mobile Banking: Pros and Cons*, Available from: http://www.identitytheft.com/index.php/article/mobile_banking; accessed 2nd October 2009.

[7] Sunet Technology., *Mobile banking*. , Available from: www.sunet.uz/mbanking.html; accessed 2nd October 2009.

[8] TowerGroup., *Indonesia Overtakes Japan to Take Third Place in Asia Subscriber Rankings*, Available from: http://www.itu.int/ITU-D/ict/newslog/Indonesia+Overtakes+Japan+To+Take+Third+Place+In+Asia+Subscriber+Rankings.aspx; accessed 28 September 2009.

[9] Van, L ., *Use of WAP technology in Mobile Banking*, Available from: http://www.mobile-financial.com/node/2520/Use-of-WAP-technology-in-Mobile-Banking; accessed 3rd October 2009.

[10] Wikipedia, *Mobile Banking*, Available from: http://en.wikipedia.org/wiki/Mobile_bankin g ; accessed 28th September 2009.

[11] ZTE corporation, *Mobile Banking Based on Short Message Service*, Available from: http://wwwen.zte.com.cn/endata/magazine/ztetechnologies/2001year/no10/articles/200312/t20031212_161135.html; accessed 3rd October 2009.

DESIGN AND IMPLEMENTATION OF SMS GATEWAY AS A LEARNING SUPPORT TOOL

Burhanuddin Dirgantoro

Fakultas Elektro & Komunikasi, Institut Teknologi Telkom
brh@ittelkom.ac.id

Abstract

Learning processes need some simple test (”quist”) to make feedback for students among the material course conveying. This activity should be given regularly, but because of the business of the lecturer, it could be impracticable, especially in the assessment credit process. This research has been tried to make that activicy become simpler. The assessment process can be done by computer. The computer will jugde the assessment value by matching the quis answers with the correct answers. The lecturers need not check the answers by themself. The answers can be sent by the students via SMS, which will be collect by an SMS Gateway. This SMS Gateway will process the answers then, and give the correct assessment credit value that is keep in the database. The SMS Gateway can be able to send the assignment value to the students, too, to be a feedback to the students. This research has been tested in the laboratory and in the class room. The test give a result that the system can be used to support learning process, especially in simple test activity. Some development must be done before it can be launch to the learning society, including the interface, the question databases, and some other aspects to give the user simplicity dan friendliness of using this tool.

Keyword: SMS gateway, learning tool, automatic assessment

1. Introduction

Class-room (perkuliahan) activity is one of the core process in the learning process. This activity consists of : material course conveying (pemberian materi perkuliahan), course material test (pengujian materi) and assessment credit value process (pemberian nilai). Material course conveying can be given along the semester period, following with dan akhir semester).

Although giving some homework and/or simple tests, the lecturer often doesn’t have much time to check the answers, because of the time needed to do that activity. On the other hand, people is more and more using handphone as their communication tools. Almost all the people, including students, use design dan implement a learning process tool, especially in the answer check and assessment process for simple tests. The implementation of the system will take profit from using SMS gateway. Besides, this tool can be used as a communication tool also.

2. Learning process and the SMS Gateway

2.1. The Learning Process

The initial process of learning process in the universities is the course registration process, that must be done by all students. Then, the students will be given a suitable material course, depend on the course taken. In this process, it is familiar for students to take some practical task and job in the laboratory. The lecturers, besides, will also give the students some homework and simple test among their material course conveying, to consider the students’ understanding about the course material some homework and simple test in the class room. In the middle and the end of the semester, it is familiar to do the middle test and final test to make the final value / credit assignment.(Pemberian materi pada umumnya dilakukan sepanjang semester, yang diikuti dengan pengujian materi dalam bentuk Pekerjaan Rumah dan kuis pada beberapa pertemuan serta ujian tengah

handphone as their daily device at this moment. So, it is an opportunity for lecturers to take benefit from this situation for using handphone as their tool for learning process. The handphone can be used as an aid to transmit the answers to the lecturer’s computer, and then the computer will finish the rest.

This reseach has an aim to and give them some feedback. Sometimes, the lecturers give them discussions, special tasks and presentations.

The leaning process in the class room will be terminated with tests, to give the students credits for their course. The assessment process can be in some forms, including spoken and written test, working task (take home test, paper work and/or presentation), or combination of those methods. The learning process can be drawn in Figure 1.

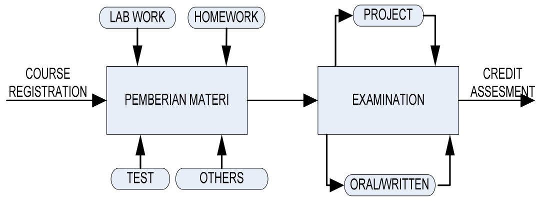


Figure 1. The Learning Process Diagram

2.2. SMS Gateway

SMS gateway is one of cellular communication-based technologies using SMS as communication media. SMS can be used effectively to support learning process, by developing an SMS gateway as a transmitting and receiving agent.

SMS Gateway works as folow :

a. A computer is installed with a program that congrol the handphone connected to the computer to receive and transmit SMSes.
b. An embedded program is developed to process the SMSes, and then transmit tSMSes to respon the message.
c. The embedded program can be able to make a text file, that can be loaded by spreadsheet like MS-Excel. So, this file can be processed by lecturers easily.

The structure of the computer dan handphone integration can be drawn in Figure 2.

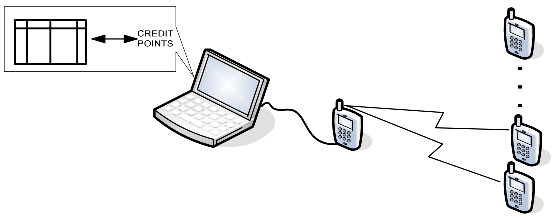


Figure 2. SMS Gateway Structure

3. System Design and Implementation

3.1. SMS Gateway Subsystem

SMS gateway subsystem can be developed from a *low-end* laptop, with specification as Table 1. The modem is a handphone from Nokia, and written in Table 2. The controller program is a freeware program from gammu.org, as indicate in Table 3.

Table 1. Spesificasion of Laptop Used.

No	Spesification	Value
1	Manufacture/type	Asus /EeePC 900
2	Processor	Intel Celeron 900 MHz
3	Memory	2 GB
4	Harddisk	20 GB(Flash memory)

Table 2. Handphone Used as Modem

No	Spesification	Value
1	Manufacture /tipe	Nokia 6030
2	Manufacture Year	2006
3	Data cable	Mini USB, CA-45

Table 3. Controller Program

No	Spesification	Value
1	Name	Gammu
2	Version	1.09.00, 6 Des 2006
3	Developer	Marcin Wiacek, and team
3	Operating System	Windows XP SP 2
4	Database & server	MySQL, *xampp bundle*

SMS gateway subsystem will work in the steps as follows :

a. Connect the data cable from computer to handphone, and install the driver program.
b. Activate MySQL server. In this research, we use xampp bundle, in Windows, that consists of Apache server, and MySQL database. Other programs in that packet are disabled.
c. Activate gammu program with command :
   Gammu –smsd MYSQL smsdrc
d. The database is ready to be processed.

This command will run the program, so that the handphone will connect to the database. Smsdrc is a configuration file consists of data needed to be loaded by the program, including database' name.

This program will be run along the system usage, that will receive SMS from and to the handphone.

3.2. SMS Tranceiver and Processor Subsystem

This subsystem is a program that made to handle incoming SMS and transfer to the database. The program was built in PHP script, that is familiar with the MySQL database. This program consists of Data Input menu, Data Watching, SMS Sending dan Test Answer Awaiting.

Data Input menu is used to input students data, and the answer keys. Data Wacthing is needed to see the content of the tables in database, that are : Inbox, Outbox, SentItems and Credit Test. There is a function menu to convert from Credit Test to Test.txt file, too.

SMS Sending menu, as other SMS gateway system, was created to send the SMS to other handphone. We can choose to send individual number, group students, to one or more of the group.

Test Asnwer Awaiting is the central point of this system. This subprogram can be run continuously, to receive the SMS from the students. Not all of the SMS sending by students are to answer the test given for them, but they sent question SMS or confirmation SMS too.

This menu works as follow:

a. Program is run, and wait for SMS sent by students,
b. If an SMS is came to the handphone, this SMS will be tranfer to Inbox table, and then be delivered to Inprocess table.
c. The subprogram will process the SMS, to see the format inside. If there is a word 'answer', the SMS will be processed as a test answer.
d. The subprogram will process the answer, and match the message with the correct answer given in the database. From this process, the system will has a credit value for that student.

e. The program will transfer the credit value to the database, and send SMS to the student about this value.
f. If there are some messages sent from the same student, the system will send message that the student had sent a messege before.

4. System Testing

4.1. System testing results with simulation data

The system was tested with simulation data, and taken with scenario in 3 ways : data in the same format as directed, data in the same format as directed with little error, data outside in the format.

a. The simulation data in the same format as directed

The answers in SMS were made in this format :

answer(space)IDNumber/Name/Question Code/answers(devided by comma)

This test showed that the results were fit with initial plan. Some notes can be remembered are written in Table 3.

Tabel 3. Notes from Simulation Data Test

No	Message	Result
1	Number of the answers was equal with the planned answer	Credit value was the same as the number of question
2	A part of the answers was coorect	Credit value was suitable with the correct answers
3	Number of the answers was less than the planned answer	Credit value was suitable with the correct answers, program wasn't error
4	A part of the answers was emptied between the commas	The answers were considered false, program wasn't error

In this case, program was worked as planned dan gave the credit value as expected.

b. The simulation data followed the given format, with a little error

The test ini this scenarion gave results as written in Table 4.

Table 4. Results for Little Erroneous Format Data

No	Type of error	Result	Solution
1	Capital and lower case difference	Detection / assessment error	Converse the code & answers to lower case
2	Spaces in the answers	Detection / assessment error	Spaces cancellation
3	Quotation mark in the answer	Program error / stopped	Quotation mark cancellation

Solution had done with recode and rerun the program. The results showed that the program can worked as planned.

c. The simulation format data were not follow the planned format

This test used message that were not in suitable format. As a result, the assessment was not suit with the plan. Some notes for this test are written in Table 5.

Table 5. Unfitted Format in The Answers

No	Message	Result
1	The unfit order	The credit value were not suitable
2	No space between word 'answer' and IDNumber	Message was not processed
3	No code	Credit value : 0
4	No comma between answers	Credit value : 0
5	There was another character in the end of the answers	The last answer was considered wrong
6	The answer devider were not commas	Credit value : 0

This test results can be understood as the formats were not recognized as planned.

4.2. Field testing

Field tests had done in the class romms. The test scenarios were done as follow ::

a. The lecturer made some groups of test, with different code for each test.
b. The seats of student were arranged so that they had a little chance to cooperate each other.
c. Studens must carried out the test in a limited time allocation.
d. In the end of the test, the students were given some minutes to write message in the right format with their own handphone and sent their SMS to the SMS gateway.
e. The SMS gateway appraised the message, which compared the messase with the correct answer in the database. Then, the system sent the credit value back to the students.

The tests were done in 2 classes of Algorithm course, with 64 and 70 srtudents each. The results from that tests are written in Table 6.

Table 6. Results from Classroom Tests

No	Case happened	Amount
1	Major format errors	2
2	Minor format errors	5
3	SMS delivered more than once	7

As a addition, in the evaluation process of answer matching, there was a mistake in the correct answer writing to the database. In this case, the lecturer corrected the answer and then rerun the program once again. This correction activity was done with PHPMyadmin program.

The result of the test were kept in the database, that can be converted to file "quis.txt". This file can be called by Ms-Excel, and can be process to get information needed easyly.

5. Conclusion dan Suggestion

5.1. Conclusion

a. The system can be worked as initial plan.
b. The system will give the correct credit value when the answer in database is correct and the format sent by students are correct too.

5.2. Suggestion

a. Development of the interface, so that lecturers can used this system easier.
b. Development of addendum program, such as answer analyzer to give students more feed back about the wrong answers.
c. Development of question test databases, so that lecturer can make test sets easier to avoid student's cooperation in doing the test.

List of Divining Manuals

[1] www.gammu.org
[2] www.cihar.com
[3] www.php_freak.com
[4] www.zend.com

EASTERN CYBERLAW EXPOSED: THE PORT SCANNING WAY

David Allister[1], Michael[2], Charles Lim[3]

[1,2,3]Swiss German University
[1]**david.allister@student.sgu.ac.id**, [2]**michael@student.sgu.ac.id**, [3]**charles.lim@sgu.ac.id**

Abstract

The security of information is vital. Based on the vulnerability assessment exercise, there are many reports about the susceptibility of information leak within a network. Nowadays, port scanning is a common method to evaluate the vulnerabilities and weaknesses in a computer system and network. However, as important as it may seem, some the ambiguity in port scanning may be regarded as illegal, based on the Indonesian law, which was passed 2008 about the protection of computer activities and computer usage in Indonesia. This law is known as the ITE law (Law of Information and Electronic Transaction) and it refers to Budapest Convention (Convention on Cybercrime).This paper discuss the legality of conducting port scanning in Indonesia as well as other countries in ASEAN.

Keywords: Port scanning, cyber law, Undang-Undang ITE, Budapest Convention

1. Introduction

The advance of networking technologies leads into the possibility among computers to communicate between each other. For computers to communicate with one another over the internet or across an organization, they must speak the same language. This language is referred to as a protocol, and port is the logical, not physical, component of protocol connection. A port itself needed to identifies the service that is running.

It is similar to a security guard going through the house and does the routine check to every door and window as the part of the house to check which ones are suppose to be open and which one are locked. So do the port scanning is method used by the security tester to determine what ports are open or in use on a system or network.

By using various tools the tester can send data to TCP or UDP ports one at a time. Based on the response received the port scan utility can determine if that port is in use. Using this information the tester can identify the vulnerabilities of the system and focus on the opened ports so that unauthorized user cannot attack these ports and try to exploit any weaknesses to gain access. In this case, port scanning look as noninvasive and nondestructive in nature and deem it legal. In some cases, there are also companies who suffer lose because this methodology was conducted and being misused by unauthorized party.

Thus, Indonesia as a based on law country has just setup a law in 2008 to protect the computer activities and computer use in Indonesia. This law also known as Undang-Undang ITE (Law of Information and Electronic Transaction). This paper discusses whether"is it actually legal to conduct port scanning in Indonesia? "And "what are actually the ethic code in performing this particular test? " Other than that, as a part of The Association of Southeast Asian Nations or ASEAN, this paper will also discover kind of cyber law is actually being used by other countries in ASEAN. Since, the law was setup and adjust based on the problem in each countries, within this paper we could discover which problem is the main concern from each of these ASEAN countries.

2. Port Scanning Overview

2.1 TCP/IP and UDP[1]

Let's start with the basic. First, we must know about the connection protocol. There are two types of connection protocol, TCP and UDP.

TCP or Transmission Control Protocol is a connection-oriented protocol which means the sender doesn't send any data to the destination target until the target acknowledge that it's listening to the sender. We can say in another word that the connection must be established before sending data. The process called three-way handshaking. The processes are:

- Host A send TCP packet with the SYN flag to host B
- Host B retrieve SYN flag from host A and then send the SYN from host A with an ACK flag from host B. this flag usually called SYN-ACK.
- Host A retrieve SYN-ACK flag from host B and host A send back the TCP packet with the ACK flag.

There are six types of TCP flags. There are:

- SYN flag - The Synthesis flag signifies the beginning of a session.
- ACK flag - The Acknowledgement flag a connection and is sent by a host after receiving SYN-ACK.
- PSH flag - The Push flag is used to deliver packet data directly to an application.

- URG flag- This flag is used to signify urgent data.
- RST flag- The Reset flag reset or drops the connection.
- FIN flag - The finish flag signifies that the connection is finished

The next protocol is UDP. UDP or User Diagram Protocol is a connectionless protocol which means the sender doesn't need to create a connection first to send the data. It just sends the data to the destination target. This method is fast but unreliable because using UDP, the sender doesn't need to verify whether the receiver is listening or ready to accept the packet also the receiver retrieve the packet or not.

2.2 Ports

Port is the logical component in a computer system that identifies the service that is running. Each service has its own port. In a computer there are 65.535 ports that can be used but there's only 1023 ports are considered well-known ports. Some of them are:

- TCP 20 and 21 (File Transfer Protocol, FTP)
- TCP 22 (Secure Shell, SSH)
- TCP 23 (Telnet)
- TCP 25 (Simple Mail Transfer Protocol, SMTP)
- TCP and UDP 53 (Domain Name System, DNS)
- UDP 69 (Trivial File Transfer Protocol, tftp)
- TCP 79 (finger)
- TCP 80 (Hypertext Transfer Protocol, HTTP)

2.3 Definition and Process

Port Scanning also referred as service scanning is a method of finding out which ports are being used in the system to find its vulnerability. The penetration tester and hackers use port scanning to scan all ports that are open in the system to find the system vulnerability.

There are seven types of port scans that is used:

- SYN scan - a packet with SYN flag send to the target. If the target replies with SYN/ACK, the attacker sends RST flag so the SYN scan is stopped. From that we can see if the target replay, it means that the port is opened.
- Connect scan - this method is a three-way handshaking. It's easily detected.
- Null scan - in null scan, all flag are closed. A closed port will replay with an RST so if no packet retrieved, best guest that the port is open.
- XMAS scan - in this type of scan, FIN,PSH, and URG are set. Close port will respond with RST flag.
- ACK scan - Hackers usually use ACK scan to pass the firewall. If the attacked port returns RST flag, the packet filter was fooled or there is no packet filtering device.
- FIN scan - a FIN packet send to target computer. If the reply is RST packet, it means that the port is closed.
- UDP scan - UDP packet is send to target computer. If the target send back ICMP "Port unreachable", the port is close.

2.4 Port Scanning Tools

There are so many tools for doing port scanning. The best tool yet is NMAP. NMAP originally written for Pharack magazine in 1997 by Fyodor. Its features are OS detection and fast multiple probe scanning.

The other tool is Nessus. Its developed in 1998. Like NMAP, Nessus is open source and can be downloaded from www.nessus.org.

Port scanners can be also used to conduct a ping sweep of a large network to find the host of the IP. The purpose is to see whether the IP is alive or not. Several tools can be use doing this ping sweep. They are FPing and Hping.

With FPing, you can ping multiple IP address simultaneously. But it has one disadvantages. It can't through the firewall. To do that, we can use HPing. HPing is similar to FPing with an advantages it can bypass firewall.

3. Port Scanning Legality in Indonesia

3.1 Undang-Undang ITE (Law of Information and Electronic Transaction) Background

The use of Information Technology, media, and communication has changed human behavior globally. The development of Information Technology and media has made the connection in the world borderless and has changed the economy, social, and culture rapidly. Information Technology nowadays could be seen in positive perspective as a media that contribute in increasing human prosperity and human civilization. Otherwise, in negative perspective it is also can be a supporting device to commit activities against the law.

Related to this issue, we have to look after the security and law certainty in the use of information technology, media, and communication in order to optimize the development of technology. Thus, there are three approaches to secure the cyberspace:

- Law Aspect
- Technology Aspect
- Social, Culture, and Ethic Aspect

In order to solve the security interference in implementing system based on electronic, the certainty law is needed to optimize the use of technology.

3.2 Related Points in UU-ITE

Undang-Undang ITE[2] consist of 13 chapters and 54 points, within these points and chapters there are some points that we could approach in applying port scanning. This is stated in Chapter VII about FORBIDDEN ACTIVITIES:

- **Point 30**
  (1) Every unauthorized people intentionally against the law with accessing Computer and/or Electronic System which is owned by other people in any kind of way.
  (2) Every unauthorized people intentionally against the law with accessing Computer and/or Electronic System which is owned by other people in any kind of way in order to get Electronic Information and/or Electronic Document.
  (3) Every unauthorized people intentionally against the law with accessing Computer and/or Electronic System which is owned by other people in any kind of way, breaking into, pass over, or penetrate the security system.

- **Point 31**
  (1) Every unauthorized people intentionally against the do interception or tapping into Electronic Information and/or Electronic Document in a Computer and/or Electronic System owned by other people.
  (2) Every unauthorized people intentionally against the do interception or tapping into Electronic Information and/or Electronic Document in a Computer and/or Electronic System owned by other people even does not make any changes and/or make any changes, make lost, and/or stopping Electronic Information and/or Electronic Document that currently being transmitted.
  (3) Except interception as it mentioned in point (1) and (2), the interception related with the need of law enforcement based on request by police department, attorney, and/or other institutional law that being stated in law.
  (4) The interception policy is based on Government Law.

- **Point 33**
  Every unauthorized people intentionally against the law to do activities that cause disturbance into Electronic System and/or cause the Electronic System work improperly.

Next are the future actions after doing port scanning. There are several possibilities actions such as attack on mail server (port 25), DOS attack (port 445), SQL attack (port 1433), virus injection, backdoor installation (Trojan), etc.

All those action already define in chapter VII on points 30, 31, and 33 as illegal access, interception, and cause computer system work improperly.

4. Port Scanning on other ASEAN Countries

4.1 Brunei Darusalam

Based on **"Computer Misuse Act 2007 of Brunei Darussalam"[3]** in Article 2 about illegal acces:

(1) Subject to subsection (2), any person who knowingly causes a computer to perform any function for the purpose of securing access without authority to any program or data held in any computer is guilty of an offence and liable on conviction to a fine not exceeding five thousand dollars, imprisonment for a term not exceeding two years or both and in the case of a second or subsequent conviction, to a fine not exceeding ten thousand dollars, imprisonment for a term not exceeding three years or both.
(2) If any damage is caused as a result of an offence under this section, the person convicted of the offence is liable to a fine not exceeding fifty thousand dollars, imprisonment for a term not exceeding seven years or both.
(3) For the purposes of this section, it is immaterial that the act in question was not directed at —
(a) any particular program or data;
(b) a program or data of any kind; or
(c) a program or data held in any particular computer.

About "**Access with intent to commit or facilitate commission of offence."**

(1) Any person who causes a computer to perform any function for the purpose of securing access to any program or data held in a computer with intent to commit an offence to which this section applies is guilty of an offence.
(2) This section applies to an offence involving property, fraud, dishonesty or which causes bodily harm and which is

punishable on conviction with imprisonment for a term of not less than two years.
(3) Any person guilty of an offence under this section is liable on conviction to a fine not exceeding fifty thousand dollars, imprisonment for a term not exceeding ten years or both.
(4) For the purposes of this section, it is immaterial whether —
(a) the access referred to in subsection (1) was authorised or unauthorised;
(b) the offence to which this section applies was committed at the same time when the access was secured or at any other time.

4.2 Malaysia

Based on Computer **"Malaysia Computer Crime Act 1997 act 563"[2]** about Unauthorized Access to Computer Material:

(1) A person shall be guilty of an offence if-

(a) he causes a computer to perform any function with intent to secure access to any program or data held in any computer;
(b) the access he intends to secure is unauthorised; and
(c) he knows at the time when he causes the computer to perform the function that that is the case.

(2) The intent a person has to have to commit an offence under this section need not be directed at-
(a) any particular program or data;
(b) a program or data of any particular kind; or
(c) a program or data held in any particular computer.

(3) A person guilty of an offence under this section shall on conviction be liable to a fine not exceeding fifty thousand ringgit or to imprisonment for a term not exceeding five years or to both.

Unauthorised access with intent to commit or facilitate commission of further offence

(1) A person shall be guilty of an offence under this section if he commits an offence referred to in section 3 with intent-

(a) to commit an offence involving fraud or dishonesty or which causes injury as defined in the Penal Code; or

(b) to facilitate the commission of such an offence whether by himself or by any other person.

(2) For the purposes of this section, it is immaterial whether the offence to which this section applies is to be committed at the same time when the unauthorised access is secured or on any future occasion.

(3) A person guilty of an offence under this section shall on conviction be liable to a fine not exceeding one hundred and fifty thousand ringgit or to imprisonment for a term not exceeding ten years or to both.

4.3 Myanmar

Based on **"The Union of Myanmar The State Peace and Development Council The Electronic Transactions Law (The State Peace and Development Council Law No. 5/2004)"[4].**

Chapter XII
Offences and Penalties

34. Whoever commits any of the following acts shall, on conviction be punished with imprisonment for a term which may extend to 5 years or with fine or with both:

(a) Sending, hacking, modifying, altering, destroying, stealing, or causing loss and damage to the electronic record, electronic data message, or the whole or part of the computer programme dishonestly;

(b) Intercepting of any communication within the computer network, using or giving access to any person of any fact in any communication without permission of the originator and the addressee;

(c) Communicating to any other person directly or indirectly with a security number, password or electronic signature of any person without permission or consent of such person;

(d) Creating, modifying or altering of information or distributing of information created, modified or altered by electronic technology to be detrimental to the interest of or to lower the dignity of any organization or any person.

38. Whoever attempts to commit any offence of this Law or conspires amounting to an offence or abets the commission of an offence shall be punished with the punishment provided for such offence in this Law.

4.4 Philippines

Based on "**Electronic Commerce Act of 2000"[5].**

SEC. 33. Penalties.
- The following Acts shall be penalized by fine and/or imprisonment, as follows:

(a) Hacking or cracking which refers to unauthorized access into or interference in a computer system/server or information and communication system; or any access in order to corrupt, alter, steal, or destroy using a computer or other similar information and communication devices, without the knowledge and consent of the owner of the computer or information and communications system, including the introduction of computer viruses and the like, resulting in the corruption, destruction, alteration, theft or loss of electronic data messages or electronic document shall be punished by a minimum fine of one hundred thousand pesos (P100,000.00) and a maximum commensurate to the damage incurred and a mandatory imprisonment of six (6) months to three (3) years;

4.5 Singapore

Based on **"Computer Misuse Act o Singapore"**. **Unauthorized use or interception of computer service.**

3. (1) Subject to subsection (2), any person who knowingly —
(a) secures access without authority to any computer for the purpose of obtaining, directly or indirectly, any computer service;
(b) intercepts or causes to be intercepted without authority, directly or indirectly, any function of a computer by means of an electro-magnetic, acoustic, mechanical or other device; or
(c) uses or causes to be used, directly or indirectly, the computer or any other device for the purpose of committing an offence under paragraph (*a*) or (*b*), shall be guilty of an offence and shall be liable on conviction to a fine not exceeding $10,000 or to imprisonment for a term not exceeding 3 years or to both and, in the case of a second or subsequent conviction, to a fine not exceeding $20,000 or to imprisonment for a term not exceeding 5 years or to both.

(2) If any damage is caused as a result of an offence under this section, a person convicted of the offence shall be liable to a fine not exceeding $50,000 or to imprisonment for a term not exceeding 7 years or to both.

(3) For the purposes of this section, it is immaterial that the unauthorised access or interception is not directed at —
(a) any particular program or data;
(b) a program or data of any kind; or
(c) a program or data held in any particular computer.

Enhanced punishment for offences involving protected computers

9. (1) Where access to any protected computer is obtained in the course of the commission of an offence under section 3, 5, 6 or 7, the person convicted of such an offence shall, in lieu of the punishment prescribed in those sections, be liable on conviction to a fine not exceeding $100,000 or to imprisonment for a term not exceeding 20 years or to both.

(2) For the purposes of subsection (1), a computer shall be treated as a "protected computer" if the person committing the offence knew, or ought reasonably to have known, that the computer or program or data is used directly in connection with or necessary for —
(a) the security, defence or international relations of Singapore;
(b) the existence or identity of a confidential source of information relating to the enforcement of a criminal law;
(c) the provision of services directly related to communications infrastructure, banking and financial services, public utilities, public transportation or public key infrastructure; or
(d) the protection of public safety including systems related to essential emergency services such as police, civil defence and medical services.

(3) For the purposes of any prosecution under this section, it shall be presumed, until the contrary is proved, that the accused has the requisite knowledge referred to in subsection (2) if there is, in respect of the computer, program or data, an electronic or other warning exhibited to the accused stating that unauthorised access to that computer, program or data attracts an enhanced penalty under this section.

Based on **"Electronic Transactions Act of Singapore" [6]**

Obligation of confidentiality
48. —(1) Except for the purposes of this Act or for any prosecution for an offence under any written law or pursuant to an order of court, no person who has, pursuant to any powers conferred under this Part, obtained access to any electronic record, book, register, correspondence, information, document or other material shall disclose such electronic record, book, register, correspondence, information, document or other material to any other person.

4.6 Thailand

Based on **"Computer Crime Act B.E 2550 (2007)"[8]**

Section 5.
Any person illegally accessing a computer system for which a specific access prevention measure that is not intended for their own use is available shall be subject to imprisonment for no longer than six months or a fine of not more than ten thousand baht or both.

4.7 Vietnam

Based on **"Constitution of the Socialist Republic of Vietnam (2001)"[9]**

SECT 226
Article 226.- Illegally using information in computer networks
1. Those who illegally use information in computer networks and computers as well as put information into computer networks in contravention of law provisions, causing serious consequences, who have already been disciplined, administratively sanctioned but continue to commit it, shall be subject to a fine of between five million dong and fifty million dong, non-custodial reform for up to three years or a prison term of between six months and three years.

2. Committing the crime in one of the following circumstances, the offenders shall be sentenced to between two and five years of imprisonment:

a) In an organized manner;

b) Causing very serious or particularly serious consequences.

3. The offenders may also be subject to a fine of between three million dong and thirty million dong, a ban from holding certain posts, practicing certain occupations or doing certain jobs for one to five years.

4.8 Others

Others coutry in ASEAN like Laos and Cambodia don't have any law that manage computer system and network access. Indonesia already explained at Section 3.

5 Conclusion

In Indonesia, in the law it self never disscus about port scanning. All point that we disscuss (point 30, point 31, point 33) from chapter VII about forbiden activity, disscuss about the future action after doing port scanning.

In point 30, it disscusses about illegal access into a computer system or network to get electronic document/information, breaking into a sistem, pass over, or penetrate the security system. port 25 attack (mail server attack) is one of many cases that the attacker will charged by point 30 because with this attack, the attacker will retrive all information and mail from the company mail server.

In point 31, it disscuss more in interception in computer system or network. One of the example is interception on port 80. The attacker intercept the port 80 to get all data between the server and client.

In point 33, it disscuss more on harming a computer system. this case is more on installing virus, worm, and trojan on computer system.

And about the other ASEAN countries, they have its own law except Laos and Cambodia. The main point of their law basically just disscuss about accessing the computer system or network. But each country law has its own advantages and disadvantages which makes every laws unique.

References

[1] Simpson, Michael T.*Hands-on Ethical Hacking and Network Defense, 1st edition,* , Course Technology, CengageLearning, 2006

[2] *"UUITE"* http://www.scribd.com/doc/2362550/UU-ITE, accessed September 10, 2009

[3] *"Computer Misuse Act 2007 of Brunei Darussalam"*http://www.agc.gov.bn/LOB/PDF/Computer%20Misuse.pdf, accessed October 29, 2009

[4] *"ACT 563 Computer Crime ACT 1997"* http://www.ktak.gov.my/system/uploaded/files/Computer%20Crimes%201997%20-%20Act%20563.pdf, accessed October 5, 2009

[5] *"The Union of Myanmar The State Peace and Development Council The Electronic Transactions Law (The State Peace and Development Council Law No. 5/2004)"* http://www.burmalibrary.org/docs/Electronic-transactions.htm, accessed October 29, 2009

[6] *"Electronic Comerce ACT of 2000"* http://www.chanrobles.com/republicactno8792.htm, accesed October 29, 2009

[7] *"Computer Misuse Act (Singapore)"* http://unpan1.un.org/intradoc/groups/public/documents/APCITY/UNPAN002107.pdf, accessed October 29. 2009

[8] *"Electronic Transaction Act (Singapore)"* http://unpan1.un.org/intradoc/groups/public/documents/APCITY/UNPAN025623.pdf, October 29, 2009

[9] *"Computer Related Crime Act B.E 2550 (2007)"*http://www.uih.co.th/download/cc_act_en_unofficial.pdf, accessed October 29, 2009

OPTIMIZE QOS OF METRO ETHERNET NETWORK WITH PACKET SCHEDULLER USING WEIGHT ROUND ROUBIN ALGORITHM (WRR)

Franciska Sukmadewi Megawati[1], Sahrul Arif[2], Puput Heryayo[3]

[1,2,3]Dept. Electrical Engineering, Post Graduate of Institut Teknologi TELKOM
[1]sukmadewi_mega@yahoo.com, [2]sahrularif@yahoo.co.id, [3]ppt.herya@yahoo.com

Abstract

Classification of Traffic are Real-time (voice and video streaming) and Non-Real-time (email and data transfer FTP), very sensitive with the delay and packet drop trouble when network get bottleneck. The function of Packet Scheduler is to divide resource (bandwidth) so it can minimize delay and packet drop trouble in network. Metro Ethernet is kind of MAN's (Metro Area Network) coverage. With assumed full traffic with various weight, role of the scheduler is traffic with higher weight will have most priority. The Simulation using OPNET Modeler 14 software. There are 7 scenarios with 3 output of comparison. First, comparing traffic that generated by system with that given a weight. Result shows that weighted with WRR is more optimize. Second, giving arrival traffic's model, there are data is given traffic pareto, voice and video with poison. Than It was compare with no offset traffic model. With the same weighted, No offset model giving packet drop 2 times more, delay and jitter more high. Third, dynamic weight between data, video and voice. Performance that *acceptable* based on ITU-T G.114 are given weight for ata (5), video streaming (40) and voice (55). When data significantly increase and real time decrease, the performance isn't acceptable by standard anymore.

Keyword: packet scheduler, QOS Metro Ethernet, algoritma WRR

1. Introduction

1.1 Ethernet

Ethernet is known as interface to connect user device (laptop or computer), almost in LAN. Based on definition of Metro Ethernet, It can be known as Ethernet's Network that implemented in metropolitan area. Big companies can operate this technology to link up branch office to the internet system.

1.2 Packet Scheduler

Scheduling Algorithm is used to manage computer's network, Operating system and real time application. Packet scheduling purpose is to divide resource (bandwidth) to the individual user, decide which packet that require more Bandwidth, "inspect" packet that leaving and comparing It to resource.

1.3 Algorithm Round Robin

Scheduling in Algorithm determine which packet to leave and wait in line. All nodes using the same schedule format.

The most easy algorithm is determine the packet's length for that arrive to the *timeslot* in *Round-Robin fashion*. Packets are allocated rotately from the first wavelength (λ_1) then second wavelength (λ_2), third wavelength (λ_3) until the last wavelength operates.
In this queue packet, possibly there is packet loss. It's because packet are queuing to satisfied channel. Packet information essential to inform:

(1) Sum of last length which is packet before was decided
(2) Length of *wavelength buffer*

1.4 QoS (Quality of Services)

Network performance is standardization that can be attaining by user in a various service that provide by network. Network parameter specified as follow:

- Delay is time that need to through from source to destination.

$$Delay = \frac{packet\_size \times 8}{line\_speed} \times 1000ms$$

- Packet Loss / Packet Drop
  Is amount packet that loss or dumped by network.
- Jitter
  Is delay variation from end-to-end delay.
- Throughput
  Is amount from requests/packets which deliver through network per time unit, amount in *Bps*. *Throughput* connect to bandwidth which served by network, that not all of bandwidth used in network application.

Table 1 Recommend ITU-T G.114 for Delay

Range in Milisecon	Description
0 – 150 msec	Acceptable for most user application
150 – 400 msec	Acceptable provided that administrators are aware of the transmission time and it's impact on transmission quality of user application
> 400 msec	Unacceptable for general network planning purpose, it is recoqnized that in some exceptional cases this limit will be exceeded.

2. Design System

Using software OPNET to design metro Ethernet network.

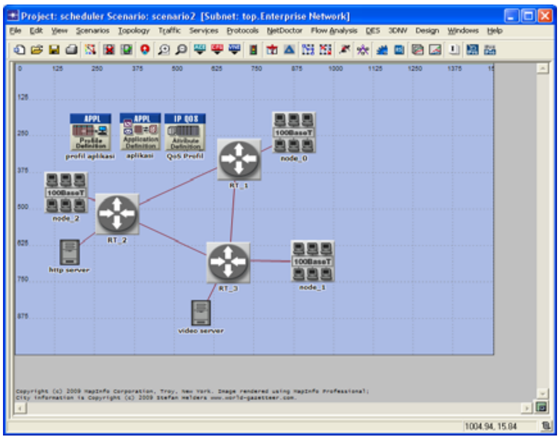


Figure 1. OPNET's Metro Ethernet Design

The simulation's attribute consist of :

1. Three nodes with It's specification
   - Node 0: LAN with 25 workstation. Switch rate up to 500 Mbps. Generated application is voice.
   - Node 1: LAN with 10 workstation. Switch rate up to 500 Mbps. Generated application is voice.
   - Node 2: LAN with 25 workstation. Switch rate up to 500 Mbps. Generated application is voice.
2. HTTP Server is Server that used to run data application
3. Video Server is server that used to run real time application.

There are 7 scenarios that run in this research, as follow:

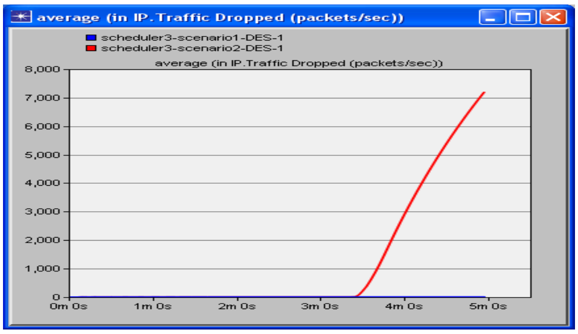
1. Scenario 1
   Network is not filled by QoS. The Traffic is generated by system.

2. Scenario 2
   QoS is given by allocate different weight in each triple play services
3. Scenario 2_with Link Traffic
   Weighted Network and also given arrival pattern that are pareto traffic for data, poison for voice and video.
4. Scenario 2 with Link No Offset
   In same condition before, but arrival pattern is No offset, traffic generated randomly by network system.
5. Scenario 2 single application
   Networks just run 1 application. Queue is given with FIFO models
6. Scenario 3 WRR Single Application
   Queue is given with WRR model
7. Scenario3_weight_v1,Scenario3_weight_v2 and Scenario 3_weight_v3
   The simulation is vary network with various weighted, so It can known which various that do well.

3. Analysis

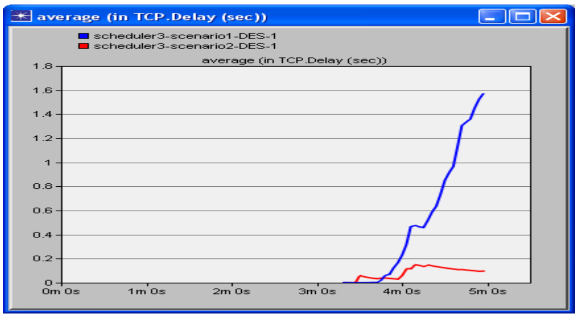
1. Comparison between Scenario 1 and Scenario 2

Packet Drop



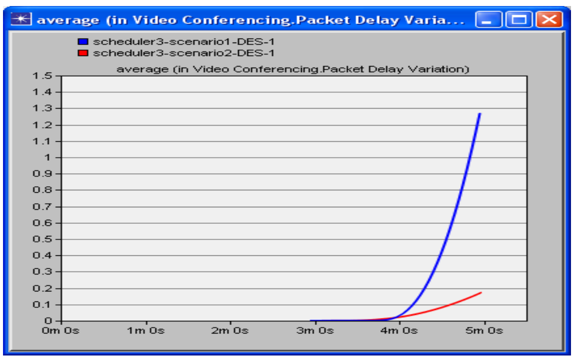
In scenario 2 the packet was droped in period 3.5 minutes. But in scenario 1 there is no packet drop.

Delay



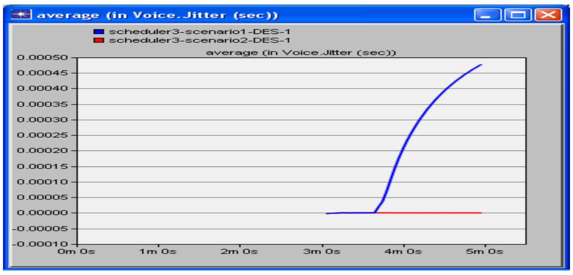
Scenario 1 shows delay more extrem. It's as consequence, when there's no drop the delay become so enormous.

Packet Delay (Video conferencing)



Scenario 1which is not given schedule, the delay so enormous.

Jitter (voice)

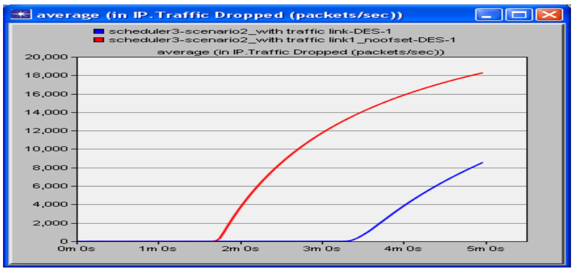


Scenario 1 increasing rapidly and scenario 2 inclined constantly.

2. Scenario2_with traffic link : Scenario2_with traffic link no offset

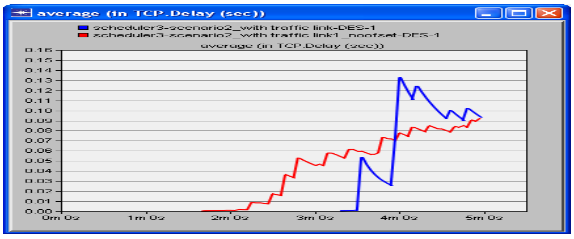
Packet Drop

Packet that droped by weighted link with no offset 2X more higher than weighted link without no offset.



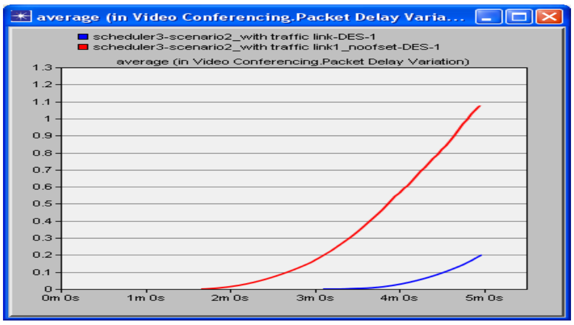
Delay

Delay with no offset lesser and more variation traffic pattern.



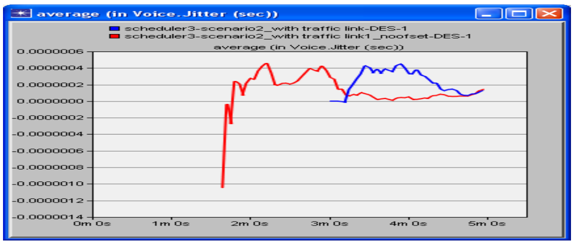
Packet delay Variation (video conferencing)

Delay in application realtime, with arrival pattern traffic is no offset, show result more enormous.



Jitter (voice)

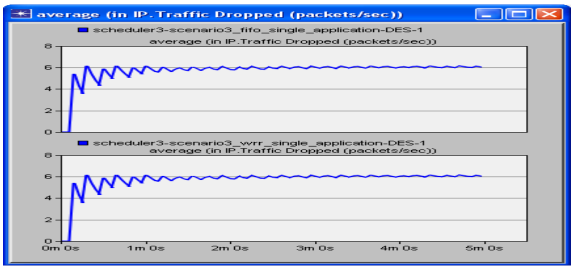
Jitter with no offset more constant than non no offset pattern.



3. Comparison Scenario FIFO single application with WRR single application

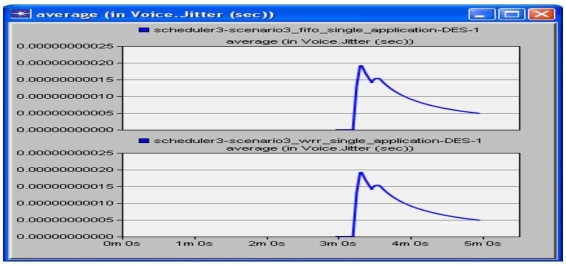
Traffic drop (packet/second)

Both of scenarios have same result to the packet drop effect.



Jitter

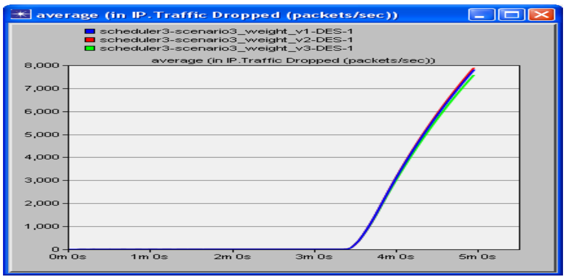
As well to the jitter, both scenario show the same result



4. Comparison between traffic weighted with v1, v2, and v3

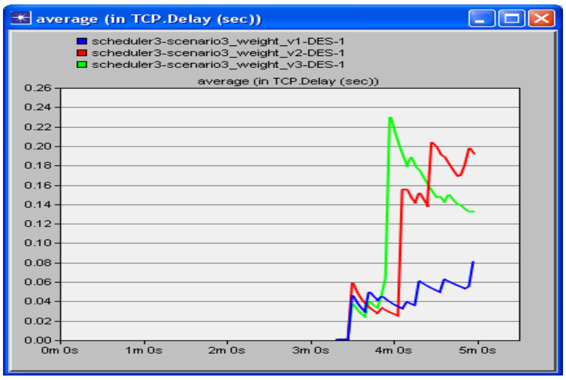
Packet drop

Packet drop from the all three scenarios are not too different.



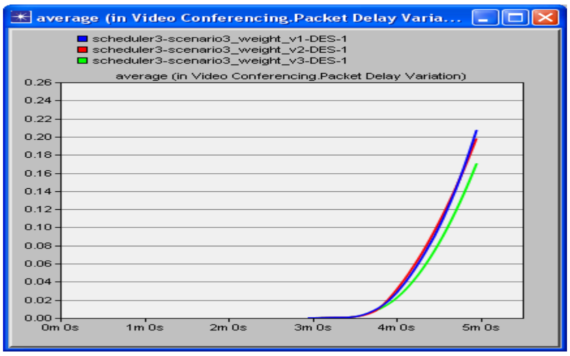
Delay

Turn to the set weighted, the most significant delay is v3. Followed by v2 and the most low is v1.



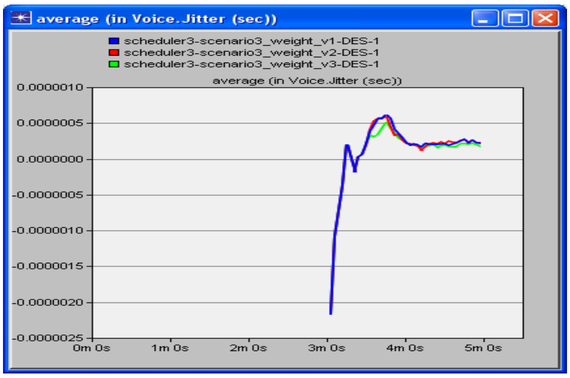
Packet Delay Variation (video)

Delay realtime in scenario v1 and v2 are show not too different. Scenario v3 shows the most low delay.



Jitter (voice)

The all three scenarios show the same result.



4. Conclusion

From the simulation, can conclude as:

1. To the first comparison, the traffic which generated by system was no drop, as consequence vey high delay and jitter. To the weighted network, there was packet drop that can be tolerated. Delay and packet loss less than 150 s.
2. Given arrival traffic pattern that are pareto for data, poison for video and voice, would be given better performance than No Offset pattern. With No Offset pattern, packet drop 2X higher.
3. Traffic's comparison in scenario 3 with variation weighted given (v1, v2, v3). V1 shows the most good performance. It's prove to the delay that allowed by ITU-T standardization which is less than 150 s.

References

[1] Duato, Jose, 2003, *Interconnection Network An Engineering approach,* Morgan Kauffman Publisher, United States of America.

[2] Zhao, Zhuang, 2003, *A Method and Simulation Study: Network Dimensioning of the Broadband Residential Ethernet-based Access Network,*Multimedia Communication Network Engineering Research Center Nokia, The Chinese Academy of Science.

.

PERFORMANCE ANALYSIS OF H.264/AVC VIDEO STREAMING OVER WIRELESS LAN WITH IEEE 802.11e ENHANCED DISTRIBUTED CHANNEL ACCESS QOS SUPPORT

Hadi Wihardja[1], Ida Wahidah[2], Sofia Naning Hertiana[3]

[1,2,3]Electrical and Communication Faculty, Institut Teknologi Telkom
[1]djadja.hw@gmail.com, [2]idw@stttelkom.ac.id, [3]snh@stttelkom.ac.id

Abstract

Recently more and more popularity of wireless LAN promote diferent kind transmision of multimedia application, streaming video is one of the most important applications. In video streaming service, user does not have to wait the download process, but at the same time user can download and play. And then, the changing condition in network IEEE 802.11e EDCA can affect the video streaming quality sent by server. In this research, we investigate the video quality attained in streaming H.264 video over IEEE 802.11e EDCA networks using an integrated tool environment called EvalVid, which comprises an H.264/AVC encoder/decoder, a network simulator and video quality evaluation tools. The benefit of such an integrated tool environment is that it allows the evaluation of real video sources compressed using an H.264 encoder. This final project will analyze video streaming quality content received by user/client using parameters like throughput, delay, loss packet, jitter and PSNR. Then, Mean Opinion Score is used for subjective evaluation. Simulation results show that effect of EDCA implementation visible when the traffic equals or greater than maximum capacity of network. EDCA implementing the higher priority for video H.264 can increase throughput 39% larger than that without EDCA, packet loss under 3%, delay under 27 ms, and jitter 0.04 second. Video quality evaluation reaches Y-PSNR more than 25 dB. In the end, implementing EDCA for video streaming H.264 can result in QoS parameter values suitable to ITU-T recommendation.

Keywords: H.264/AVC, IEEE 802.11e, wirelessLAN, EDCA, video streamming, QoS

1. Introduction

The development of communication world requires accessibility and quality guarantee of various services, especially for real time service. The popular services include Youtube, Skype, etc. The IEEE 802.11 standard was not designed to provide differentiation and priority based on service type. In order to meet this requirement, the standard added QoS support later.

Video streaming demands the fulfillment of bandwidth to maintain the high quality service. QoS gives quality to video streaming service for particular number of user as long as the bandwidth is guaranteed. Therefore, the efficacy and efficiency of bandwidth needs to be concerned. This advantage has to be traded with computational load that requires better and faster processor equipment.

2. Theoretical Background

2.1 Video Coding Concept

Basically, video can be illustrated as a pile of images with identical frame size, which are displayed sequentially with certain frequency. Hence, video possesses three dimensions, i.e. two spatial dimensions (horizontal and vertical) and one time dimension. Two important things can be utilized to compress the video, spatial and temporal redundancy. The removal of spatial redundancy is conducted by taking advantage from the fact that human eyes cannot distinguish colors rather than brightness, so that the images can be compressed (this technique is the same as lossy color reduction compression for image). The removal of temporal redundancy (interframe compression) is done by sending and coding only the changed frame, whereas the similar data is kept.

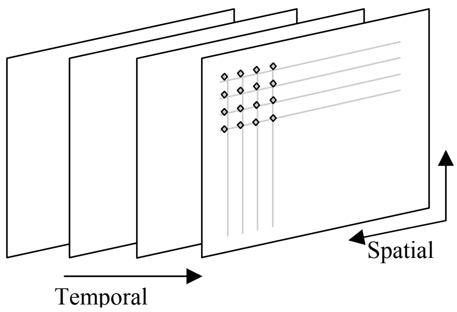


Figure 1. Temporal and Spatial Concept

2.2 H.264/AVC Coding Standard

H.264 (MPEG-4 *Part* 10) or well known as Advanced Video Coding (AVC) is a digital video codec that performs high compression ratio by employing adaptive transformation block method. ITU-T Video Coding Expert Group (VCEG) together with ISO/IEC Moving Picture Expert Group (MPEG) develops H.264. This group is named as Joint Video Team (JVT) in 2003.

The objectives of H.264 development are to make a digital video standard that can yield good video quality at low bit rate compared with previous digital video standard (MPEG-2, H.263, or MPEG-4 Part 2) without significant changes and easy to implement. The other objective is feasibility for various applications, such as video broadcast, DVB storage, RTP/IP packet networks, and ITU-T multimedia telephony systems.

H.264/AVC coding standard is stacked on two conceptual layers, i.e. video coding layer (VCL) aimed to achieve video content efficiency, and network abstraction layer (NAL) that is formatting the output of VCL, appending suitable header information, and forwarding it to the transport layer or storage media.

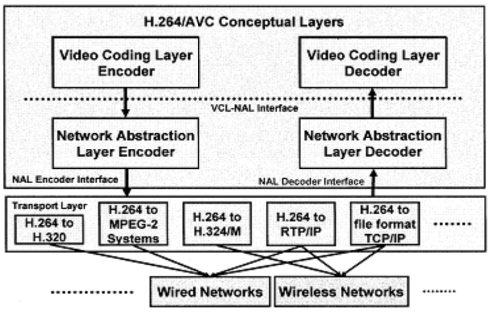


Figure 2. H.264/AVC Conceptual Structure

2.2.1 H.264 video structure

H.263 video structure comprises video sequence that has one or more GOP (Group of Pictures) started by header and closed by end of sequence. Picture (frame) that is part of GOP is called primary coding unit of the video sequence, representing luminance (Y) and two chrominance (Cb and Cr) values. In H.264/AVC coding, we use 4:2:0 format, in which chrominance component has half portion of luminance component. Furthermore, macroblock is known as basic coding unit in MPEG algorithm, i.e. 16x16 pixel segment in a frame. Macroblock covers square area with 16x16 pixel size for luminance component and 8x8 pixel for each chrominance component. Block is the smallest coding unit in MPEG algorithm. In addition, 8x8 pixel or 4x4 pixel can be one of the luminance, red chrominance, or blue chrominance. A number of macroblocks is called slice to be processed later. There are five types of slice, i.e. I, P, B, SI, and SP. The sequence is from left to right, then top to bottom. It is important for error handling. If error occurred, the decoder would skip to the next slice.

2.2.2 H.264/AVC Codec

As the previous coding standards (H.263 and MPEG-1,2), H.264/AVC is a coding standard based on hybrid video coding.

Input image is divided into macroblocks. Each macroblock consists of three components, i.e. Y, Cr, and Cb. Y represents brightness, while Cb and Cr represent color intensity. In I-slice, all macroblocks is coded by intra mode. For P-slice, all macroblocks is predicted by employing motion compensation with one reference frame. B-slice resembles P-slice with two reference frames. SI- and SP-slice are special slices that cannot be found in the predecessors. SP-slice is coded to gain efficiency. SI-slice is coded to correct errors when using intra prediction.

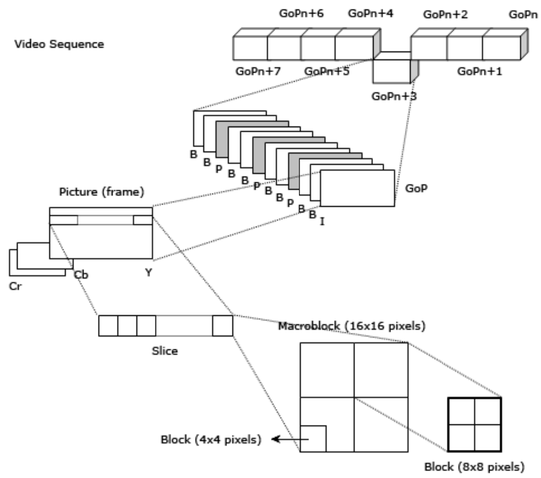


Figure 3. H.264/AVC Video Structure

In H.264/AVC destination process, entropy decoder will decode the quantization coefficient and movement data. Like the encoding process, prediction signal is resulted from intraframe or motion compensated prediction that is added with the inverse of transformation coefficient. After deblocking filter, the decoding process is finished and the macroblock is stored in memory for successive prediction.

2.2.3 *Deblocking filter*

Deblocking filter is a new element in MPEG video compression standard. The main function of deblocking filter is to reduce the blocking distortion in each decoded macroblock. At encoder, decoding filter is applied after inverse transform and before reconstruction and storing process. At the decoder, deblocking filter application is done after inverse transform and before reconstruction and displaying process.

Deblocking filter is employed to improve image quality by removing the blocking artifact commonly occurred to digital video. Deblocking filter is applied in each 4x4 block or 16x16 macroblock to achieve better video quality.

This filter has two advantages:

- The edges of block and macroblock will be finer
- Filtered macroblock will be used to predict next frame (at encoder), so that fewer residues are found in the prediction process

3. System Modeling and Simulation

System modeling in this research is aimed to build a network simulation of wireless LAN 802.11e EDCA and non-EDCA in order to analyze their performances. Specifically, the capability in video streaming transmission for various compression formats, mainly H.264/AVC compression. Beside H.264, H.263 and MPEG-4 Visual Part 2 are also observed as comparison.

3.1 System Modeling

In this part, we describe the modeling for video streaming simulation over WLAN 802.11e EDCA network by means of Network Simulator. Globally, the system model in this research is shown below.

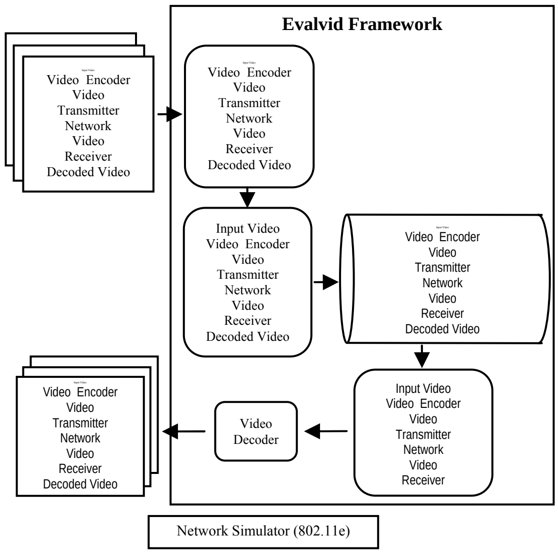


Figure 4. System Modeling

As can be seen in the figure, the system input is video. The video is first compressed by encoder into several formats. The compressed video is extracted to statistical information that represents the packet size of each video frame that will be transmitted. This information is utilized as the inputs to Network Simulator. The data output of NS is adjusted to communication channel. At last, both videos are compared to notice the influence of network to video quality.

4. Performance Analysis of Video Streaming over IEEE 802.11E EDCA Network

4.1 Testing and Analysis

In this research, performance test simulation is conducted for video quality as well as traffic in network. The traffic sources are H.264, MPEG-4, and H.263 video with predetermined bitrate generated by NS software. Measured parameters include PSNR, MOS, packet loss, delay, bandwidth, and jitter.

First stage is the evaluation of WLAN network, i.e. IEEE 802.11 versus its amendment (802.11e EDCA). In this stage, the number of users is varied from 1 to 16. From this, we obtain the specific number of users in which the performance is still acceptable. The second stage is the investigation of video streaming quality for the three coding schemes.

4.1.1 Network Traffic

In this research, performance parameters concerned embrace:

- Throughput, number of successfully received bits divided by observation time
- Packet loss, the percentage of missing packets in data transmission process from traffic source to destination node. Packet loss for voice and multimedia application that can be tolerated is 1-3% (ITU-T G.1010 standard)
- Delay, the time interval between packet departure and arrival at receiver. Good delay is 0-150 ms (ITU-T standard)
- Jitter, the variability of packet arrival. This jitter value affects the video quality so that a buffer is needed to store out-of-sequence incoming packets. Jitter should be less than 5 ms (ITU-T G.1010 standard)

4.1.2 Video Quality

Performance parameters for video quality embrace:

- Peak Signal to Noise Ratio (PSNR) is used to describe the image degradation due to encoding, compression, and transmission error. PSNR value is expressed in dB, where good quality is achieved when PSNR greater than 25 dB. PSNR parameter has three major components, i.e. Y-PSNR, U-PSNR, and V-PSNR.
- Mean Opinion Score (MOS) is the result of visual observation by some respondents to the displayed video.

4.2 Traffic analysis of IEEE 802.11e EDCA and non-EDCA networks

The scenario of network performance measurement is to compare between IEEE 802.11 and its amendment. Each user generates three data types, i.e. video source simulation with access category 1 and Constant Bit Rate (CBR) as much as 350 kbps, audio source with access category 2 and CBR of 128 kbps, and background traffic with access category 3 and FTP. Measured parameters include throughput, packet loss, delay, and jitter.

4.2.1 Throughput

In this experiment, data is sent from source, WLAN user, to destination via access point for 30 seconds. This experiment is done for 1 to 16 WLAN users over both networks, i.e. IEEE 802.11 EDCA and non-EDCA.

From Figure 5 and 6, it can be seen that the increase number of users yields CBR throughput improvement. It occurs continuously until the network maximum throughput is achieved, ranging from 4200 kbps to 5100 kbps.

If observed thoroughly, the experiment on more than 11 users results in relatively stable throughput, i.e. 1.26-1.44 Mbps for CBR 128 and 3.49-3.92 Mbps for CBR 350. This implies that the chance of data transmission given to both data types is in accordance with the generated traffic. On the other hand, CBR throughput in IEEE 802.11e EDCA network decreases since 11-user experiment. This implies that CBR-128 traffic with AC 2 moves back when there is CBR-350 traffic with AC 1 in the network.

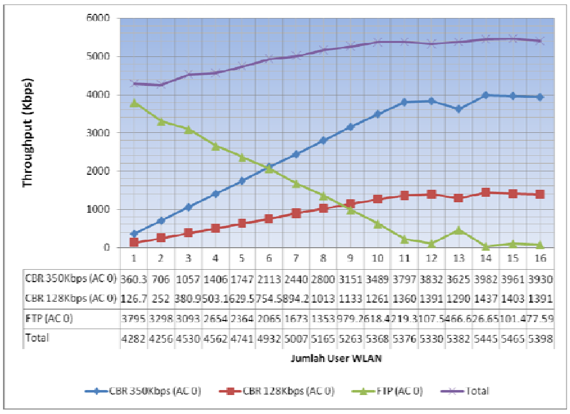


Figure 5. Throughput of IEEE 802.11 Non EDCA

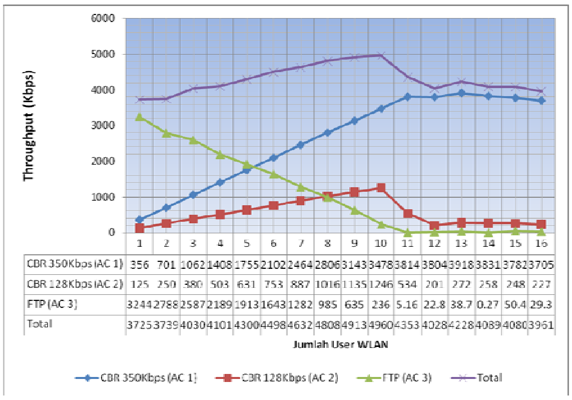


Figure 6. Througput of IEEE 802.11e EDCA

4.2.2 Packet Loss

Mean packet loss in simulation on both types of network has fluctuant values during observation. In non-EDCA network, the increase number of WLAN users significantly affects the packet loss from 0.7% to 29%, especially above 11 users. However, for EDCA network, CBR-128 traffic has packet loss of 82% due to packet discarding in buffer.

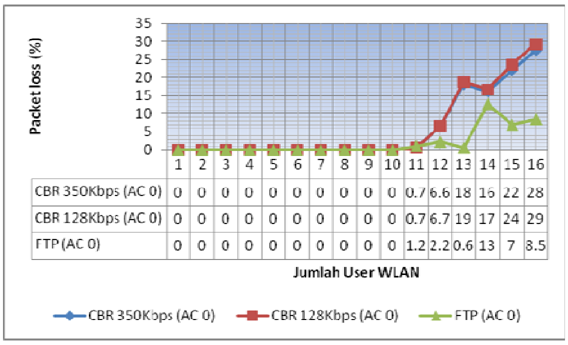


Figure 7. Packet Loss of IEEE 802.11 Non-EDCA

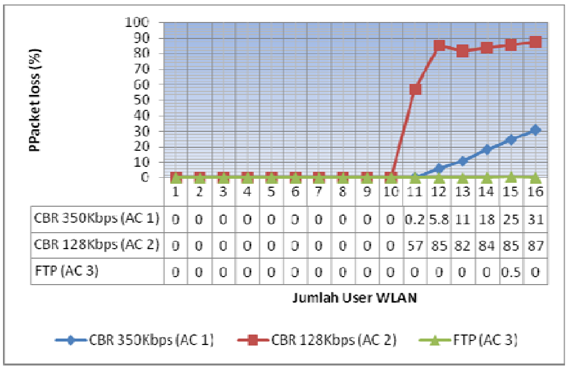


Figure 8. Packet Loss of IEEE 802.11 EDCA

4.2.3 Delay

Mean delay data measured in this research is merely buffer and transmission delay. It shows time interval from packet transmission, buffering to packet reception.

The result from delay measurement for both types of network seems that the increase number of WLAN users does not influence mean delay. This is probably due to sudden packet discard when waiting too long in buffer or capacity overload.

4.2.4 Jitter

Jitter from the measurement has average value below 1 ms in all experiments. The increase number of WLAN users does not affect jitter.

4.3 Performance Analysis of Video Streaming Over IEEE 802.11e EDCA Network

This analysis is resulted from measurement scenario for IEEE 802.11 and its amendment in input video transmission for different compression formats. In this research, we use three formats, i.e. H.264/AVC, MPEG-4 Visual Part 2, and H.263. Video streaming performance testing is done in EDCA addition for above-12-user environment. This scenario assigns highest priority level for video traffic with access category 0, where the other 12 users function as background traffic.

In order to produce three video compression formats, encoding process is conducted by setting the output bitrate from 64 to 768 kbps for each format. Then, the encoding result is measured for Y-PSNR and a video input with similar Y-PSNR will be selected. H.264 compression format yields nearly unchanged Y-PSNR for lower bitrate compared with MPEG-4 Visual Part 2 and H.263 compression.

4.3.1 Performance Analysis of Video Streaming Traffic

4.3.1.1 Throughput

Based on Figure 9 and Table 1 below, throughput comparison between non-EDCA and EDCA network with frame size of QCIF, it is apparent that throughput is improved for all video input. For Akiyo, H.263 reaches 76.5 kbps, MPEG-4 56.2 kbps, and H.264 94.3 kbps. For Foreman input, H.263 gains 105.1 kbps, MPEG-4 143.9 kbps, and H.264 105.5 kbps. The percentage of throughput change in this experiment is ranging from 22.2% up to 39.3%.

Table 1. Throughput of QCIF Video Traffic (Kbps)

	QCIF					
	Akiyo			Foreman		
	H.263	MPEG4	H.264	H.263	MPEG4	H.264
Sender	270.304	253.501	277.5	473.54	463.66	268.57
Receiver No EDCA	191.601	196.518	181.48	365.78	317.23	162.644
Receiver EDCA	268.125	252.757	275.81	470.86	461.11	268.1
margin	78.7032	56.9832	96.019	107.76	146.42	105.926
margin	2.1792	0.744	1.6912	2.6856	2.5496	0.4696
% perubahan	28.3103	22.185	33.992	22.189	31.03	39.2658

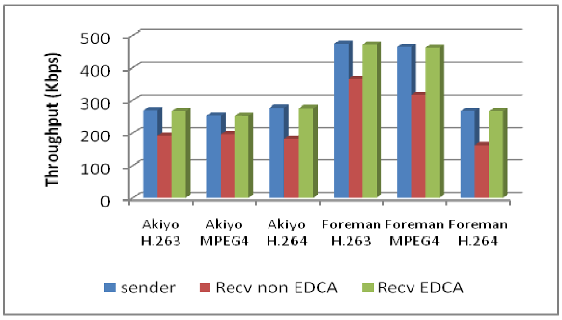


Figure 9. Throughput with QCIF Frame Size

Based on that, we conclude that EDCA network is capable to increase throughput for higher priority traffic type. The greatest rise is experienced by H.264 video, i.e. 39.3% for Foreman and 34% for Akiyo.

4.3.1.2 Frame Loss

The next analysis is video frame loss by observing the number of frame discarded in the transmission process. Figure 10 and Table 2 shows the frame loss reduction, ranging from 17.4% to 38.6%. The highest frame loss decrease is experienced by MPEG-4 video for Foreman, 38.6%, and H.264 video for Akiyo, 28%. This is because packets with highest priority are not stacked in buffer. Packet loss for EDCA network supports ITU-T recommendation, i.e. under 3%.

Table 2. Frame Loss of QCIF Video Traffic (%)

	QCIF					
	Akiyo			Foreman		
	H.263	MPEG4	H.264	H.263	MPEG4	H.264
non EDCA	24.7	17.7	29.3	33.7	39.3	34.3
EDCA	1.3	0.3	1.3	1.3	0.7	0.3
margin	23.4	17.4	28	32.4	38.6	34

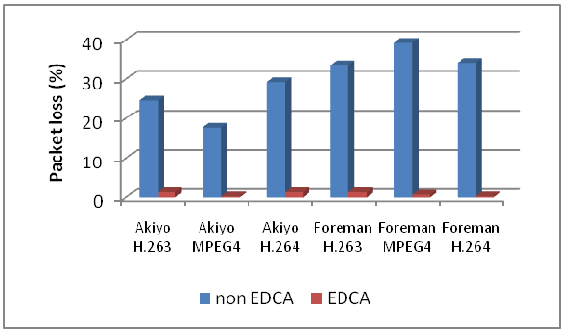


Figure 10. Frame Loss with QCIF Frame Size

4.3.1.3 Delay

According to Figure 11 and Table 3, delay comparison between non-EDCA and EDCA network with QCIF size is quite different. For Akiyo video, H.263 has delay of 750 ms, MPEG-4 is 690 ms, and H.264 is 900 ms. For Foreman video input, H.263 has delay of 450 ms, MPEG-4 is 620 ms, and H.264 is 910 ms.

Table 3. QCIF Video Delay (sec)

	QCIF					
	Akiyo			Foreman		
	H.263	MPEG4	H.264	H.263	MPEG4	H.264
non EDCA	0.77163	0.712716	0.930656	0.520682	0.674825	0.9378
EDCA	0.026	0.0242	0.027923	0.067171	0.058937	0.0253
margin	0.74564	0.688516	0.902733	0.453511	0.615888	0.9125

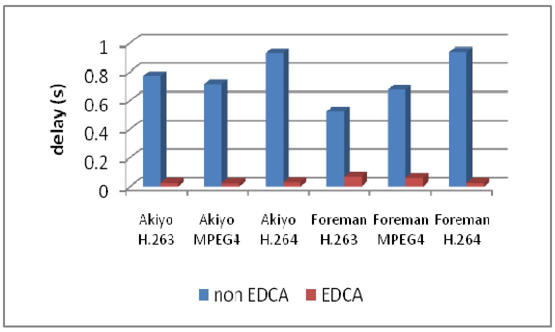


Figure 11. Delay for QCIF Frame Size (sec)

Based on the simulation data, EDCA network can decrease video streaming delay for higher priority traffic type as much as 260-910 ms. The greatest delay reduction is for H.264 compression using Foreman video, 910 ms, and Akiyo, 900 ms. This is caused by limited number of packets queued in the buffer. The resulted delay is lower than the recommended maximum delay of 150 ms.

4.3.1.4 Jitter

Figure 12 and Table 4 shows the jitter comparison between non-EDCA and EDCA network with QCIF frame size.

Table 4. QCIF Video Jitter (sec)

	QCIF					
	Akiyo			Foreman		
	H.263	MPEG4	H.264	H.263	MPEG4	H.264
Non EDCA	0.6316	0.64016	0.5737	0.0097	0.4439	0.56698
EDCA	0.0052	0.00718	0.0251	0.0083	0.0155	0.04007
margin	0.6264	0.63298	0.5486	0.0014	0.4284	0.52691

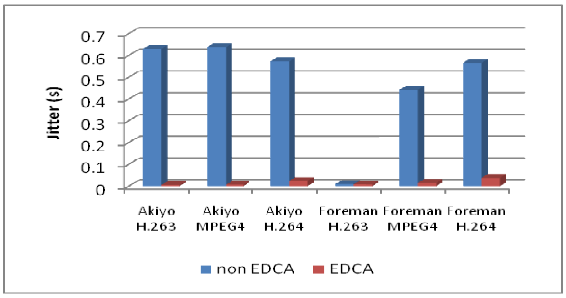


Figure 12. Jitter for QCIF Frame Size (sec)

EDCA network can decrease jitter for higher priority traffic, ranging from 1.4 ms to 550 ms. The highest jitter improvement is achieved in MPEG-4-compressed Akiyo video, as much as 630 ms, and H.264-compressed Foreman video, i.e. 520 ms. Generally, jitter value does not satisfy ITU recommendation of 5 ms.

4.3.2 Video Streaming Quality Analysis

Video quality parameters observed in this research include PSNR and MOS. Reference is made to encode PSNR before transmission over IEEE 802.11 network. Subsequent research accentuates on Y-PSNR, because human visual system is more sensitive to luminance than chrominance.

4.3.2.1 Peak Signal to Noise Ratio (PSNR) Analysis

Video impairment and poor picture received by WLAN user can be represented by Y-PSNR. For non-EDCA system, Y-PSNR value goes down below normal for Foreman file, which has low redundancy and requires high bitrate. This is because Foreman owns low temporal and spatial redundancy. Hence, a few packet loss impacts decoding process significantly.

Y-PSNR variation in non-EDCA network ranges between 16 and 36 dB. The greatest Y-PSNR is found in Akiyo video using H.263 format, i.e. 36 dB, while the lowest Y-PSNR is for Foreman using H.264, i.e. 16 dB.

Table 5. Average Video Streaming Y-PSNR (dB)

		Akiyo			Foreman		
		H.263	MPEG4	H.264	H.263	MPEG4	H.264
CIF	avg Y-PSNR ref	44.0170	44.6340	46.9250	37.9730	37.8732	37.510
	avg Y-PSNR output non EDCA	35.5920	35.4650	33.0000	18.8160	18.0469	15.920
	avg Y-PSNR output EDCA	41.8080	41.0100	41.3690	25.3120	25.6528	37.340
	margin non EDCA	8.4249	9.1689	13.9250	19.1570	19.8263	21.590
	margin EDCA	2.2089	3.6242	5.5559	12.6610	12.2204	0.169
	margin EDCA and EDCA	6.2160	5.5447	8.3690	6.4962	7.6059	21.420

In the simulation of IEEE 802.11e EDCA network, Y-PSNR changes significantly for Foreman video. This is in line with the packet loss reduction that eventually affects the video quality.

5. Conclusion

1. *Loss* does not only occurs in transmission process, but also in encode/decode process. It could influence the PSNR of video.
2. Based on EDCA network performance, the effect of increasing priority is obvious if offered traffic exceeds maximum throughput, i.e. 4-5.4 Mbps. It can be observed when more than 10 users attempt to access the network.
3. Best performance for each parameters in Akiyo: throughput of 99.7%, frame loss of 0.3%, and delay of 24 ms (all of these are achieved by MPEG-4), while jitter of 5.2 ms is performed by H.263.
4. Best performance for each parameters in Foreman: throughput of 99.8%, frame loss of 0.3%, and delay of 25 ms (won by H.264/AVC), while jitter of 8 ms is yielded by H.263.
5. Optimum decode PSNR in EDCA network is 37.3 dB for Foreman and 41.8 dB for Akiyo. The improvement factor achieved by EDCA reaches 21.4 dB.
6. The combination between 802.11e EDCA network and H.264 compression presents the best performance in our experiments.

References

[1] Adam M, Boy. 2008. ***Simulasi Pengiriman Video H.264/AVC Pada Jaringan CDMA2000 1XEV-DO.***Bandung : Tugas Akhir Jurusan Teknik Elektro IT Telkom

[2] Al-Mualla, Mohammed Ebrahim, C. Nishan Canagarajah and David R. Bull. 2002. ***Video Coding for Mobile Communictions-Efficiency, Complexity, and Resilience***.San Diego, CA: Academic Press.

[3] Dajiang He, Charles Q. Shen. 2003. ***Simulation Study of IEEE 802.11e EDCF.*** Singapore: National University of Singapore.

[4] Klaue, Jirka, Berthold Rathke, and Adam Wolisz. September 2003. ***EvalVid - A Framework for Video Transmission and Quality Evaluation.*** Berlin: Technical University of Berlin, Telecommunication Networks Group (TKN)

[5] Lo, Anthony., Geert Heijenk, Ignas Niemegeers. 2005. ***Performance Evaluation of MPEG-4 Video Streaming over UMTS Networks using an Integrated Tool Environment***. Twente : University of Twente.

[6] Stockhammer, Thomas, Miska M. H., and Thomas Wiegand. July 2003. ***H.264/AVC in Wireless Environments.*** IEEE Transaction on Circuit and System for Video Technology, Vol. 13, No. 7

[7] Richardson, Iain E. G. 2003. ***H.264 and MPEG-4 Video Compression - Video Coding for Next-generation Multimedia.*** England: John Wiley & Sons Ltd

[8] Wiegand, Thomas, Gary J. Sullivan. July 2003. ***Overview of the H.264/AVC Video Coding Standard.*** IEEE Transaction on Circuit and System for Video Technology, Vol. 13, No. 7.

[9] Wiethoelter, Sven. Et al. 2006. ***TKN EDCA Model for ns-2.*** Berlin : Technical University of Berlin, Telecommunication Networks Group (TKN).

[10] H-Series. ITU-T. ***Audiovisual and multimedia systems***

[11] G-Series. ITU-T. ***Transmission systems and media, digital systems and networks***

CONTENT PROVIDER SERVER FOR WIRELESS MOBILE DEVICE

Iwan Handoyo Putro[1], Resmana Lim [2], Stefanus Rocky Fabriano[3]

[1,2]Electrical Engineering Department, Faculty of Industrial Technology, Petra Christian University
[3]Computer Engineering Department, San Jose State University
[1]iwanhp@petra.ac.id, [2]resmana@petra.ac.id, [3]stefanusrocky.ongkowidjojo@students.sjsu.edu

Abstract

This paper described the process of building a content provider server to give a service for mobile user wirelessly. That service could be accessed from a personal digital assistant (PDA). The system is built from ASP.NET technology combined with the Content Management System from DotNetNuke. This will make administrator's managing the system; such as adding content, creating new voucher, viewing user billing and viewing statistic, easy. The system is also using the SQL Server technology, which hopefully compatible to most PDA; which use Windows Mobile Operating System. The system has been tested by using a PDA to access contents via wireless LAN. The result is server is able to provide the data that was requested by the user. Furthermore, all of the features that have been developed could also work well.

Keyword: Content Provider, Mobile Device, Wireless Access

1. Introduction

Rapid growth of technology development lead to the increased of people demand for electronic devices; especially to the equipment that has a mobile capability. Wireless technology is also developed with the emergence of devices that support these technologies. This situation then create a new service called mobile content access. However, this service still have some problems, such as :

- The limited number of mobile content providers.
- Content that is provided by providers is quite expensive because the price is charged by the amount of volume download.
- Mobile content offered by some service providers currently using cellular network, making content access price is quite expensive.

As the solution of these problems, an application that can serve the demand for mobile content using wireless LAN media is built.

Wireless LAN was chosen as the network platform for the system based on the consideration that frequency bands used in wireless LAN network is cost-free frequency band. This means users no longer burdened by the cost of bandwidth. In addition, the development of wireless LAN speed provides enough opportunity to support this application. This is because the radius of coverage offered in the specification far enough, for example the IEEE 802.11n specification that just launched offers a range of up to 300 ft and throughput between 50 to 144 Mbit/s [1] .

2. Content Management System

Content Management System (CMS) is a software system used for content regulation, including a file of documents, audio-video, computer files and web content [6]. The main idea was to make these files available and can be accessed easily either through inter-office network or via the web.

A Content Management System must have at least the features below:

- The ability to retrieve and / or create documents and multimedia contents.
- The ability to identify key users and the rights of each user to access and manage content.
- The ability to determine the role of each different content categories or types.
- The ability for sending a message to the content managers.
- The ability for managing content that have several versions.
- The ability to store content in a matter that future access could be done easily.
- The ability to manage user interface the layout of its user interface.

One of the existing CMS, DotNetNuke, is built based on "Microsoft ASP. NET" platform using "VB.NET" programming languages [2,3]. DotNetNuke is an open source application that already has an agreement under BSD license (BSD agreement) [6]. By having an open source license, DotNetNuke can be used by anyone without charge and with its support for "ASP. NET" platform makes DotNetNuke as a very powerful CMS.

3. ASP.NET 2.0

ASP. NET 2.0 is a web programming language that is part of . NET Framework 2.0. ASP. NET 2.0 has several goals in its development, namely: developer productivity, administration and management, and performance and scalability [5]. In terms of increasing developer productivity, for example: "ASP. NET 2.0" successfully reduced nearly 2/3 of the number of lines needed to build an ASP.NET application. In the field of administration and management, ASP. NET 2.0 provides the Microsoft Management Console (MMC) that enables an administrator to do the editing of the configuration settings on the fly [4].

Other feature of ASP.NET 2.0 that is found to be useful in web programming is the Membership and Role Management. Briefly speaking, when making a web site that requires users authentication the code an be quite complex. In ASP.Net 2.0, all of that task could be done by using the API provided or with built-in Web tool called Web Site Administration Tool [4]. Therefore, managing the user with each of their role, such as create, delete or edit could be done easily.

4. System Design

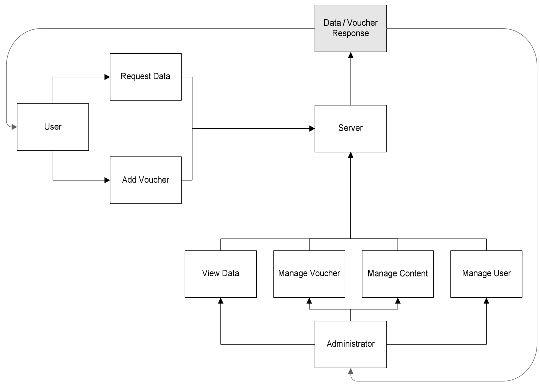


Figure 1. Block Diagram of The System

The system was designed to deliver data to the mobile client as it's requested. Server provides contents that could be downloaded by the user. Each content has a price and every client should buy a voucher before they could download the contents. The amount of the money in clients account then will be debt each time a transaction is made. When the client's account is not sufficient for further transaction, client could ask Administrator to issue another voucher. Each voucher will have a unique number according to its nominal value. Figure 1 shows the complete block diagram of the system.

4.1 Architectural Software Design

Server will be built by using components as follows:

- Windows Server 2003 as an Operating System.
- IIS as web server and SQL Server Express Edition as a database server application.
- . NET Framework version 2.0 as the basic platform ASP technology.
- DotNetNuke as Content Management System application. DotNetNuke features used in project include : LDAP Authentication, Problem Notification, Developer Community, Online Help, Mass Upload, Style Wizard, Advanced Caching, Load Balancing, Page Caching, Themes / Skins, Inline and Online Administration, Link Management, Contact Management, Database Reports, Events Management, Search Engine and Site Map.
- Visual Studio. NET 2005, SQLLink CE and SQL Server Management Studio Express Edition (SSMS-EE).

For administrators, the software required is Windows XP as the operating system, Internet Explorer version 6 or 7 as a web browser and SQL Server Management Studio Express Edition (SSMS-EE) to provide direct update capability to the database server.

4.2 Administrator Feature Design

This system allows administrator and regular users log on to the system. An administrator has the authority to make arrangements, changes and deletion of data as well as viewing history and statistics of usage. Data settings include:

- Content management, namely adding the file, changing the identity of the file, changing the status of files, view statistics and view history of downloaded files.
- Voucher management, namely adding vouchers, changing the status of vouchers, see the history and statistics of sold vouchers.
- User management, namely adding users, change user status and delete user.

Figure 2 shows the flow diagram at the Administrator side.

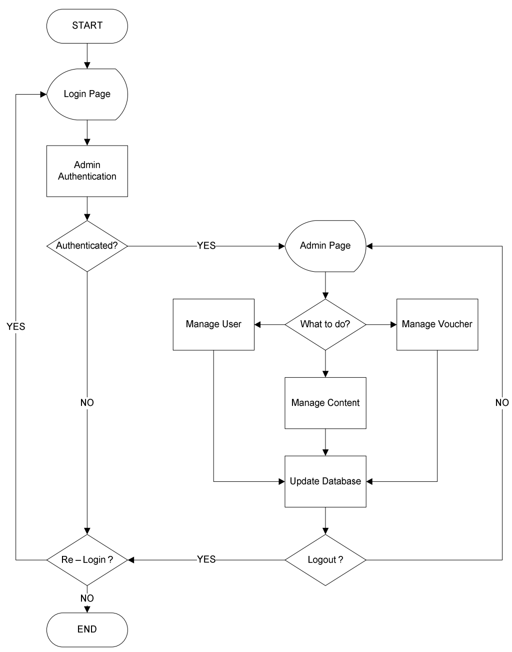


Figure 2. Flow Diagram at Administrator Side

4.3 Database Design

One of the most important aspects in the developing stage of this system is the database design. In the database design process, the requirements of the system should be considered first. In this project, the system will be built according to the following requirements:

- Information of registered users.
- File Settings and detailed information of each file, such as file type, price, description, and file storage location.
- Balance information of each user and their details usage.
- Voucher balance that can be used to add the users balance for downloading files.

By considering those requirements, the next step is to determine the entities that will be used as the basic information for creating a table. Here are the entities that will be used:

- User. This entity contains information about registered users. That information include : username, password, full name, and email address. In this project, the user entity will use the database model from CMS DotNetNuke.
- File. This entity contains the information of files that are available. That information include : file name, file type, price files, file location, and files description.
- File Type. This entity only contains information about file types and its descriptions. This entity is required to create a file type master.
- *Voucher*. This entity contains of voucher number, voucher value and active status.
- UserAddVoucher. This entity contains information about the activities of credit (the addition of balance) conducted a user, the date and the number of nominal credits.
- UserDownload. This entity contains information about the activities debit (deduction balances) conducted a user, namely: the date and amount of nominal debit.
- UserBalance. This entity contains information about the user balances and usage activity. This entity will keep all transactions ever undertaken by the user, such as the addition of balance, download activities and the final balance information.

Figure 3 show the Entity Relationship Diagram from the entities that were used.

In Figure 3, entities that are within the shaded box are additional entities that were created in this project. While entities outside the shaded box is the entity from the CMS DotNetNuke (Chris Smith, 2007). Users management by DotNetNuke will be integrated with the Microsoft ASP Membership. Therefore, the user management namely registration, change passwords, user authentication and user management will be processed by the DotNetNuke and ASP Membership.

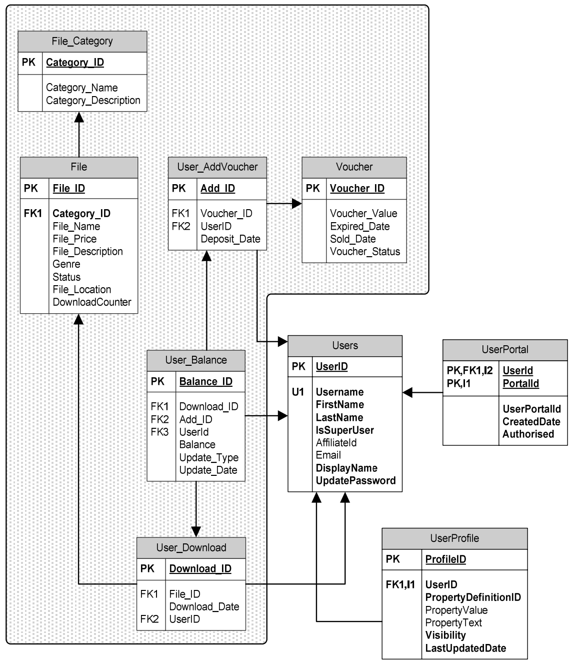


Figure 3. Entity Relationship Diagram

5. Testing and Result

5.1. Administrator Authentication

Administrator authentication is done on the DotNetNuke portal login module, and it is entirely handled by the ASP Membership. Therefore, to log in as administrator, the user must provide the right username and password. Figure 4 shows the login process on the DotNetNuke portal.

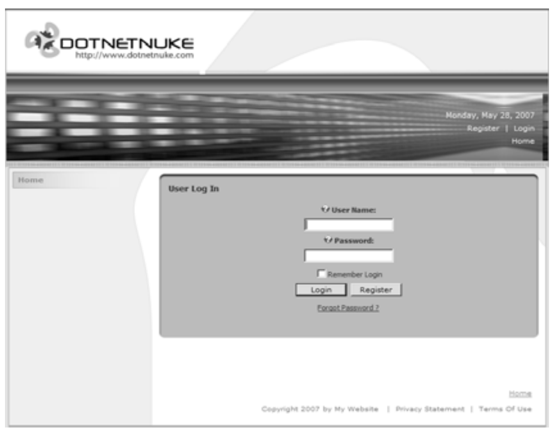


Figure 4. Login Page

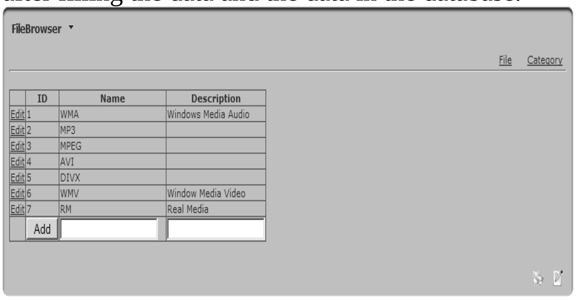
5.2. Adding *Category*

Before adding a new file, the administrator must first determine what category will be used in each file. Figure 5 shows the module files when there is no data at all.

Figure 5. An Empty Module File

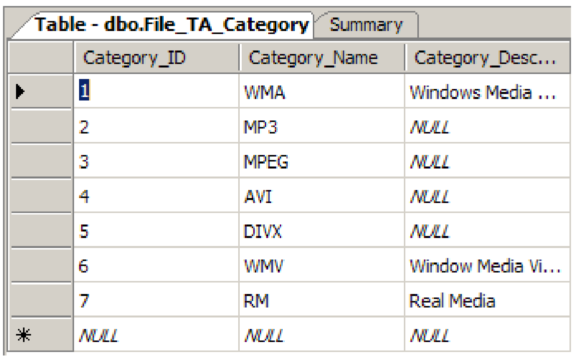
To fill in the category, click the category links on the top right corner and administrator should fill in the fields "Name" and "Description". Data on the column name must be filled while the description field is an optional. If the column name is not optional, at the time of pressing the Add button will display a sign error.

If the data has been successfully added, the new data will be instantly displayed on the module category and the data are recorded directly in the database. Figure 6 and 7 shows the module category after filling the data and the data in the database.



	ID	Name	Description
Edit	1	WMA	Windows Media Audio
Edit	2	MP3	
Edit	3	MPEG	
Edit	4	AVI	
Edit	5	DIVX	
Edit	6	WMV	Window Media Video
Edit	7	RM	Real Media

Figure 6. Module Category interface after the data is inserted



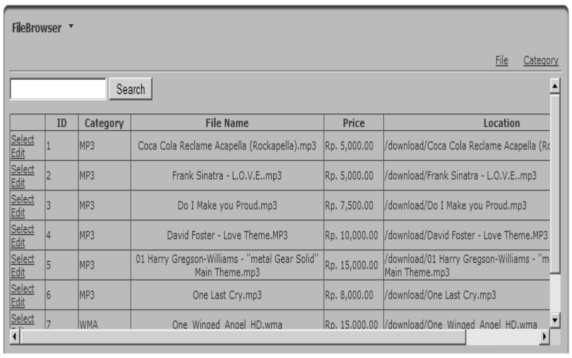
	Category_ID	Category_Name	Category_Desc...
▶	1	WMA	Windows Media ...
	2	MP3	NULL
	3	MPEG	NULL
	4	AVI	NULL
	5	DIVX	NULL
	6	WMV	Window Media Vi...
	7	RM	Real Media
*	NULL	NULL	NULL

Figure 7. Data has been successfully inserted into the database

5.3. Adding File

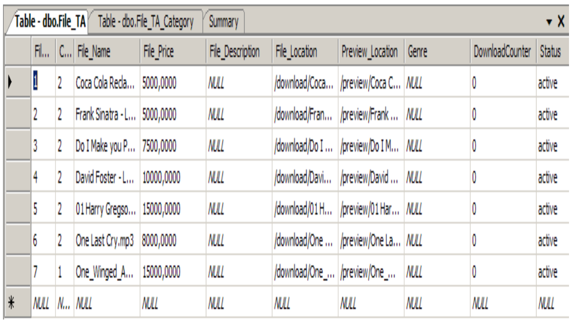
After a category has been added, administrators can add new files. Dropdown list in the category column in the module's files can be selected and adminis-trators can add new files by inserting the data on the existing parameters in each column.

In order to add data, the administrator must complete all the existing parameters. For the filename parameter, the administrator does not need to fill it manually because it has been automatically filled when the file location was inputted. direct charged automatically at the time of the input-file location. Figure 8 and 9 shows the module after the addition of data files and when data is already entered into the database.



	ID	Category	File Name	Price	Location
Select Edit	1	MP3	Coca Cola Reclame Acapella (Rockapella).mp3	Rp. 5,000.00	/download/Coca Cola Reclame Acapella (Ro
Select Edit	2	MP3	Frank Sinatra - L.O.V.E..mp3	Rp. 5,000.00	/download/Frank Sinatra - L.O.V.E..mp3
Select Edit	3	MP3	Do I Make you Proud.mp3	Rp. 7,500.00	/download/Do I Make you Proud.mp3
Select Edit	4	MP3	David Foster - Love Theme.MP3	Rp. 10,000.00	/download/David Foster - Love Theme.MP3
Select Edit	5	MP3	01 Harry Gregson-Williams - "metal Gear Solid" Main Theme.mp3	Rp. 15,000.00	/download/01 Harry Gregson-Williams - "m Main Theme.mp3
Select Edit	6	MP3	One Last Cry.mp3	Rp. 8,000.00	/download/One Last Cry.mp3
Select	7	WMA	One Winged Angel HD.wma	Rp. 15,000.00	/download/One Winged Angel HD.wma

Figure 8. Module File After Data was Inserted



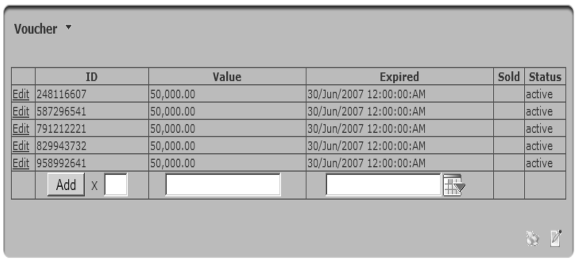
	Fi...	C...	File_Name	File_Price	File_Description	File_Location	Preview_Location	Genre	DownloadCounter	Status
▶	1	2	Coca Cola Recla...	5000,0000	NULL	/download/Coca...	/preview/Coca C...	NULL	0	active
	2	2	Frank Sinatra - L...	5000,0000	NULL	/download/Fran...	/preview/Frank ...	NULL	0	active
	3	2	Do I Make you P...	7500,0000	NULL	/download/Do I ...	/preview/Do I M...	NULL	0	active
	4	2	David Foster - L...	10000,0000	NULL	/download/Davi...	/preview/David ...	NULL	0	active
	5	2	01 Harry Gregso...	15000,0000	NULL	/download/01 H...	/preview/01 Har...	NULL	0	active
	6	2	One Last Cry.mp3	8000,0000	NULL	/download/One ...	/preview/One La...	NULL	0	active
	7	1	One_Winged_A...	15000,0000	NULL	/download/One_...	/preview/One_...	NULL	0	active
*	NULL	N...	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL

Figure 9. File Table After The Data was Inserted

5.4. Adding Voucher

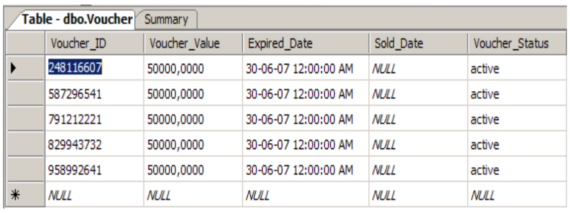
To issue a new voucher administrator could enter the voucher page and add the voucher data via the supplied form. Same as previous modules, in this module data form must be completed before pressing the button add. If the add button is pressed before all the optional parameters will display an error sign.

To fill the new voucher, administrators could fill in the ID field values with some vouchers with the same value at once. Figure 10 and 11 shows module voucher and table voucher after the data was inserted.



	ID	Value	Expired	Sold	Status
Edit	248116607	50,000.00	30/Jun/2007 12:00:00:AM		active
Edit	587296541	50,000.00	30/Jun/2007 12:00:00:AM		active
Edit	791212221	50,000.00	30/Jun/2007 12:00:00:AM		active
Edit	829943732	50,000.00	30/Jun/2007 12:00:00:AM		active
Edit	958992641	50,000.00	30/Jun/2007 12:00:00:AM		active

Figure 10. Module Voucher has been filled

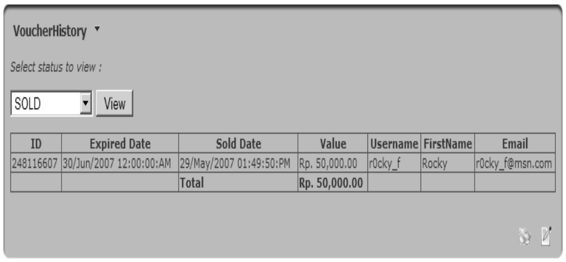


	Voucher_ID	Voucher_Value	Expired_Date	Sold_Date	Voucher_Status
▸	248116607	50000,0000	30-06-07 12:00:00 AM	NULL	active
	587296541	50000,0000	30-06-07 12:00:00 AM	NULL	active
	791212221	50000,0000	30-06-07 12:00:00 AM	NULL	active
	829943732	50000,0000	30-06-07 12:00:00 AM	NULL	active
	958992641	50000,0000	30-06-07 12:00:00 AM	NULL	active
*	NULL	NULL	NULL	NULL	NULL

Figure 11. Voucher table after the data was Inserted

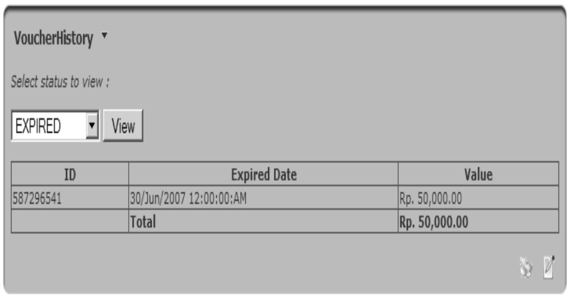
5.5. Monitoring Voucher History

In the voucher history module, administrators can see the history of the vouchers that have been sold or have expired. Figure 12 and 13 shows the screenshot of a voucher that has been sold and expired respectively.



ID	Expired Date	Sold Date	Value	Username	FirstName	Email
248116607	30/Jun/2007 12:00:00:AM	29/May/2007 01:49:50:PM	Rp. 50,000.00	r0cky_f	Rocky	r0cky_f@msn.com
		Total	Rp. 50,000.00			

Figure 12. Module Voucher history - SOLD Voucher

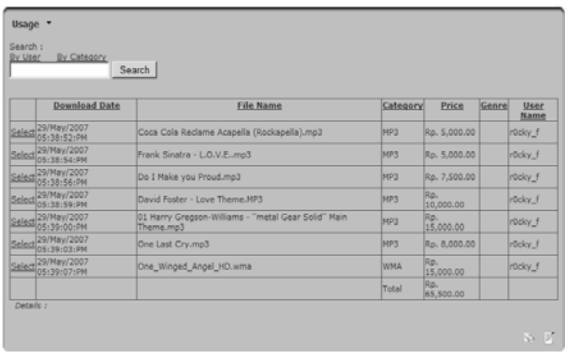


ID	Expired Date	Value
587296541	30/Jun/2007 12:00:00:AM	Rp. 50,000.00
	Total	Rp. 50,000.00

Figure 13. Module Voucher history - EXPIRED Voucher

5.6. Monitoring Download History

On this page, administrators can view the logs of all transactions that were done by the user. There are two ways that can be used to see the history, according to a registered username or by category files. Figure 14 shows the download page history.

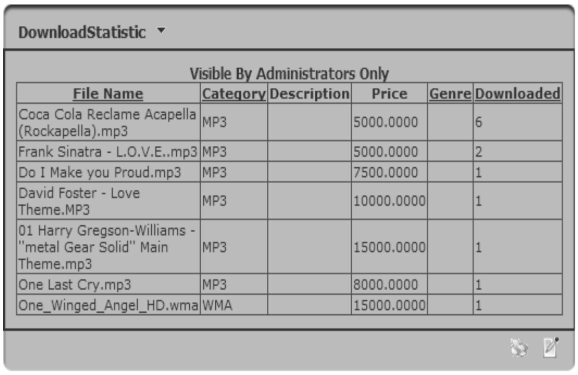


	Download Date	File Name	Category	Price	Genre	User Name
Select	29/May/2007 05:38:52:PM	Coca Cola Reclame Acapella (Rockapella).mp3	MP3	Rp. 5,000.00		r0cky_f
Select	29/May/2007 05:38:54:PM	Frank Sinatra - L.O.V.E..mp3	MP3	Rp. 5,000.00		r0cky_f
Select	29/May/2007 05:38:56:PM	Do I Make you Proud.mp3	MP3	Rp. 7,500.00		r0cky_f
Select	29/May/2007 05:38:59:PM	David Foster - Love Theme.MP3	MP3	Rp. 10,000.00		r0cky_f
Select	29/May/2007 05:39:00:PM	01 Harry Gregson-Williams - "metal Gear Solid" Main Theme.mp3	MP3	Rp. 15,000.00		r0cky_f
Select	29/May/2007 05:39:03:PM	One Last Cry.mp3	MP3	Rp. 8,000.00		r0cky_f
Select	29/May/2007 05:39:07:PM	One_Winged_Angel_HD.wma	WMA	Rp. 15,000.00		r0cky_f
			Total	Rp. 65,500.00		

Figure 14. Download history page

5.7. Monitoring Download Statistic

The administrator could also view on the users download statistics. From this page, it can be seen how many times a file has been downloaded. This information will become a parameter of the popularity of a file and can be taken into consideration for the next policy. Figure 15 shows the download page statistic.

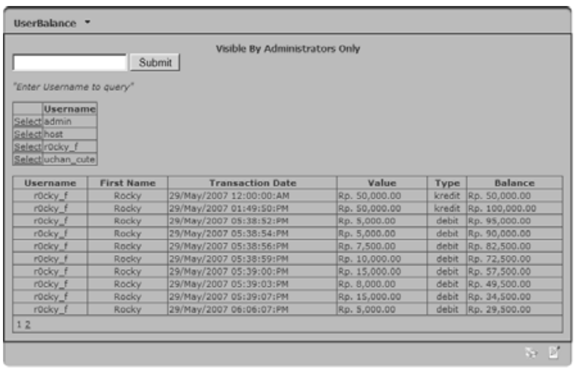


File Name	Category	Description	Price	Genre	Downloaded
Coca Cola Reclame Acapella (Rockapella).mp3	MP3		5000.0000		6
Frank Sinatra - L.O.V.E..mp3	MP3		5000.0000		2
Do I Make you Proud.mp3	MP3		7500.0000		1
David Foster - Love Theme.MP3	MP3		10000.0000		1
01 Harry Gregson-Williams - "metal Gear Solid" Main Theme.mp3	MP3		15000.0000		1
One Last Cry.mp3	MP3		8000.0000		1
One_Winged_Angel_HD.wma	WMA		15000.0000		1

Figure 15. Download Statistic Page

5.8. User Balance Log

Figure 16 shows the appearance of balance user logs. On this page, the administrator can view information about the transaction that has been done by the user from the beginning. To view the transaction history of users, the administrator must perform 2 step. The first stage is the administrator must determine the username of the user to view transactions. Username search conducted by typing a keyword (username initials) on the textbox that was provided. After that, a list of usernames that meet these criteria will be appeared. If the administrators do not enter keywords into the textbox, then all registered username will appear. After that, select one of the username and the and the history of the transaction of the user will be displayed.



	Username
Select	admin
Select	host
Select	r0cky_f
Select	uchan_cute

Username	First Name	Transaction Date	Value	Type	Balance
r0cky_f	Rocky	29/May/2007 12:00:00:AM	Rp. 50,000.00	kredit	Rp. 50,000.00
r0cky_f	Rocky	29/May/2007 01:49:50:PM	Rp. 50,000.00	kredit	Rp. 100,000.00
r0cky_f	Rocky	29/May/2007 05:38:52:PM	Rp. 5,000.00	debit	Rp. 95,000.00
r0cky_f	Rocky	29/May/2007 05:38:54:PM	Rp. 5,000.00	debit	Rp. 90,000.00
r0cky_f	Rocky	29/May/2007 05:38:56:PM	Rp. 7,500.00	debit	Rp. 82,500.00
r0cky_f	Rocky	29/May/2007 05:38:59:PM	Rp. 10,000.00	debit	Rp. 72,500.00
r0cky_f	Rocky	29/May/2007 05:39:00:PM	Rp. 15,000.00	debit	Rp. 57,500.00
r0cky_f	Rocky	29/May/2007 05:39:03:PM	Rp. 8,000.00	debit	Rp. 49,500.00
r0cky_f	Rocky	29/May/2007 05:39:07:PM	Rp. 15,000.00	debit	Rp. 34,500.00
r0cky_f	Rocky	29/May/2007 06:06:07:PM	Rp. 5,000.00	debit	Rp. 29,500.00

Figure 16. User's balance log

5.9. Previewing File

Administrator could perform a preview on every file that has been inserted by pressing the button "SELECT" on the related files. After that, a detailed information will appear at the bottom of the module files as shown in Figure 17.

Figure 17. Preview File

6. Conclusion

From the testing result, some conclusions could be drawn, namely:

1. Server applications that was built based on DotNetNuke CMS could manage contents and vouchers well and also display a preview of the content.
2. Server applications can display all the information regarding to users activity, namely: transactions both debit and credit transactions and the date of the transaction.
3. Application server has an ability to display content access statistics.
4. Application server could show the history of the vouchers that either have been used or expired.

References

[1] Bellanger P., Biba, K., 2007, 802.11n Delivers Better Range, http://www.wifiplanet.com/tutorials/article.php/3680781, accessed date : 29 October 2009.

[2] Darie, C., Ruvalcaba, Z. 2006, *Build Your Own ASP.NET 2.0 Web Site Using C# & VB* (2nd ed), SitePoint Pty.Ltd, Collingwood.

[3] Egan, D.N., Michael A.W, Steve V., 2006, *Building Websites with VB.NET and DotNetNuke 4,* Packt Publishing Ltd, Birmingham.

[4] Evjen, B., et al., 2006, *Professional ASP.NET 2.0.*, Wiley, Indianapolis.

[5] Lowe, D., 2006, *ASP.NET 2.0 Everyday Apps For Dummie*, Wiley, Indianapolis.

[6] Walker, Shaun, et al., 2006, *Professional DotNetNuke 4.0 – Open Source Web Application Framework For ASP.NET 2.0.*, Wiley, Indianapolis.

IMPROVED FLOODING PROTOCOL WITH GRAVITY ANALOGY IN WIRELESS SENSOR NETWORK

Junita[1], Karyono[2], Byungho Kim[3], Ihan Martoyo[4], Jinchun Kim[5]

[1]Undergraduate Program in Electrical Engineering, University of Pelita Harapan
[2,4]Graduate Program in Electrical Engineering, University of Pelita Harapan
[3,5]Dept. of Computer Engineering, Kyungsung University

[1]junita@uph.edu, [2]karyono@unimedia.ac.id, [3]bkim@ksu.ac.kr, [4]ihan_martoyo@uph.edu, [5]jckim@ks.ac.kr

Abstract

The Wireless Sensor Network (WSN) is a promising networking system. The superiority of WSN is on the size and cost of its nodes. The WSN has small size and low cost nodes. Despite of its superiority, these nodes have critical limitation in their resources. Thus, we need to design a simple, robust, and efficient routing protocol to be implemented in the WSN. In this paper, we present Gravity Based Flooding protocol. The main purpose of this paper is to improve the Flooding protocol, which is the simplest routing protocol in WSN, and the primitive building block for many other protocols. The proposed routing protocol improves Flooding protocol by adopting and modifying the gradient based routing approach from GRAB. The transmission of sensing information will be guided to the sink based on the gravity value that each node has. This gravity value represents the minimum hop count value to the sink and the data traffic condition of each node. We can achieve higher efficiency in network resource consumption by minimizing the redundant data and by using a well distributed power consumption mechanism. To provide the quantitative evaluation and animation process of Gravity Based Flooding protocol, we use OMNeT++(Objective Modular Network) as the simulation tool.

Keywords: WSN, routing, flooding, gradient, gravity, OMNeT

1. Introduction

The Wireless Sensor Network has the same basic characteristic as Wireless Ad-hoc Network (MANET). Both networks are automatically formed based on the condition of the nodes. The WSN has improved itself in size and cost of the nodes. The decrease in size makes the nodes more flexible to be implemented, and the low cost factor makes the WSN possible to cover a large area.

In order to make all the nodes work together and communicate with each other, the WSN needs to assign a routing protocol. This routing protocol must be efficient enough to be able to deal with WSN nodes' limitations. The WSN nodes have limited battery power, which limits the nodes' computation and communication capabilities

In this paper, the Gravity Based Flooding protocol for WSN is proposed. This routing protocol is an improvement of the Flooding protocol [1]. The Gravity Based Flooding protocol can maintain the simplicity of Flooding protocol while minimizing the overhead. The Gravity Based Flooding protocol bases its activity by using the gradient values. Adapted from the Gravity routing protocol [2], the Gravity Based Flooding protocol counts the gradient values, which are also called the gravity values, based on the hop distance to the sink and the number of sensed, received, and transmitted sensing information.

This paper is organized as follows. Section 2 discusses some related works that become the basic idea of the Gravity Based Flooding protocol. In Section 3, the algorithm behavior of the proposed routing protocol will be explained. In Section 4, the Gravity Based Flooding protocol will be analyzed by comparing it with the Flooding and Gossiping protocols. Finally, the conclusions and further research possibilities are discussed in Section 5.

2. Related works

Many routing protocols have been proposed to handle the overcome data transmission in WSN. Some of those protocols that are related with the Gravity Based Flooding protocol are:

2.1 Flooding Protocol

The main advantage of the Flooding protocol is on its simplicity. Each node that receives sensing information simply forwards the sensing information to all of the node's neighbors until the sensing information reaches the destination or sink.

2.2 Gossiping Protocol

As proposed in [1], the Gossiping protocol transmits the sensing information only to one node that the receiver node chooses randomly. The receiver node is a node that receives the sensing information. Gossiping protocol develops a single-route transmission. The single-route transmission can minimize the overhead in the network by minimizing unnecessary transmission processes.

2.3 GRAdient Broadcast Protocol (GRAB)

The basic idea of GRAB that is proposed in [3] is to deliver sensing information to the sink with the direction of a descending gradient. The assignment of the gradient value is initiated and maintained by the sink and memorized by each node in the network. In GRAB, the decision whether a node should participate in forwarding the sensing information or not is made by the receiver node. The receiver node forwards sensing information only if its own cost is lower than the previous sender node.

2.4 Gravity Routing Protocol

Adapted from SPIN [4][5], Gravity routing protocol uses data negotiation in performing its algorithm. The gravity value in the Gravity routing protocol represents the initial gravity value that is provided by the sink, and added by the gravity value that the node earns or lost from receiving and transmitting sensing information. Similar to the gravity law, the sensing information flows from the node with highest gravity value to the node with the lowest gravity value.

3. Gravity Based Flooding Protocol: Improved Flooding Protocol with Gravity Analogy in Wireless Sensor Network

To accept the challenge of designing a self configuring, simple, and robust routing protocol for WSN, the Gravity Based Flooding protocol is proposed. In designing this routing protocol, the advantages of other well known routing protocols are added and combined.

3.1 Gravity Based Flooding Protocol Background

The Flooding protocol is chosen for this paper's study and comparison because in Wireless Sensor Networks, the Flooding protocol is a primitive building block for many other protocols. The flooding algorithm is a fundamental network controlling system for executing some tasks, such as notifying all nodes of some information or getting information from all nodes, with a high reliability. Basically, all gradient based routing approaches still make use of the Flooding protocol as their basic transmission method. Thus, the goal of this paper is to design a routing protocol that as simple as Flooding, but more efficient than Flooding.

Among the conventional and the simple routing protocols, Flooding is chosen over Gossiping because of its robustness. Compared to single-route mechanism, multi-route mechanism can keep its robustness in a simpler way. In the single-route mechanism, a routing protocol needs to add another mechanism in order to confirm and make sure the receiver node receives the sensing information properly. Such confirmation needs an acknowledgement and addressing system, which could be complicated. If the receiver does not receive the sensing information, the source node has to re-transmit the information. Not only will it cost more in terms of power consumption, but also in terms of delay time.

Even though the Flooding protocol has the advantages of robustness and parallel transmission, it has the disadvantage of implosion. This disadvantage can be solved by using the gradient based routing approaches. These approaches overcome the implosion by directing the transmission of sensing information to the sink using descending value of gradient. By doing this, unnecessary transmission can be minimized. From the existing gradient based routing approaches, the GRAB routing protocol mechanism is chosen to be adapted. This approach is able to minimize the implosion by limiting the number of nodes that participate in sensing information transmission, while maintaining the reliability of the network by providing high possibility for multi-route, without the necessity of acknowledgement system.

To make the GRAB routing protocol mechanism more efficient but still simple, an adaptive gravity system [2] is implemented. An implementation of adaptive gravity system causes the network to be aware of traffic transmission in the network, thus creating well distributed power consumption for the entire nodes in the network.

This routing protocol can be implemented in a single-sink as well as in a multi-sink network environment. In this paper, the Gravity Based Flooding protocol is discussed in a single-sink network environment.

3.2 Initial Gradient Assign Phase

The Gravity Based Flooding protocol has 3 phases in its processes. The initial gradient assign phase is the first phase in the Gravity Based Flooding protocol. This phase is initiated by the sink so that every node in the network can maintain its minimum hop count value to the sink, which is the initial gradient of the node.

The sink initiates this phase by broadcasting the initial message that contains a hop count field and message ID. First, the sink sets the value in hop count field into 0, and then broadcasts the initial message to the sink's entire neighbor nodes. When a node receives an initial message, the receiver node increments the value in the hop count field by 1 and takes that new hop count value as its initial gradient. Afterwards, the receiver node re-broadcasts the initial message with the new hop count value to its entire neighbor nodes.

There is more than 1 node involved in this initial gradient assign phase. The existence of the multi-nodes means each node in the network will receive more than one initial message. To prevent a node from broadcasting the initial messages every time the node receives the initial message, each node remembers the ID of the initial message when the node receives the initial message for the first time. So that when a node receives another initial message, the node can recognize and directly discards that initial message. Besides minimizing the redundant data in the network, by discarding the next received initial message, a better transmission route for each node can be maintained. There is a high possibility that the first arrived initial message contains the smaller hop count value compared to the next received initial message. By using message ID memorizing mechanism, the initial gradient phase will keep running only until all nodes in the network get their initial gravity values. The initial message will disappear from the network because all the remaining initial messages will eventually be discarded by the nodes.

3.3 Gradient Assign Phase

The gradient assign phase is an adaptive process from Gravity routing protocol. The gravity value of each node not only contains the initial gravity value from the initial gradient assign phase. A node also adopts its gravity values from the amount of the sensing information that the node senses, transmits, and receives. To give a better guidance for the transmission flow of the sensing information, and to create a more efficient routing protocol, the amount of sensing information is added into the gravity value. Each time a node senses an event or receives the sensing information, the node's gravity value increases by one. And each time a node transmits a sensing information to the other nodes or sink, the node's gravity value decreases by one.

3.4 Decision Phase

When a node senses an event, the node transforms that event into sensing information, then increments the node's gravity value by one, and broadcasts the sensing information to the node's entire neighbors. The sensing information contains data and 2 header values. These header values are the sender nodes gravity value and the sensing information ID.

Adapted from GRAB, in this phase, the receiver node decides whether to forward the sensing information or to discard it. The receiver node forwards the sensing information if the sender's gravity value contained in the sensing information header is higher than the receiver's gravity value. On the other hand, if the sender's gravity value is the same as, or lowers than the receiver's gravity value, the receiver node discards the sensing information. This way of making decision prevents the information transmission from flowing to the route that is farther from the sink.

Similar to the initial gradient assign phase, the information transmission in the decision phase also runs in the flooding mode with information ID memorizing mechanism. This memorizing minimizes the redundancy of information transmission in a very significant way. The ID memorizing mechanism prevents a node from broadcasting the same sensing information.

4. Analysis and Result

To provide the quantitative evaluation of the Gravity Based Flooding protocol, OMNeT++ (Objective Modular Network) is used [6]. OMNeT++ is one of the well established simulation tools with an Eclipse-based IDE and a graphical runtime environment. This simulator is an open-source, component-based simulation package, which is designed for modeling communication networks, multi-processors systems, or distributed systems.

A comparison is done between the Gossiping protocol, the Flooding protocol, and the Gravity Based Flooding protocol. The Gossiping protocol is included in the comparison to provide quantitative evaluation in terms of the reliability and the time efficiency between a single-route routing protocol and a multi-route routing protocol.

The full mesh network structure developed in [7] is used as the starting point for developing the simulation. For a precise comparison and analysis, instead of using the full mesh network structure, a fixed networking structure is used.

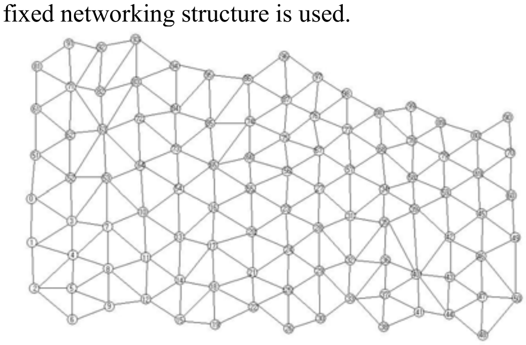


Figure 1. Fixed Network Structure

The performance comparison between the Gossiping, Flooding, and Gravity Based Flooding protocol is evaluated by simulating these routing protocols using different number of nodes, and by assigning different source node at each simulation.

For the number of nodes, 10 levels of values are used, from 10 until 100 with the increment of 10. 10 levels of values are used to represented variety size of networks, from the small area to the large area of networks. The increment of 10 is used to change the size of the network without causing a disorder in the analysis of the simulation results.

For each number of nodes, the algorithm of the Gossiping protocol, the Flooding protocol, and the Gravity Based Flooding protocol are simulated twice for 2 different assigned source nodes. An assigned source node is a node which is scheduled to get the first sensing information from the simulation program. A node can be assigned as the source node by assigning the node's ID into the simulation program. The node's ID is a certain number which is uniquely assigned to each node at the time the structure of the network is built.

First, for each number of nodes, the algorithm of the Gossiping protocol, the Flooding protocol, and the Gravity Based Flooding protocol are simulated when the source node ID number is a half of the total number of the nodes in the entire network. Afterwards, the algorithm of the Gossiping protocol, the Flooding protocol, and the Gravity Based Flooding protocol are simulated again by assigning the last node as the source node. The last node is the node with the highest ID number. That way, a random position of the source node can be achieved systematically. The result of the algorithm simulation can be analyzed in a variety of cases, based on the source node position. It is important to know that even though the number of the source node ID is a half of the total number of the entire nodes in the network, that number does not indicate that the source node is in the middle of the network. As already mentioned before, the position of a node depends on the network structure. The importance of the source node position will be explained later using the quantitative evaluation.

The node with the ID number 0 is chosen as the sink. This node is chosen because it is located at the opposite end of the line of the entire assigned source node. The analysis is done by computing the trace result from the simulation process in OMNeT++. The focus of analysis will be on the time efficiency and the power efficiency factor.

4.1 Time Efficiency

In the simulation, the value of time is counted using simulation time, which is based on how many events that occur in the network. If some events are done simultaneously, they counted as 1 simulation time. For example, in the Flooding protocol when many nodes broadcast the sensing information simultaneously, that will be counted as 1 simulation time. On the contrary, the Gossiping protocol does not display simultaneous events. The sensing information travels from one node to another in the network. Thus, without additional routing information, the Gossiping protocol needs more time to finish the simulation compared to the Flooding protocol.

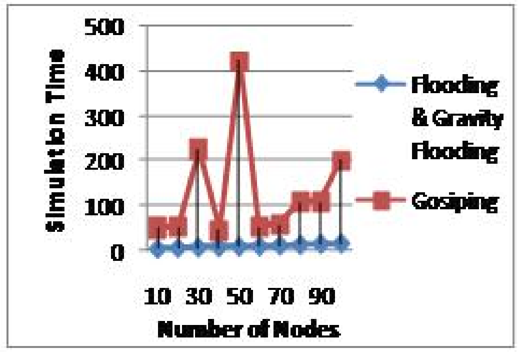


Figure 2. Simulation time comparison when the assigned source node ID number is half of the total number of nodes

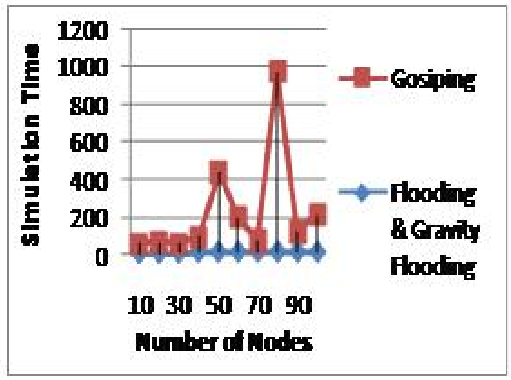


Figure 3. Simulation Time comparison with the last node as the source node

From Figure 2 and 3, it can be seen that the simulation of the Gossiping algorithm has a random simulation time. The Gossiping protocol chooses the next node randomly. If the source node in the Gossiping protocol chooses a node that is farther from the sink, then the simulation time will take more time.

This comparison is a rough comparison because the simulation program does not include the congestion delay and the computation delay into a count. In the actual network there will be congestion for both, Flooding and Gravity Based Flooding, because these routing protocols broadcast many data simultaneously.

The simulation result also shows that the Flooding protocol and the Gravity Based Flooding protocol have the same time to finish the simulation. This simulation result is not precise because the Gravity Based Flooding protocol takes a little more time compared to the Flooding protocol. The Gravity Based Routing protocol performs an extra computation every time a node receives sensing information. Also, the time that the Gravity Based Flooding simulation needs for the initial gradient assign process is not included because this initialization process does not take place during transmission phase. Thus, the algorithm improvement of the Gravity Based Flooding does not affect the time efficiency in a significant way.

4.2 Energy Efficiency

The energy efficiency is examined for the Flooding protocol and the Gravity Based Flooding protocol. The examination is based on the amount of the transmitted and the arrived sensing information of the Flooding protocol and the Gravity Based Flooding protocol. Besides transmitting sensing information, the Gravity Based Flooding protocol performs the initial message transmission. The number of the transmitted sensing information represents the cost in the network. For the initial message, it is assumed that each initial message only costs 10% of the sensing information, because the initial message is a small packet message that only contains 2 headers. The energy efficiency is computed by dividing the number of the arrived sensing information with the number of the transmitted sensing information.

The Gossiping protocol is not included in the comparison of the energy efficiency because the decision in the Gossiping protocol is made in a random way thus has no particular pattern.

Table 1. Table of Energy Efficiency Comparison When The Assigned Source Node ID Number is Half of The Total Number of Nodes

		Energy Efficiency(%)	Difference(%)
Flooding	N=10	5.13	0.14
Gravity Flooding	src=5	4.99	
Flooding	N=20	2.35	1.12
Gravity Flooding	src=10	3.47	
Flooding	N=30	1.43	0.42
Gravity Flooding	src=15	1.85	
Flooding	N=40	1.07	0.45
Gravity Flooding	src=20	1.52	
Flooding	N=50	0.84	0.44
Gravity Flooding	src=25	1.28	
Flooding	N=60	1.36	0.51
Gravity Flooding	src=30	1.87	
Flooding	N=70	1.15	0.15
Gravity Flooding	src=35	1.3	
Flooding	N=80	0.98	0.07
Gravity Flooding	src=40	1.05	
Flooding	N=90	0.86	0.05
Gravity Flooding	src=45	0.81	
Flooding	N=100	0.77	0.07
Gravity Flooding	src=50	0.7	

Table 2. Table of Energy Efficiency Comparison with The Last Node As The Source Node

		Energy Efficiency(%)	Difference(%)
Flooding	N=10	5.13	0.49
Gravity Flooding	src=9	4.64	
Flooding	N=20	2.35	0.22
Gravity Flooding	src=19	2.13	
Flooding	N=30	1.43	0.12
Gravity Flooding	src=29	1.31	
Flooding	N=40	1.07	0.1
Gravity Flooding	src=39	0.97	
Flooding	N=50	0.84	0.03
Gravity Flooding	src=49	0.87	
Flooding	N=60	1.36	0.02
Gravity Flooding	src=59	1.34	
Flooding	N=70	1.15	0.08
Gravity Flooding	src=69	1.07	
Flooding	N=80	0.98	0
Gravity Flooding	src=79	0.98	
Flooding	N=90	0.86	0
Gravity Flooding	src=89	0.86	
Flooding	N=100	0.77	1.46
Gravity Flooding	src=99	0.81	

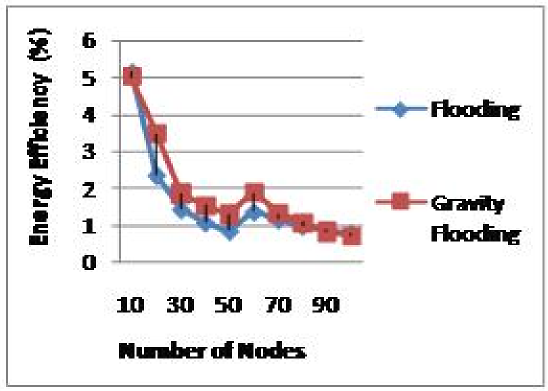


Figure 4. Energy Efficiency Comparison When Assigned Source Node ID Number is Half of The Total Number of Nodes

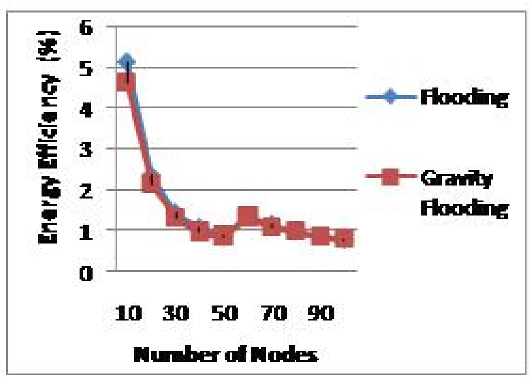


Figure 5. Simulation Time Comparison With Last Node as The Source Node

On the Table 1 and 2, the term N represents the total number of the nodes in the network, and term src represents the ID value of the source node. From the Table 1 and 2, also the Figures 4 and 5, it can be seen that the energy efficiency of the Gravity Based Flooding protocol is very dependent on the position of the source node. The closer the source node to the sink, the more efficient it will be. The Gravity Based Flooding is able to guide the sensing information transmission to flow toward the sink and not going backward to the nodes that are farther from the sink. The Flooding protocol transmits the sensing information to all nodes, without considering any direction. That is why Flooding is not affected by the position of the source node. Regardless of the source node position, the sensing information always transmitted to all nodes. If the source node is close to the sink, the Gravity Based Flooding is more efficient. It minimizes the redundant data by only transmitting sensing information towards the sink. But if the source node is not close enough to the sink, or if the source node has the highest gravity, the Gravity Based Flooding will be slightly less efficient than Flooding. Thus, In that case, similar to Flooding, it will transmit the sensing information to almost all the nodes in the network. The Gravity Based Flooding becomes less efficient than Flooding because of the initialization process that also consumes the energy of the node. But the initial message is small and it does not cost much. Thus, in general, the Gravity Based Flooding gives more efficiency than the Flooding protocol.

The Flooding and the Gravity Based Flooding have the same amount of arrived sensing information because both of them use the ID memorizing mechanism. The number of the sink's neighbor nodes represents the number of routes that can be formed by the network. Since ID memorizing mechanism is used, 1 route can only transmit the same sensing information once. Until the number of nodes reaches 50 nodes, the sink has only 2 neighbor nodes. As the network gets larger and wider, when the number of nodes reaches 60 nodes, the sink neighbor nodes increase into 4 nodes. These changes can be seen in Figure 4 and Figure 5 where sudden increase of energy efficiency happens as the number of nodes reaches 60 nodes.

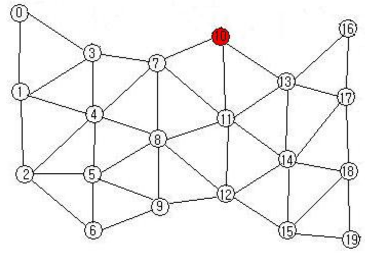


Figure 6. Network Structure With 20 Nodes, and Node 10 as The Assigned Source Node

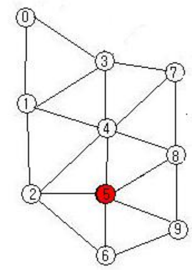


Figure 7. Network Structure With 10 Nodes, and Node 5 as The Assigned Source Node

When the WSN has a network structure as in Figure 6, the Flooding protocol produces more transmitted sensing information which is the redundant data compared to the Gravity Based Flooding protocol.

When the WSN has a small network as in Figure 7, the Gravity Based Flooding is less efficient than the Flooding protocol. The advantage of the Gravity Based Flooding is in term of producing less amount of the transmitted sensing information compared to the Flooding protocol. But, in a small network, the amount of the transmitted sensing information produced by the Gravity Based Flooding and the Flooding are relatively the same. Thus, the difference can not compensate the redundant data that the Gravity Based Flooding has from broadcasting initial message.

Gossiping is not included in the comparison of energy efficiency. Because the decision in Gossiping protocol is made in a random way. The Gossiping protocol does not have a particular pattern in producing the transmitted sensing information.

5. Conclusions

The Gravity Based Flooding protocol is an improved Flooding protocol that can be implemented in Wireless Sensor Network. This proposed routing protocol improves Flooding protocol using the implementation of an adaptive gravity system. With the implementation of an adaptive gravity system, the nodes in the network adopt their gravity values from the amount of the sensing information that the nodes sense, transmit, and receive. Instead of broadcasting the sensing information blindly, based on the gravity value, it can direct the flow of sensing information to the sink with a simple algorithm. By adopting the amount of the sensing information that the nodes sense, transmit, and receive into the nodes' gravity values, the network formed a well managed distribution of power consumption.

Through the comparison using simulation process in OMNeT++ simulator tool, the simulation result shows that Gravity Based Flooding protocol has improved Flooding protocol in terms of the energy efficiency. The largest the network and the closer the source node to the sink, the more efficient Gravity Based Flooding protocol will be. In a large network, the Gravity Based Flooding protocol produces less redundant data compared to the Flooding protocol.

References

[1] S. Hedetniemi and A. Liestman, "A survey of gossiping and broadcasting in communication networks", *IEEE* Networks,. Vol. 18, No. 4, pp. 319-349, 1988..

[2] Junita, Leonardy Setyawan, Ihan Martoyo, Byungho Kim, and Jinchun Kim, "Gravity: A Distributed Routing Protocol for Wireless Sensor Networks", *Korean Institute of Information Scientists and Engineers*, pp. 261-266, 2008.

[3] Fan Ye, Gary Zhong, Songwu Li, Lixia Zhang, "GRAdient Broadcast: A Robust Data Delivery Protocol for Large Scale Sensor Networks", *Wireless Networks* Vol. 11, pp. 285-298, 2005.

[4] J. Kulik, W. Rabiner, H. Balakrishnan, "Adaptive Protocols for Information Dissemination in Wireless Sensor Network", *ACM/IEEE Conference on Mobile Computing and Networking (MobiCom'99)*, pp. 174-185, 1999.

[5] J. Kulik, W. Heinzelman, H. Balakrishnan, "Negotiation-Based Protocols for Disseminating Information in Wireless Sensor Networks," *Wireless Networks*, Vol. 8, No. 2-3, pp. 169-185, 2002.

[6] OMNET++. Retrieved June 10, 2008, from the OMNeT++: http://www.omnetpp.org

[7] Karyono, "Simulasi Algoritma Gravity dan Flooding pada Wireless Sensor Network", *Thesis for Electrical Engineering Graduate Program, Universitas Pelita Harapan, Indonesia*, May 2009.

[8] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk, J. Anderson, "Wireless sensor networks for habitat monitoring", *ACM Int. Workshop on Wireless Sensor Networks and Applications*, pp. 88–97, 2002.

[9] I. F. Akyildiz and I. H. Kasimoglu, "Wireless sensor and actor networks: Research challenges", *Ad Hoc Networks Journal*, Vol. 2, No. 4, pp. 351–367, 2004.

[10] N. Sadagopan, B. Krishnamachari and A. Helmy, "Active Query Forwarding in Sensor Networks (ACQUIRE)", *SNPA*, 2003.

[11] Shijin Dai, Xiaorong Jing, and Lemin Li, "Research and analysis on routing protocols for wireless sensor networks", *Circuits and Systems*, 2005.

[12] J. Qiangfeng and D. Manivannan, "Routing protocols for sensor networks", *Consumer Communications and Networking Conference, CCNA*, 2004.

[13] K. Akkaya and M. Younis, "A Survey on Routing Protocols for Wireless Sensor Networks", *Ad Hoc Networks*, Vol. 3, pp. 325-349, 2005.

AN APPLICATION TO SUPPORT TDM-TO-SOFTSWITCH-BASED NGN MIGRATION ON OPNET IT GURU® NETWORK PLANNER

Kurniawan Eko Putro [1], Fazmah Arif Yulianto [2], Angkoso Suryocahyono [3]

[1,2]Faculty of Informatics, Institut Teknologi Telkom
[3] R&D Center, PT. Telekomunikasi Indonesia
[1]**wawan_smansa@yahoo.com**, [2]**fay@computer.org**, [3]**angkoso@telkom.co.id**

Abstract

OPNET IT Guru® Network Planner is equipped with amazing capability to do network planning either for TDM-based network or Softswitch-based NGN (Next Generation Network). This capability has been realized beneficial for many service providers when conducting network planning for their network infrastructure. However, in some cases, service providers require not only the ability to do planning but also for doing migration from TDM-based to Softswitch based NGN. At this moment, OPNET IT Guru® Network Planner facilities migration by manual way; this is by deleting TDM-based switch objects and creating new Softswitch object, either on canvas or through the exported XML file and then adjusting the link one by one. Performing migration by manual way is time consuming and vulnerable to error. Therefore, existence of application for automating replacement of TDM-based objects by Softswitch objects can be very useful. The application reads and modifies the XML file which consists of TDM-based network topology that is exported by OPNET IT Guru® Network Planner. The modified version of the XML file is then imported into canvas forming the new topology which is now Softswitch-based. This new Softswitch-based topology is now ready to support further processes, such as voice quality, bandwidth requrement, capacity planning and flow analysis.

Keywords: OPNET, TDM, NGN, XML, softswitch, migration

1. Introduction

Facing the multimedia and global information era, there are huge demands on telecommunication networks to support more complex and various services. Existing PSTN-based infrastructures are no longer enough for this particular purpose. Here comes Next Generation Networks (NGN).

Existing PSTN-based networks need to be migrated to NGN-based ones. The migration itself must be well-planned in order to reduce cost and optimize the result. In most cases, the planning phase will include the use of some sort of modeling tools.

One of the most popular modeling tools is OPNET IT Guru® Network Planner. This software helps the network designer in modeling and simulating new network design. Despite of its powerful features, there is a limitation in OPNET related to a process of migration of a PSTN switch into a softswitch and a process of collocation of some PSTN switch to a softswitch. OPNET has no function to enable automatic migration and collocation. Those two processes must executed manually by the network designer.

Manual migration and collocation are time consuming and subject to human error. The complexity arises when there are a lot of switch to migrate, as in the case of PT. Telkom [1]. To overcome these problems, we developed an application to do migration and collocation automatically.

2. Observation

2.1 The Migration to Softswitch-based NGN

Softswitch-based NGN has been a main target for some telecommunication company, including PT. Telekomunikasi Indonesia. PSTN switch still dominating current telecommunication network in Indonesia. To migrate them into softswitch-based ones, there are some important considerations. First, from the functional point of view, existing PSTN switch will be replaced by softswitch. We call this a replacement. Second, there will be a reduction in the number of switch elements. In other words, some of PSTN switch will be collocated into a single softswitch (collocation process).

2.2 OPNET's XML files

We can export the current network topology in OPNET IT Guru ® Network Planner. The files are in XML format. These files will be our media and interface between OPNET and our own application.

Based on our research, there are five main elements (objects) in the XML file: profile, subnet, node, link, and demand. Every object has some special attributes to describe the details of a particular object.

2.3 Manual Replacement and Collocation

The procedure of manual replacement is as follows. First, delete the PSTN switch which will be replaced by a softswitch. As a consequence of this action, all links that directly connected to that switch will also deleted. The next step will be an insertion of softswitch from object palette, rebuild the links, and define the subnet containing NGN elements, such as media gateway, access router, and backbone router.

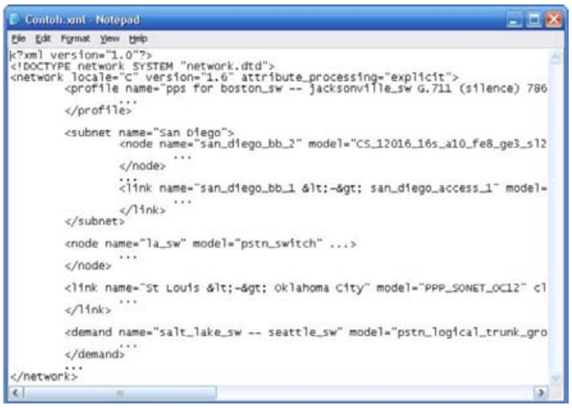


Figure 1. An example of OPNET's export file in XML format

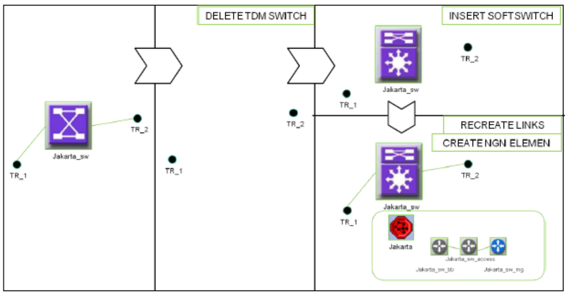


Figure 2. Manual Replacement procedure

When we want to manually collocates switch into a softswitch, the switch will be deleted from network topology. All nodes that previously connected to that switch will be handled by a softswitch. All links previously attached to the deleted switch will be attached to softswitch.

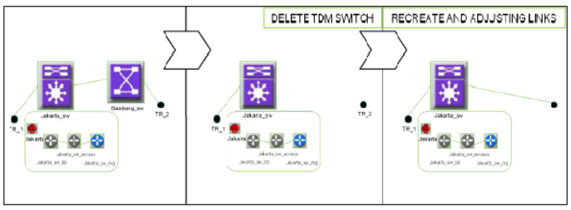


Figure 3. Manual Collocation procedure

3. Application Design

The main idea of our application is to do some modification on OPNET's exported file (in XML format). To resemble manual replacement process, our application will change node's attributes from PSTN switch into Softswitch. To do a collocation, the application deletes some PSTN switches and includes their function into a softswitch.

Later, we can import the XML file back into OPNET to have a softswitch-based NGN topology. Further analysis on the new topology, e.g. simulating current network situation, network readiness, voice quality, bandwidth requirements, capacity planning etc., then can be done using OPNET.

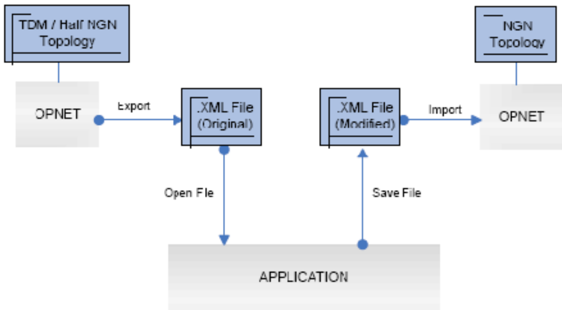


Figure 4. The Application – OPNET Interfacing

The overall process starts with pre-processing of OPNET's exported file. The application parses the XML file to separate one object from the others. The parsing went further to extract the information from object's attributes. To organize the result we use database system process called XML to dB.

The application manipulates this database to do replacement and collocation processes. When all migration processes finished, application will execute a dB to XML process. The result of all modification during the migration processes will now saved in this new XML file. The new XML containing softswitch-based NGN topology is ready to imported back into OPNET.

4. Evaluation

The goals of our evaluation are: to make sure that the results of our automatic migration processes (migration and collocation) are correct, and to get a user experience feedback. A new network topology from the automatic process using our application will be compared with the one from OPNET (using manual procedures), both visually and technically (comparing the key attributes in the XML file).

4.1 Evaluation scenarios

We have three main scenarios, each representing easy, medium, and complex scenario. The first scenario is a simple replacement and collocation. The second one consists of replacement of two connected softswitches and collocation of some links. The third scenario involves a replacement of a full mesh demand, a collocation of switches with many links, and handling of duplicated links.

The evaluation held at Service Node Lab of PT. Telkom's R&D Center Bandung with the assistance of some experts in network modeling and planning using OPNET. We use the manual migration process as the baseline.

Starts from the same source topology (Figure 5), we do some exact migration processes using OPNET and our application. The topology exported into an XML file, which then loaded into our application (Figure 6). Using the same scenario and certain rules, we want to make sure that the comparison is relatively fair.

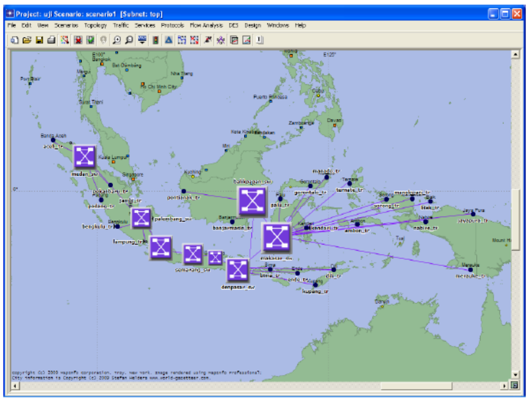


Figure 5. Source topology (PSTN switches)

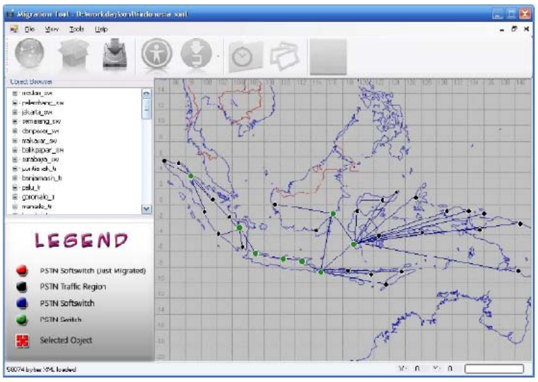


Figure 6. The source topology loaded in our application

Here are the details of our scenarios:

- Scenario 1:
  - Replace the Jakarta node with a subnet consist of NGN elements
  - Collocate Semarang node to Jakarta
- Scenario 2:
  - Replace Jakarta node with a subnet consist of NGN elements
  - Replace Semarang node with a subnet consist of NGN elements
  - Collocate Palembang node to Jakarta
- Scenario 3:
  - Replace Denpasar node with a subnet consist of NGN elements
  - Replace Surabaya node with a subnet consist of NGN elements
  - Replace Jakarta node with a subnet consist of NGN elements
  - Collocate Makasar node to Denpasar

4.2 Results

For all three scenarios above, the new topologies from the migration processes using our application are exactly the same as the ones from executing manual migration processes using OPNET. Below are some screenshots showing that either visually or technically (the detail of XML files) the automatic migration using our application resulting a same NGN topology with what OPNET does.

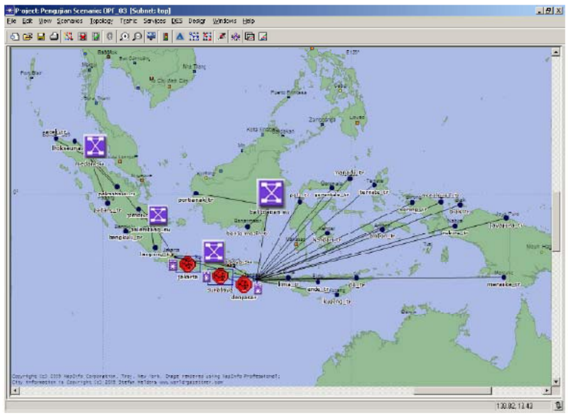


Figure 7. NGN topology of the 3rd scenario after manually migrated using OPNET

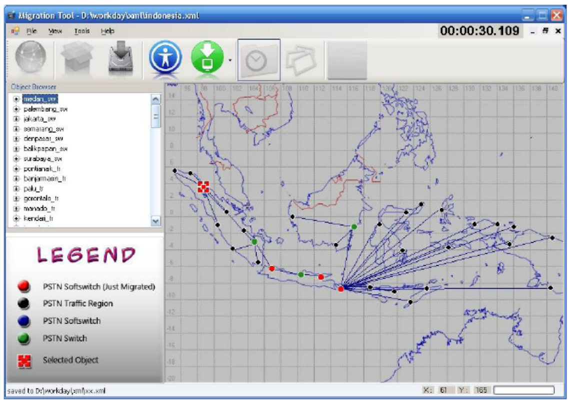


Figure 8. NGN topology of the 3rd scenario after automatically migrated using our application

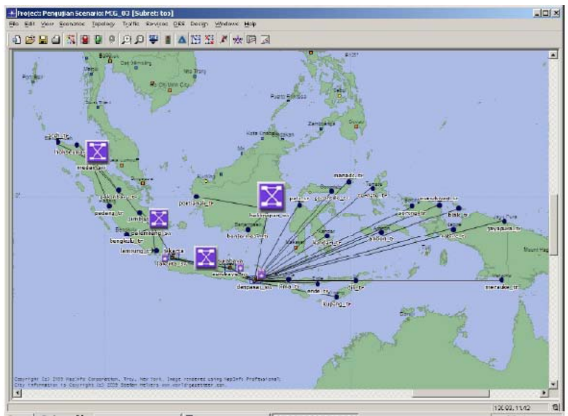


Figure 9. NGN topology from our application imported into OPNET

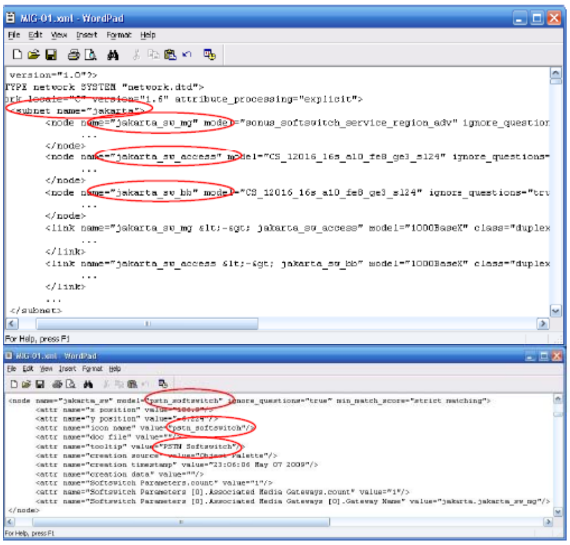


Figure 10. The modification of XML file either by using OPNET or our application

Two XML files from OPNET and from our application are identical, in term of the key attributes have the same values. The occurrence of objects in these files are not necessarily the same.
We count the number of mouse click required to accomplish the migration processes in all scenarios using OPNET and our application. The result (figure 11) shows that our application needs less click than OPNET.

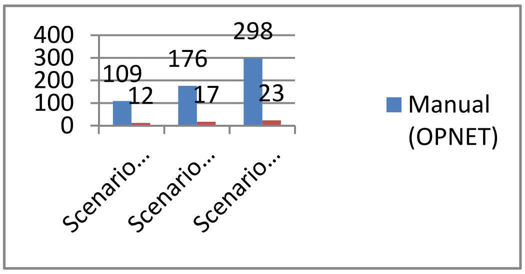


Figure 11. Number of human-computer interaction to do migration (measure: number of mouse click)

Do the migration using OPNET operated by experts (PT. Telkom's employees whose use OPNET in their daily activity) take much more time than doing the same task using our application.

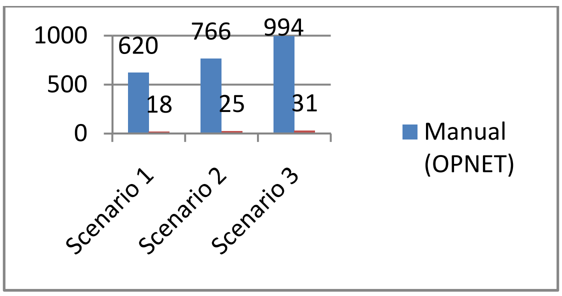


Figure 12. Migration time (in seconds)

5. Conclusion

Doing the migration from PSTN-based telecommunication topology into softswitch-based NGN topology automatically is our application's main goal. The result is exactly the same as the one when we do the process manually using OPNET.

The automatic processes using our application are much simpler than the manual one, has a lower human-error risk, and take less time to complete.

6. Future Works

Further research and development to complete the feature of our application, e.g. to detect which object should be replaced depends on some user defined rules.
With the collaboration with OPNET Technology, Inc. we can embed these functions into OPNET IT Guru.

References

[1] Suryocahyono, A., Widyowibowo, O. (2008, August 29). *Modeling and Simulation of Add-on Application for Supporting Planning on TDM-to-Softswitch-based NGN Migration*. Washington, D.C..: OPNETWORK 2008

[2] International Packet Communication Consortium. Packet Communication Refference Architecture. www.packetcomm.org, accessed on 9 Februari 2009

[3] Ludfy, Akhmad. 2003. *Sekilas Tentang Next Generation Network (NGN)*. Jakarta: Gematel

[4] Munawar. 2005. *Pemodelan Visual dengan UML*. Jakarta: Graha Ilmu

[5] Ohrtman, Franklin D. 2003. *Softswitch Architecture for VoIP*. New York: McGraw-Hill

[6] OPNET Technologies, Inc.. 2008. *Introduction to Using IT Guru Network Planner*

[7] OPNET Technologies, Inc.. 2008. *Introduction to VoIP*

[8] OPNET Technologies, Inc.. 2008. *Planning and Analyzing VoIP Migrations for Service Providers*

[9] Suryocahyono, A. (2008, August 29). *Planning on on TDM-to-Softswitch-based NGN Migration by OPNET IT Guru® Network Planner*. Washington, D.C..: OPNETWORK 2008

[11] Wahana Komputer. 2008. *Cepat Menguasai Visual Studio .NET 2008 Express*. Semarang: Andi Offset

[12] Widodo, Aji. 2002. *Softswitch Konsep Jaringan Telekomunikasi Masa Depan*. Jakarta: Gematel

A COMPARISON BETWEEN FUSC AND PUSC SUB-CHANNELIZATION TECHNIQUES FOR DOWNLINK MOBILE WiMAX IEEE 802.16e PERFORMANCE

Nomarhinta Solihah[1], Heroe Wijanto[2], Budi Prasetya[3]

[1,2,3]Institut Teknologi Telkom
[1]**nomarhintas@gmail.com**, [2]**hrw@stttelkom.co.id**, [3]**bpy@stttelkom.ac.id**

Abstract

Broadband Wireless Access (BWA) is a technology which advantages are wide coverage services, large capacities and good quality service. Nowadays, BWA technology has various kinds of superior products and one of them is Worldwide Interoperability for Microwave Access (WiMAX). WiMAX has evolved from fixed WiMAX to mobile WiMAX. Mobile WiMAX, which standard is IEEE 802.16e, is not only use for mobile user but also Non Line of Sight (NLOS) channel. The reason is some of innovative technology applications have supported IEEE 802.16e in order to cope the problem. One of the innovative technologies is Sub-Channelization. Sub-Channelization used in mobile WiMAX is Full Usage Subcarrier (FUSC) and Partial Usage Subcarrier (PUSC). This paper analyzes the comparison performance between sub-channelization technique FUSC and PUSC in downlink mobile WiMAX. Simulation is done in different number of user and velocity of user. The result of simulation shows that system with different user velocity (0, 3, 30 and 120 km/hour), the faster user, the higher error bit value. FUSC give BER repair ± 50% for same Eb/No value. Meanwhile, system with different number of user (1, 4, 16 and 32 users), FUSC can give performance correction for ± 0.37 dB. Because, FUSC process is subcarrier permutation while PUSC process are cluster permutation and subcarrier permutation. Thus, process period of PUSC is longer than FUSCs. Hence, Eb/No of PUSC is greater than FUSC. In conclusion, FUSC give better performance for different number of user and user velocity.

Key Words: Mobile WiMAX, Sub-Channelization, FUSC, PUSC, Downlink.

1. Introduction

In recent years, people call for improvement in the service quality of multimedia (voice, video, data, and picture) communication so that there is a technology named Broadband Wireless Access (BWA). BWA comes along with the increasing of broadband access that needs wide bandwidth and large capacity. However, BWA still has problems to solve and one of them is Non Line Of Sight (NLOS) channel condition. Furthermore, NLOS will cause multipath fading that causes great delay spread and Inter Symbol Interference (ISI).

One of BWA's product is Worldwide Interoperability for Microwave Access (WiMAX). Accomplishing the need of telecommunication, WiMAX has evolved from fixed WiMAX to mobile WiMAX. Mobile WiMAX, which standard is IEEE 802.16e, is suitable for NLOS channel condition. It is because mobile WiMAX is supported by some innovative technologies, for instance OFDM (Orthogonal Frequency Division Multiplexing) and sub-channelization. Sub-channelization used in mobile WiMAX are Full Usage of Subcarrier (FUSC) and Partial Usage of Subcarrier (PUSC).

The use of sub-channelization has different effect on the performance of mobile WiMAX because both of FUSC and PUSC have different process. Even though these are a subcarrier permutation method that subcarriers are distributed throughout the available spectrum, these still have attraction to be observed. Despite this reason, this paper will analyze how these sub-channelization affect the performance if there are different number of user and user velocity. The process will be done based on standard that mobile WiMAX has, IEEE 802.16e. The simulation is done in different user number (single user and multi user) and velocity (fixed and mobile).

2. System Description

Mobile WiMAX, a Broadband Wireless Access product, is a development technology from fixed WiMAX. Its standard is IEEE 802.16e that supports high-speed data access, security, service quality and mobility. Solving NLOS problem, mobile WiMAX is equipped with OFDM technology, sub-channelization, directional antenna, transceiver diversity, adaptive modulation, error channel correction, and power control [2]. Besides, it supports not only channel bandwidth 1.25 Mhz to 20 MHz but also 2048 number of subcarrier.

OFDM is designed to overcome NLOS problem because it works on multicarrier transmission concept. This concept will divide data stream into some parallel sub streams. Then, it will be placed at subcarrier frequency that is closed and orthogonal. After that, it is transmitted serially. This concept also can change the speed of data, from high speed data to low speed data. The reason is subcarrier period will be greater to overcome delay spread that is happened due to NLOS channel condition.

Transmitting many users, mobile WiMAX uses OFDMA in physical layer. OFDMA is created by using multiple access scheme to transmit many data users both downlink sub-channel and uplink sub-channel. In a simple, OFDMA will divide a group of subcarrier to be allocated for many users. In the process of OFDMA, it adds cyclic prefix to give immunity caused by multipath and to avoid error synchronization.

Besides, OFDMA has sub-channelization. It is defined as a group of subcarrier allocated in frequency spectrum. A sub-channel is consisted of data subcarrier, pilot subcarrier and null subcarrier. A sub-channel is consisted of data subcarrier, pilot subcarrier and null subcarrier.

The distribution of subcarriers in a sub-channel is called subcarrier permutation. These subcarriers have not to be adjacent or even it can be sprout in all frequency bands. That is why there are two methods of subcarrier permutation; diversity subcarrier permutation and adjacent subcarrier permutation.

Diversity subcarrier permutation distributes subcarriers randomly, for instance, Full Usage Subcarrier (FUSC) and Partial Usage Subcarrier (PUSC). Its advantages are frequency diversity and inter cell interference. Besides, it can reduce the use of same subcarrier in a sector or cell. In other hand, it is difficult to estimate channel. Meanwhile, adjacent subcarrier permutation is a group of contiguous subcarrier. The example is Adaptive Modulation and Coding (AMC). This method needs the best condition of bandwidth but it is easy to estimate channel because of adjacent subcarrier.

2.1 Downlink FUSC

Full Usage Subcarrier (FUSC) uses all data subcarrier to create some sub-channels. They are allocated for a sector or a cell to obtain diversity. Limited overhead (pilot and guard subcarrier) in FUSC makes the number of transmitted data subcarrier is greater than PUSC. In other hand, the transmitted data will be useless, if not all subcarriers are used. A slot of FUSC consists of a symbol of FUSC.

There are two kinds of pilot in FUSC; constant set pilot and variable set pilot. The difference of these is in the index of pilot subcarrier. Constant set pilot has permanent index while variable set pilot has random index. Variable set pilot is possible to predict channel response precisely, especially in great delay spread channel (small coherent bandwidth)[7].

In FUSC process, first, pilot subcarriers (constant set and variable set) are allocated using equation below:

$$Pilot\ Locations = Variabel\ sets\ \# + (6 \times (FUSC\ symbol\ number\ mod\ 2))\ [6] \qquad (1)$$

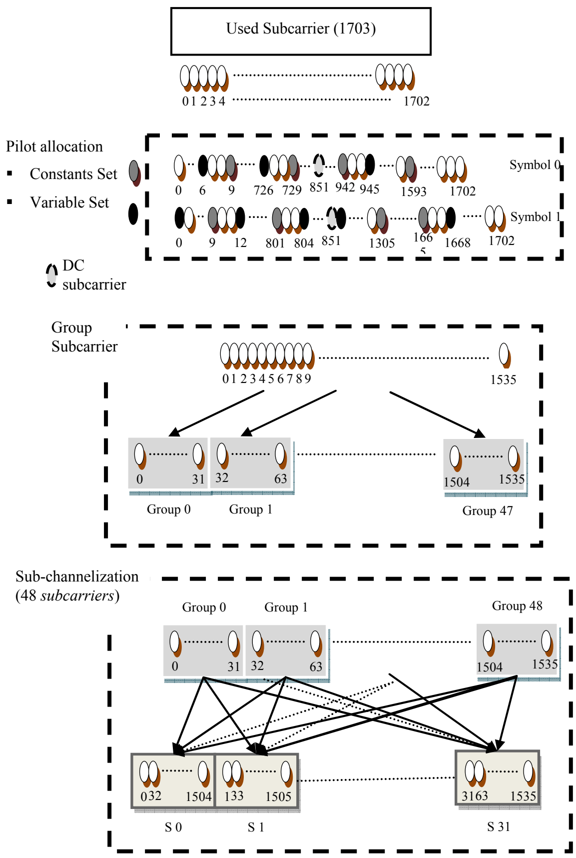


Figure 1. FUSC Permutation Process

Then, adjacent subcarriers are formed to be some groups and are allocated into a sub-channel. The number of sub-channel is as same as the number of subcarriers in a group. After that, remaining data subcarriers that are allocated are based on the equation below:

$$S(k,s) = N_{cn}.\eta_k + \{P_s(\eta_k\ mod\ N_{cn}) + DL\_PBase\}\ mod\ N_{cn} \qquad (2)$$

Where,

S(k,s) the index of subcarrier k among all available data subcarriers. With 2048-FFT, index of subcarrier ranges from 0-1535.

s the index number of a sub-channel in a group, from the set [0...Nsubchannels-1] i.e., with 2048-FFT, Nsubchannels ranges from 0-32.

k the subcarrier-in-sub-channel index from the set [0...Nsubcarrier - 1]. It ranges from 0 - 47.

N_{cn} The number of sub-channel.

η_k (k + 13s) mod $N_{subcarrier}$.

DL PBase	*Range* 0 – 31.
$P_s[.]$	is the series obtained by rotating basic permutation sequence cyclically to the left s times (Table-311 [3]). Ps[.]= {3, 18, 2, 8, 16, 10, 11, 15, 26, 22, 6, 9, 27, 20, 25, 1, 29, 7, 21, 5, 28, 31, 23, 17, 4, 24, 0, 13, 12, 19, 14, 30}.

FUSC permutation scheme has different parameter for each FFT size. It is shown from table below:

Table 1. FUSC Permutation Scheme Permutasi

FFT	128	256	512	1024	2048
Data subcarriers used	96	192	384	768	1536
• *Subcarriers per subchannels*	48	N/A	48	48	48
• *Number of subchannels*	2	N/A	8	16	32
Pilot Subcarriers in constant set	1	8	6	11	24
Pilot Subcarriers in variable set	9	N/A	36	71	142
Left-guard subcarriers	11	28	43	87	173
Right-guard subcarriers	10	27	42	86	172
DC subcarrier	1	1	1	1	1

2.2 Downlink PUSC

In Partial Usage Subcarrier (PUSC), some of data subcarriers are used to form sub-channel to obtain diversity. There is only one group that uses in one cell or sector depend on traffic condition in order to reduce interference. A slot of PUSC consists of two OFDMA symbol (odd and even symbol). The process of PUSC is divided into some steps, they are:

1. Outer Permutation
   It consists of physical cluster formation and logical cluster formation.
2. Inner Permutation
   A process of allocating subcarriers has group formation and sub-channelization.

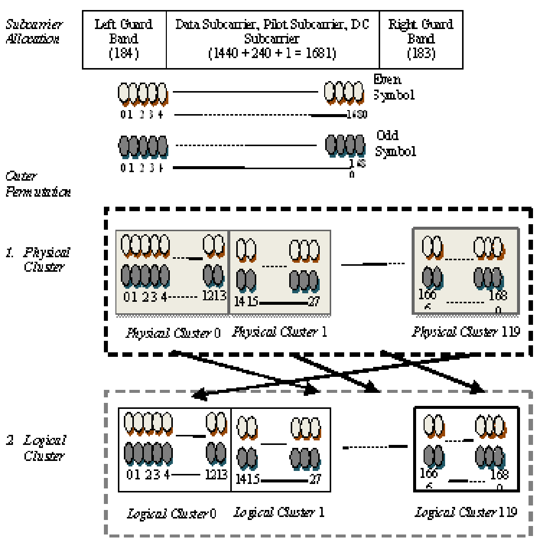


Figure 2. Outer Permutation

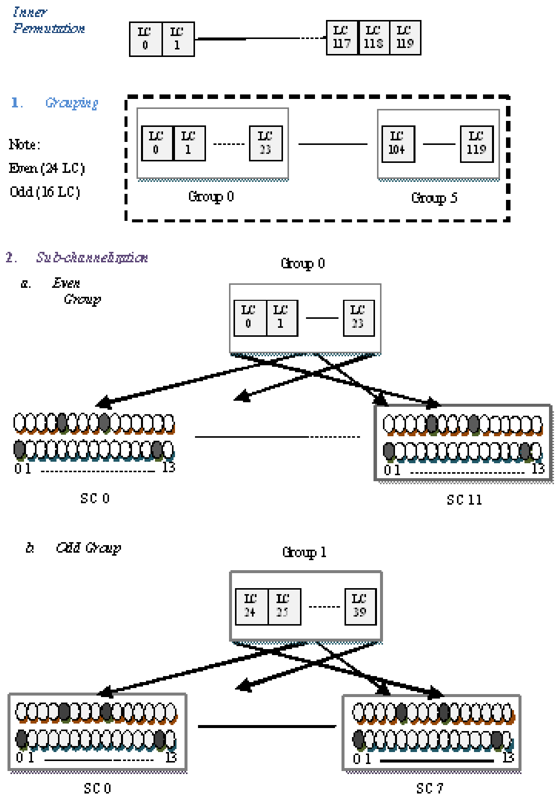


Figure 3. Inner Permutation

The process starts with physical cluster that is a group of fourteen adjacent subcarriers. Then, logical cluster is created by renumbering physical cluster. Logical cluster is done by equation 3.

$$Logical\ cluster = Renumbering\ Sequence\ (PC) + 13.DL\_PBase[6] \quad (3)$$

After that, logical clusters are formed by six groups that have even and odd group (group 0 to group 5). The even group consists twenty-four clusters while the odd group consist sixteen clusters. The even group will create twelve sub-channels and the odd group will create eight sub-channels. Pilot subcarriers position spread out for odd and even OFDM symbol. The equation for inner permutation shows below:

$$S(k,s) = N_{cn}.\eta_k + \{P_s(\eta_k\ mod\ N_{cn}) + DL\_PBase\}\ mod\ N_{cn} \quad (4)$$

Where,

S (k,s)	The index of subcarrier k within a group (i.e., with 2048-FFT, in group `0', it ranges from (0-287).
s	The index number of a subchannel in a group, from the set [0 - Nsubchannels-1] (i.e., with 2048-FFT, Nsubchannels for even and odd groups are 12 and 8 respectively).

k — The subcarrier in sub-channel index from the set. [0 - Nsubcarrier-1]

N_{cn} — The number of sub-channel for odd group is eight and for even group is twelve.

η_k — (k + 13s) mod $N_{subcarrier}$.

DL PermBase — *Range* 0 – 31.

$P_s[.]$ — The series obtained by rotating basic permutation sequence cyclically to the left s times Ps[:] = {6; 9; 4; 8; 10; 11; 5; 2; 7; 3; 1; 0} and {7; 4; 0; 2; 1; 5; 3; 6} for even and odd groups respectively.

The allocated subcarrier should be appropriate with the PUSC parameter below:

Table 2. PUSC Permutation Scheme Parameter

FFT	128	512	1024	2048
Data Subcarriers used	72	360	720	1440
• *Data Subcarriers per subchannels per symbol*	12	12	12	12
• *Number of subchannels*	3	15	30	60
Pilot Subcarriers	12	60	120	240
• *Pilot Subcarriers per subchannels per symbol*	2	2	2	2
Left-guard subcarriers	22	46	32	184
Right-guard subcarriers	21	45	91	183

3. Model System

3.1 Transceiver System

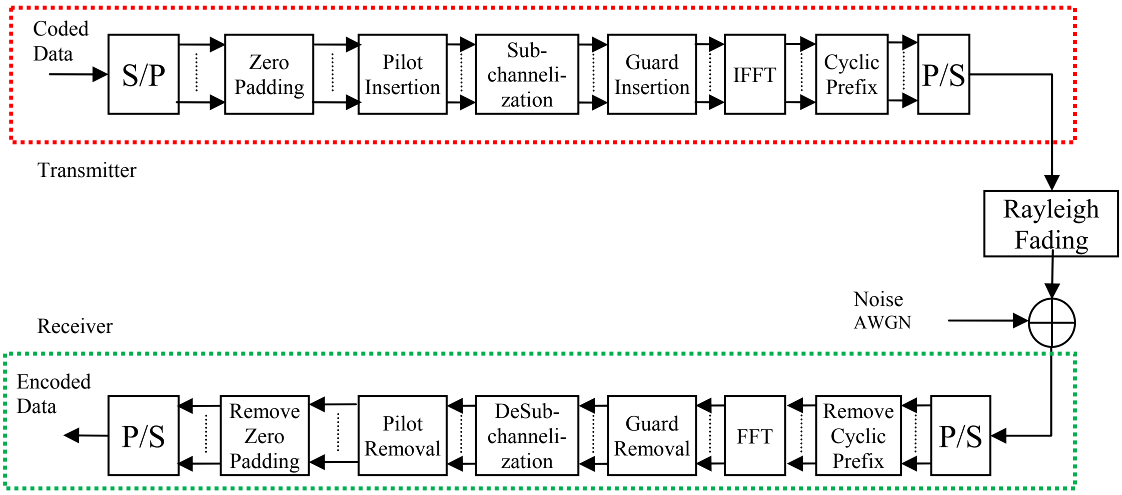


Figure 5. Transceiver OFDMA Diagram

Figure 6. Channel Coding Process

The simulation uses same method in number of bit, bit rate code, modulation type and channel condition. It was done to make simulation easier. The used numbers of bit are 1024. The parameters below used in the simulation according to IEEE 802.16e standard.

- Frequency : 3,5 GHz
- Channel Bandwidth : 20 MHz
- FFT Size : 2048
- Sampling Factor : 8/7
- Sampling Frequency : 22,8 MHz
- Spacing Subcarrier : 11,16 KHz
- Used Symbol Period : 89,6 µs
- Cyclic Prefix Ratio : 1/8
- Cyclic Prefix Period : 11,2 µs
- Periode simbol OFDMA : 100,8 µs

3.2 Transmission Channel

Channel model used in simulation are Rayleigh distributed multipath fading and AWGN (Additive White Gaussian Noise) channel. The chosen Jakes Model uses six delay taps and relative path power.

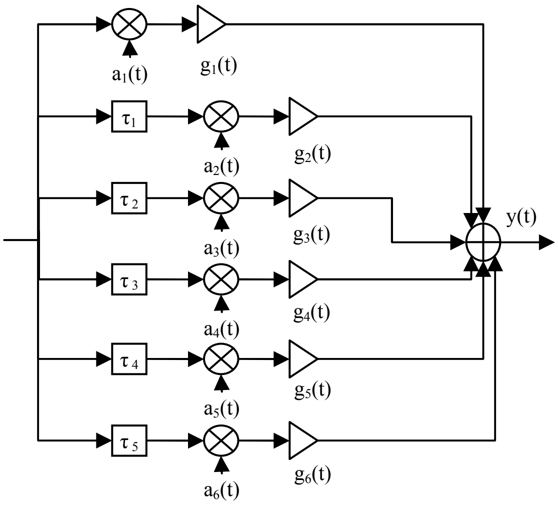


Figure 7. Multipath Fading Channel in Six-Path Delay

Where,

s(t)	= transmitted signal
τ_i	= multipath delay
$a_i(t)$	= Rayleigh variable coefisien
$g_i(t)$	= given gain
y(t)	= summary of received signal
I	= 1,2,3,...,6

Table 3. Multipath Delay Channel

Taps	*#1*	*#2*	*#3*	*#4*	*#5*	*#6*
Delay (ns)	*0*	*310*	*710*	*1090*	*1730*	*2150*
Relative path power (dB)	*0*	*-1*	*-9*	*-10*	*-15*	*-20*

4. Simulation Results

4.1 Single User

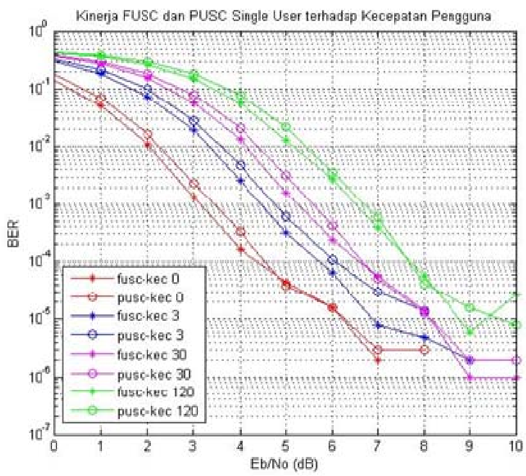


Figure 8. Single User Comparison between FUSC and PUSC

Table 4. Single User FUSC and PUSC Performance

Velocity	FUSC		PUSC	
	BER	Eb/No (dB)	BER	Eb/No (dB)
0 km/hour	10^{-5}	6.24	10^{-5}	6.26
3 km/hour	10^{-4}	5.75	10^{-4}	6.07
30 km/hour	10^{-5}	8.12	10^{-5}	8.13
120 km/hour	10^{-5}	8.75	10^{-5}	9.65

Figure 8 shows that the faster user moves, the greater Eb/No is needed. Besides, FUSC has better BER than PUSCs; it shows when these Eb/No are less than 4 dB. Table 4 shows that the faster user moves, the greater BER value is. It happens because mobile user will make mean time alteration in radio mobile channel that causes Doppler spread. This condition appears because there is a phenomenon of user frequency shift. The faster user move, the higher Doppler shift is. Furthermore, there will be an Inter Carrier Interference (ICI) effect. ICI happens in non- linear channel condition that will influence symbol detection process and received power in antenna.

4.2 Four Users System

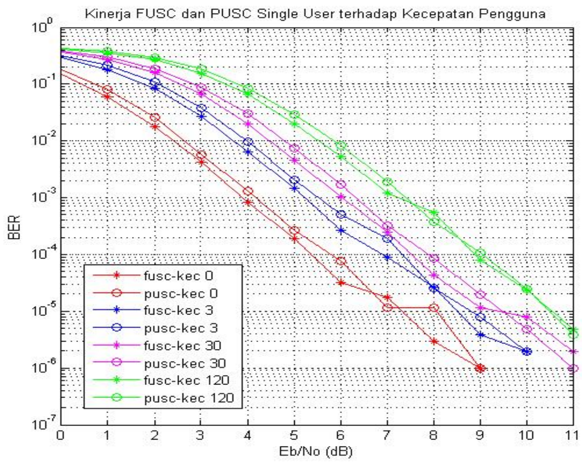


Figure 9. Four Users Comparison between FUSC and PUSC

Table 5. Four Users FUSC and PUSC Performance

Velocity	FUSC		PUSC	
	BER	Eb/No (dB)	BER	Eb/No (dB)
0 km/hour	10^{-6}	9	10^{-6}	9
3 km/hour	10^{-5}	8.5	10^{-5}	8.75
30 km/hour	10^{-6}	12	10^{-6}	11
120 km/hour	10^{-5}	10.51	10^{-5}	10.49

Figure 9 shows the faster users move, the greater Eb/No is. Mobile users will cause the alteration of mobile radio channel mean time so that Doppler spread appears. The faster users move, the higher Doppler shift happens. That is why signal will have distortion. Besides, the faster user move, the greater BER value is.

Table 5 shows an insignificantly differenced performance of FUSC and PUSC with same BER and Eb/No. They have almost same use of power signal. However, PUSC has better performance than FUSC when user moves for 30 km/hour. It happens because PUSC is more stable than FUSC confront to random channel. PUSC does twice as many as FUSC to maintain its system stability.

4.3 Sixteen Users System

Table 6. Sixteen Users FUSC and PUSC Performance

Kecepatan	FUSC		PUSC	
	BER	Eb/No (dB)	BER	Eb/No (dB)
0 km/jam	10^{-4}	7,26	10^{-4}	7,58
3 km/jam	10^{-6}	8,95	10^{-6}	9,95
30 km/jam	10^{-5}	9	10^{-5}	9,5
120 km/jam	10^{-4}	8,39	10^{-4}	8,76

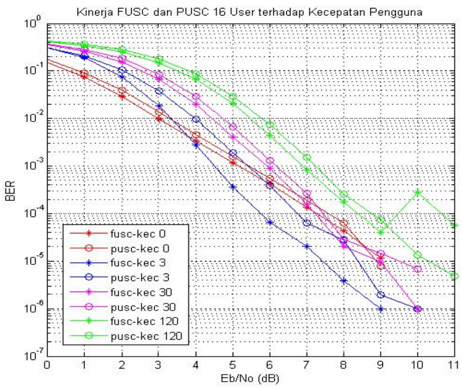


Figure 10. Sixteen Users Comparison between FUSC and PUSC

Table 6 and Figure 10 show fix users need greater bit energy to overcome random channel. It also because noise for mobile user is greater than fix users so that Eb/No for mobile user is smaller. Eb/No will be greater when users move faster. Moreover, figure 10 shows that system start to be unstable. It happens because subcarrier allocation is less along with the increasing of number of user. Hence, user subcarrier interference is easier to appear.

4.4 Thirty-two Users System

Table 7. Thirty-two Users FUSC and PUSC Performance

Kecepatan	FUSC		PUSC	
	BER	Eb/No (dB)	BER	Eb/No (dB)
0 km/jam	10^{-5}	7,35	10^{-5}	7,7
3 km/jam	10^{-5}	8,5	10^{-5}	8,76
30 km/jam	10^{-4}	7,46	10^{-4}	7,8
120 km/jam	10^{-5}	10,4	10^{-5}	10,42

Figure 11 and Table 7 shows the needed Eb/No will be greater if users move faster. The system seems unstable because allocated subcarrier is smaller. It will lead to user-subcarrier-interference appearance. Furthermore, more error will appear and it will be difficult to be detected and to be corrected in receiver.

5. Conclusion

1. Single user system, FUSC has better performance than PUSC. It shows in 4 dB Eb/No for fix user, FUSC BER is 1.67 10^{-4} while PUSCs is 3.33 10^{-4}.
2. Four users system, FUSC performance is better than PUSCs. For instance, when Eb/No is less than 7 dB, 3 km/hour user BER for FUSC is 8.79 10^{-5} while PUSCs is 1.88 10^{-4}.
3. Sixteen users system, FUSC Eb/No is better than FUSCs. It shows that Eb/No for FUSC fixed user and mobile user are 7.26 dB, 8.95 dB, 9 dB dan 8.39 dB while PUSCs are 7.58 dB, 9.95 dB, 9.5 dB dan 8.76 dB.
4. Thirty-two users system, FUSC is better than PUSC because FUSC need smaller signal power than PUSCs. For instance, BER value is 10^{-6} and FUSC's signal power is 8.95 dB while PUSCs is 9.95 dB.

References

[1] Gunawan Wibisono and Gunadi Dwi Hantoro, 2008, *Mobile Broadband: Tren Teknologi Wireless Saat Ini dan Masa Datang*, Informatika, Bandung.

[2] Gunawan Wibisono and Gunadi Dwi Hantoro, 2006, *WiMAX: Teknologi Broadband Wireless Access (BWA) Kini dan Masa Datang*, Informatika, Bandung.

[3] IEEE, 2004, *Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands and Corrigendum 1*, New York.

[4] IEEE, 2004, *IEEE Standard for Local and Metropolitan Area Networks - Part 16: Air Interface for Fixed Broadband Wireless Access Systems*, New York.

[5] Ilyasa, Tito, *OFDM pada Komunikasi Digital Pita Lebar*, University of Indonesia, Depok.

[6] Maqbool Masood, Marceau Coupechoux and Philippe Godlewski, 2008, *Subcarrier Permutation Types in IEEE 802.16e*, Departement Informatique et Reseaux TELECOM ParisTech, Paris.

[7] Mobilecomm.Lab, 2007, *WiMAX Fundamental: Technology, Bussiness, & Implementation*, Telkom Institute of Tecnology, Bandung.

[8] Rappaport, Theodore S, 2002, *Wireless Communication Principles and Practice*, Prentice-Hall, Inc., America.
[9] Schulze, Henrik dan Christian Lűders, 2005, *Theory and Applications of OFDM and CDMA*, John Wiley & Sons Ltd, England.
[10] Sirait, Emyana Ruth Eritha, 2007, *Analisa dan Simulasi Sistem OFDMA pada Standard Mobile WiMAX*, Telkom Institute of Tecnology, Bandung.
[11] Surahmat, Andi Ervan, 2008, *Perancangan Algoritma Subakanalisasi pada Mobile WiMAX IEEE 802.16e*, Telkom Institute of Tecnology, Bandung.

ANALYSIS PERFORMANCE OF BANDWIDTH REQUEST-GRANT MECHANISMS IN WIMAX NETWORKS

Dini Annisa[1], Sofia Naning Hertiana[2], Sony Sumaryo[3]

[1,2,3]Electrical and communication faculty, Institut Teknologi Telkom
[1]dini.annisa@gmail.com, [2]snh@ittelkom.ac.id, [3]sny@ittelkom.ac.id

Abstract

Broadband technology and wireless is two telecommunications technologies that growing swiftly. One of them is WiMAX technology, guaranteed high-speed with broader access coverage than technology wireless before. This thing causes more and more user to obtain amenity to access various applications every time and everywhere. Requirement base of service user is varying and there are many users in one Base Station (BS), hence allocation process of bandwidth in communications system wireless will become very complex. To overcome this IEEE 80216 as standard for WiMAX has a Bandwidth Request-grant mechanism for uplink bandwidth allocation between base station (BS) and subscriber station (SS). In this research, evaluation of bandwidth request-grant mechanism is done by measuring throughput, delay and packet loss system when overcoming VoIP traffic based on result of simulation using network software OPNET 14.0. Simulation result done at this research embraces : first scenario, maximum throughput is obtained when bandwidth request 64 kbps for amounts user 10, 20, 30, 40, 50 equal to 52.93 kbps,32.67 kbps,21.87 kbps,16.49 kbps, and 13.21 kbps. Packet loss is, when bandwidth request 64 Kbps and 128 Kbps with number of user 10, 20, and 30 which appropriate to ITU G.107, less than 20%. Delay, when bandwidth request 128 kbps, 192 kbps, 256 kbps, and 320 kbps with number of user increases, there are still be tolerated by residing in at range 0-150 ms (ITU G.114). Scenario 2, maximum throughput is obtained when bandwidth request 64 Kbps for speed of user 0 km/h,5 km/h,60 km/h and 80 km/h equal to 52.93 kbps, 51.21 kbps, 48.04 kbps and 31.97 kbps. Packet loss is, when bandwidth request 64 kbps and 128 kbps with speed of user 0 km/h, 5 km/h, and 60 km/h still be tolerated, less than 20% (ITU G.107). At the time of bandwidth request 128 Kbps, 192 Kbps, 256 Kbps, and 320 Kbps with speed of user increases, delay still be tolerated by residing in at range 0-150 ms (ITU G.114).

Key words : WiMAX, Bandwidth Request-grant, VoIP

1. Introduction

Broadband technology and wireless telecommunications technologies are two of the fastest growth today. Both technologies offer convenience to the user to access various services and information whenever and wherever they need. Differences cause broadband services needs different bandwidth. A number of bandwidth provided to a service must meet the quality of those services that involve several parameters such as throughput, error rate, delay, and jitter.

Based on the needs of various user services and the large number of users in a BS, the bandwidth allocation process in the wireless communication system will become very complex and complicated, especially when allocating uplink bandwidth.

WiMAX (Worldwide Interoperability for Microwave Access) technology is a Broadband Wireless Access (BWA) that has high speed access to the wide range. IEEE 802.16 standard for BWA systems. WiMAX in the design to meet the needs of Quality of Service (QoS) on the uplink and downlink connections. In the IEEE 802.16 standards, there is a mechanisms Request-grant bandwidth that is used to accommodate various QoS requirements in heterogeneous traffic. Request-grant protocol is used for uplink bandwidth allocation between the BS and SS. At the time of the SS to send data, the first one sends bandwidth request to BS. BS will accept the request and the bandwidth that will be determined whether there is enough bandwidth to satisfy the bandwidth request. This study aims to analyze the performance / performance mechanism for bandwidth request-grant specifically for Voice services ever Internet Protocol (VoIP) in a WiMAX network. Known parameters so that it can be throughput, delay and packet loss.

2. WiMAX Network

WiMAX is an international standard of Broadband Wireless Access (BWA), which refers to the IEEE 802.16 standard. The difference between WiMAX and wireless technology is high-speed access and wide coverage.

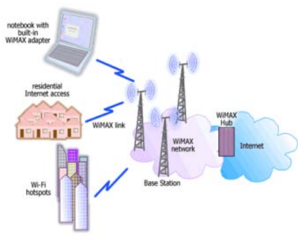


Figure 1. WiMAX Network Topology

WIMAX technology has reached the distance range 50 km and handle data rates up to 75 Mbps. WiMAX is also a technology with open standards. In a sense, WiMAX communication between multiple devices of different vendors can still be done.

2.1 Quality of Service (QoS) in WiMAX

With the birth of new technologies in wireless networks like WiMAX, it is typically accompanied by a greater ability when compared with previous generation technologies. Many advantages offered by WiMAX allow more and more users take advantage of this technology. That requires a guarantee of service quality so that the user requests can be met varies. WiMAX features Quality of Service (QoS) as a guarantee service performance from BS to SS. WiMAX offers a variety of QoS that can guarantee the quality of services provided to users. There are 4 types of service classes provided by WiMAX, namely:

a. Unsolicited Grant Service (UGS)
   UGS is used for services that require guaranteed data transfer with a top priority. Service with UGS criteria has characteristic that require guaranteed real-time and effective for the services that are sensitive to throughput, maximum latency and jitter. Examples: for VoIP applications, T1/E1, ATM.
b. Real-time Polling Service (rtPS)
   Effective for services that are sensitive to throughput and latency, but with a looser tolerance when compared with UGS. Example: MPEG video applications and video conferencing.
c. Non-Real Time Polling Service (nrtPS)
   Effective for the application of non real-time that requires guarantee of performance. Examples for non-real time services such as streaming video and audio.
d. Best Effort (BE)
   In general, for any traffic that does not require collateral speed data (best effort). There is no guarantee (requirements) on the rate or delay it. Example for internet applications (web browsing), email, FTP.

2.2 Bandwidth Request-Grant Method

2.2.1 Bandwidth Request

SS bandwidth request mechanism is conducted for uplink bandwidth allocation in BS. Bandwidth demand is always done at every connection established basis. Bandwidth request can be made by sending a request packet bandwidth in the uplink sub-frames or by laying (piggybacking) on packet data. Figure 2 shows the Bandwidth Request Contention Slots in Uplink sub-frame.

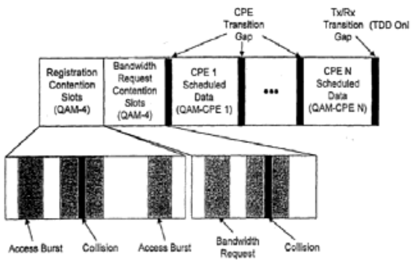


Figure 2. Bandwidth Request Contention Slot in Uplink Sub-Frame

Uplink subframe is used to transmit information, including bandwidth requests to the BS. In the uplink subframe, there are three MAC control messages, namely:

a. Registration contention slot that functions to register SS.
b. Bandwidth request contention slot
c. CPE schedule slot data

2.2.2 Grant

BS will check the availability of bandwidth in accordance with the SS bandwidth request. If the network does not have enough resources, then the bandwidth request will be rejected. Whereas if the network has sufficient resources, the BS will allow the SS to use these resources. IEEE 802.16 set two modes of bandwidth request fulfillment, namely:

a. Grant Per Subscriber Station (GPSS)
   In the GPSS mode, BS-bandwidth needed at of each SS. This mode is suitable for the lot number of connections per terminal, and allows to meet the QoS needs of such complex real-time applications that require quick response.
b. Grant Per Connection (GPC)
   BS Scheduler handles each connection separately and bandwidth will be met for each connection. GPC mode according to the number of users per terminal in small amounts

3. Designing Simulation Model

3.1 Network Configuration

Network configuration used in this simulation is as follows:

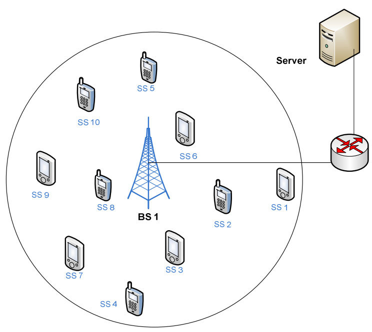


Figure 3. Network Topology

3.2 Parameter Models

Parameter Models in this simulation are follows:

Link between BS and the router = E1 (2048Mbps) [3]
Links between routers and servers = E1(2048Mbps) [3]
The flow of data from the SS to the server.
The size of voice packet is sent 160 byte (ITU-T G.711) [1]
Pattern arrival time based on exponential distribution.
Length of simulation runs for 200 sec.
Polling methods used for bandwidth request mechanism

3.3 Scenario Simulation

Scenario 1: Analysis of changes in the number of users and bandwidth request. In this scenario, how changing the number of users and bandwidth request of the parameters of throughput, packet loss, and delay. Source consists of sending traffic to be analyzed and spam traffic.

Scenario 2: Analysis of changes in speed and the amount of bandwidth a user request. In this scenario, how changing the amount of bandwidth requests to the parameters throughput, packet loss, and delay of the move user and quiet. Source consists of sending traffic to be analyzed and spam traffic.

3.4 Making Simulation

In this simulation it is used OPNET Modeler 14.0 Educational Version which works on Windows XP operating system. This software has advantages for the network design based on existing devices on the market, protocols, services and technologies. Simulation results can be made in a few scenarios that can be relied upon in planning a packet-based network.

4. Analysis of Simulation Results

4.1 QoS Parameters

The parameters analyzed include:

a. Throughput is defined as the average speed of effective data received by the receiver node at a certain observation time interval.
b. Packet loss percentage is the number of packets lost in the process of sending data from source to destination nodes to the total number of packets sent. Packet loss voice and multimedia applications can be tolerated up to 20% (ITU standard G.107)
c. Delay is the amount of time a packet to travel from source to destination (end-to-end delay). Delay for voice and multimedia applications can be tolerated up to 0 to 150 ms (ITU standard G.114)

4.2 Analysis of Changes in The Number of Users and Bandwidth Request

In this scenario to test how changing the number of users and bandwidth request of the parameters of throughput, packet loss and delay. Source consists of sending traffic to be analyzed and spam traffic. The following user terms used:

a. Changing the number of users consisted of 10, 20, 30, 40, and 50 users that use VoIP traffic and background -traffic FTP.
b. Determination of the amount of bandwidth request is:

$$n \text{ x } DS0 = n \text{ x } 64 \text{ Kbps}; \qquad n = 1, 2, 3,$$

c. Timeout bandwidth request packet (T3) are 50 ms - 200 ms (Table 342 Standard of 802.16e-2005). In this simulation package used bandwidth timeout 50 ms.

4.3 The Influence of Changes in The Number of Users and Bandwidth Request of The Throughput

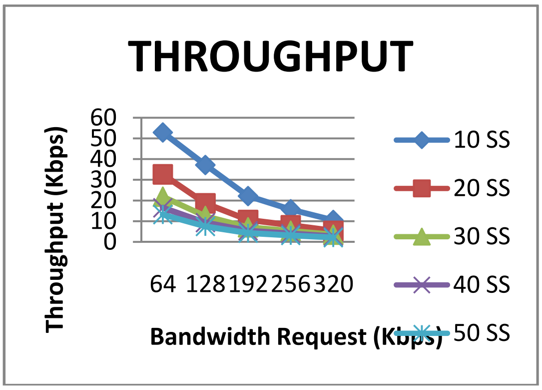


Figure 4. Throughput Scenario 1

It can be seen that the greater number of users then the average throughput per user getting smaller. This is caused by the number of users served in a more BS, which resulted in a solid traffic. Greatest throughput when the number of users 10, 20, 30, 40, and 50 produced at the request bandwidth equal to 64 Kbps, in the amount of 52.93 Kbps, 32.67 Kbps, 21.87 Kbps 16:49, and 13:21 Kbps.

4.4 The Influence of Changes in The Number of Users and Bandwidth Request of Packet Loss

Adding the number of users and bandwidth request in this scenario causes packet loss percentage is increasing. This is because the user is assumed to grow and continue to access the service.

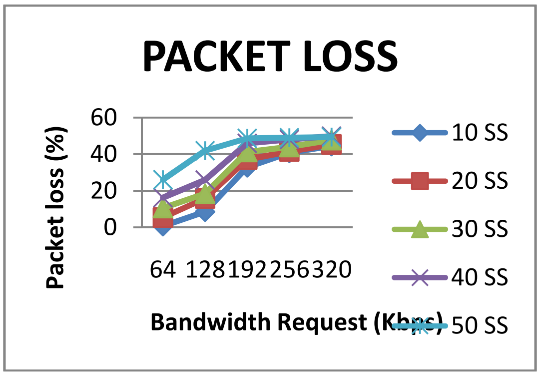


Figure 5. Packet Loss Scenario 1

The lowest packet loss as the number of users consisted of 10, 20, 30, 40, and 50 produced at the request bandwidth equal to 64 Kbps, in the amount of 0.67%, 5:36%, 10.5%, 16.2%, and 26.1%. Referring to the ITU G.107 standard, the best conditions obtained at 64 Kbps bandwidth requests to the number of users 10, 20, 30, and 40. While for 128 Kbps bandwidth requests when the number of users 10, 20, and 30.

4.5 The Influence of Changes in The Number of Users and Bandwidth Request to Delay

Figure 6 shows the influence of bandwidth change request and change the number of users to the end-to-end delay. For any number of users 10, 20, 30, 40, and 50 largest delay occurs at 64 Kbps bandwidth request is 337.75 ms, 556.22 ms, 785.4 ms, 892.4 ms and 984.3 ms. This condition is not in accordance with the ITU G.114 standard, where delay for voice and multimedia applications can be tolerated up to 0 to 150 ms. This happens because the amount of bandwidth request is smaller than the total bandwidth required to pass VoIP traffic is 85.6 Kbps.

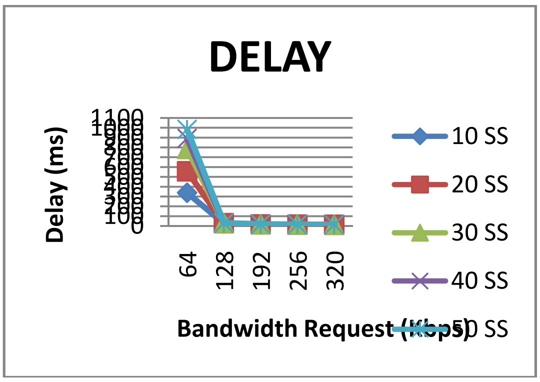


Figure 6. Delay Scenario 1

4.6 Analysis of The Influence of Velocity Changes and The Amount of Bandwidth of User Request

In this scenario to test how changing the amount of bandwidth requests to the parameters throughput, packet loss, and delay of move user and quiet. Source consists of sending traffic to be analyzed and spam traffic. The following terms used in the user scenario:

a. Consists of 10 users with VoIP traffic specification and backgorund traffic-FTP.
b. Simulation done for 3 (three) conditions of the user, not the move, walking distance (5 km / h), and drive (60 km / h).
c. Determination of the amount of bandwidth request is:

n x DS0 = n x 64 Kbps

d. Timeout bandwidth request packet (T3) is 50 ms - 200 ms (Table 342 Standard of 802.16e-2005). For this simulation package used bandwidth timeout 50 ms.

4.7 The Influence of Velocity Changes and The Amount of Bandwidth A User Requests to The Throughput

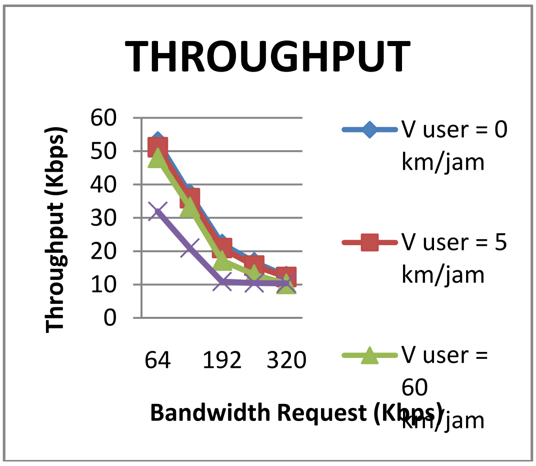


Figure 6. Throughput Scenario 2

Figure 6 above shows the throughput scenario 2 when the amount of bandwidth change request with the user speed which different intervals of observation in 200 seconds.

The higher speed will generate the lower throughput of user. This is because the signal quality changes the user to move or migrate. These conditions apply to any amount of bandwidth request. At the same amount of bandwidth requests with 128 Kbps, 192 Kbps, 256 Kbps and 320 Kbps produces lower throughput.

4.8 The Influence of Velocity Changes and The Amount of Bandwidth A User Requests to Packet Loss

Adding speed and bandwidth user requests in this scenario causes the percentage of packet loss increases. This is because the speed of users increases and the user is assumed to continue to access the service, causing the signal quality decreases.

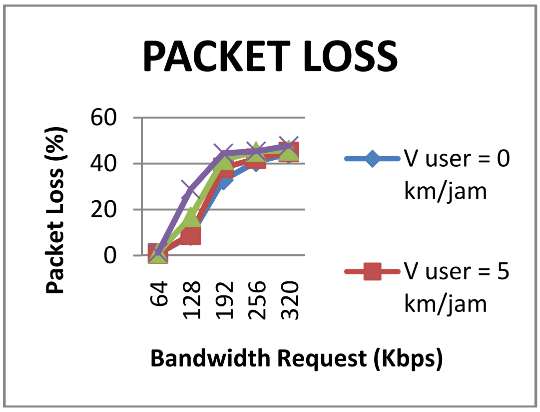


Figure 7. Packet Loss Scenario 2

Packet loss at the lowest speed of the user 0 km / h, 15 km / h, 60 km / h, and 80 km / hour produced at the request bandwidth equal to 64 Kbps, in the amount of 0.67%, 0.86%, 0.96%, and 1:22 %. Referring to the ITU G.107 standard, the best conditions obtained at 64 Kbps bandwidth requests the user speed 0 km / h, 5 km / h, 60 km / h, and 80 km / hour. While for 128 Kbps bandwidth request when the user speed 0 km / h, 15 km / h, and 60 km / h, where packet loss is less than 20%.

4.9 The Influence of Changes in Speed and Bandwidth of A User Request to Delay

For each user speed 0 km / h, 5 km / h, 60 km / h, and 80 km / hour, the greatest delay occurs at 64 Kbps bandwidth request is 337.75 ms, 408.69 ms, 453.81 ms and 491.25 ms. This condition is not in accordance with the ITU G.114 standard, where delay for voice and multimedia applications can be tolerated up to 0 to 150 ms

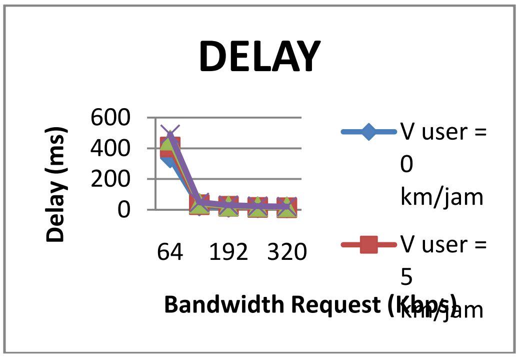


Figure 8 Delay Scenario 2

Lowest delay occurs when 320 Kbps bandwidth request of 11:35 ms, 13,082 ms, 18,407 ms and 19,805 ms. Conditions in accordance with the ITU G.114 standard is when the bandwidth request 128 Kbps, 192 Kbps, 256 Kbps and 320 Kbps. Figure 4.6 shows the influence of bandwidth change request and change the number of users of the delay.

5. Conclusion

Based on the simulation and analysis has been done, can be concluded as follows:

a. The addition amount of bandwidth throughput of requests resulted in a decrease of 80.14% for the number of users 10, 83.68% for the number of users 20, 84.04% for the number of users 30, 83.74% for the number of users 40, and 84.17% for the number of users 50.
b. The addition amount of bandwidth and the number of user requests result in an increasing number of packet loss. Conditions in accordance with the ITU G.107 standard is a bandwidth request at 64 Kbps and 128 Kbps with the number of users 10, 20, and 30 where the packet loss is less than 20%.
c. The addition of lead to lower bandwidth request delay, bandwidth request at 128 Kbps, 192 Kbps, 256 Kbps and 320 Kbps with the number of users increases, the delay is still tolerable because it is in the range 0-150 ms (ITU G.114)
d. The addition amount of bandwidth requests lead to lower throughput up to 76.36% for silent user conditions (0 km / h), 76.02% for the user to walk (5 km / h), 78.8% for users drive (60 km / h), and 67.59% for the users drive (80 km / h).
e. The addition amount of bandwidth and speed of the user requests result in an increasing number of packet loss. At the request of bandwidth and 64 kbps to 128 kbps speed user 0 km / h, 5 km / h, and 60 km / h, packet loss is tolerable because it is less than 20% (ITU G.107).
f. The addition of lead to lower bandwidth request delay, bandwidth request at 128 Kbps, 192 Kbps, 256 Kbps, 320 Kbps and the user speed increases, the delay is still tolerable because it is in the range 0-150 ms (ITU G.114)

References

[1] _______ . ITU Standard. 2005.

[2] Ahson, Syed. 2008. WiMAX Standard and Security. CRC Press : New York.

[3] Dwi Hantoro, Gunadi. 2006. WiMAX Teknologi Wireless Access (BWA) Kini dan Masa Depan. INFORMATIKA : Bandung.

[4] IEEE Computer Society. 2006. Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands. IEEE : New York.

[5] Indriani, Wulan. 2008. ANALISA QoS WiMAX IEEE 802.16e. Institut Teknologi Telkom : Bandung.

[6] Kim, Hwangnam. 2008. Dynamic Bandwidth Request-Allocation Algorithm for Real-Time Services in IEEE 802.16 Broadband Wireless Access Networks. Telecommunication & Network Division, Samsung Electronics : Korea.

[7] Maheshwari, Supriya. 2005. An Efficient QoS Schedulling Architecture for IEEE 802.16 Wireless MANs. Kanwal Rekhi School of Information Technology : India.

[8] www.wikipedia.com/Forum/WiMAX. 2007/White Paper WiMAX.

[9] Opnetwork. 2005. WLAN Lab Opnet Tutorial. Washington D.C.

[10] Switching Technique Laboratory. 2009. Modul Praktikum EXT Jaringan Telekomunikasi. Institut Teknologi Telkom : Bandung.

[11] Zimmerman, Ofer. 2005. Method and Apparatus for Bandwidth Request-Grant Protocols in a Wireless Communication System. Procopio : San Diego.

DESIGN ON INFORMATION TECHNOLOGY STRATEGIC PLANNING FRAMEWORK BY ACTIVITY VALUES IDENTIFICATION

Yudha Purwanto[1], Arry Akhmad Arman[2]

[1]Institut Teknologi Telkom
[2]Bandung Institute of Technology
[1]ydp@ittelkom.ac.id, [2]arman@kupalima.com

Abstract

In our business area, strategy is a vital need to determine any pattern of corporate actions to achieve the vision. But in many cases, there're many strategies which fail to be implemented because the strategies didn't plan or design well. Strategic planning is an important step, start with good strategic planning will produce good strategies that will guide corporate actions lead to achieve corporate vision. In this case, need a planning framework which will give guidance through the strategic planning activities. From many existing strategic planning framework, each had their own value/excellence. This research has analyzed four strategic planning methodologies, combined them in to a synthesized strategic planning framework. The combination done trough analysis of each methodology values, and synthesize them in to a single framework which had more complete values. The proposed framework has tested by benchmark to the guideline from CCTA. The research output is a strategic planning framework which have more complete values, well structured and implementable.

Keyword: information technology, strategic planning, framework, values, CCTA

1. Introduction

In strategy management cycle, there're some phase that we usually done, starting with strategy formulation, strategy implementation and followed by evaluation and control. The result of earlier cycle of strategy management will always connected to the next cycle, by the evaluation and control of strategy implementation which will give feedback to the strategy formulation in the next cycle. Each strategy formulation will give a strategic planning, which will be implemented in the corporate or business area.

The problem is how to formulate a good strategy plan. In many failed implementation cases, they started with a poor strategic planning which didn't notice any influence/consideration which can cause problems in implementation, or even worst didn't notice any consideration which will happen in the future. That's why it's very important to make strategic planning framework which will guide the user to many important things to consider in strategic planning formulation. Not only to formulate the strategy but also identify the strategic steps which have clear and measurable objectives. In some known IT Strategy Planning (ITSP) methodologies, each has their own benefits which will drive the values of each methods.

This research based on hypothesis; by combination of well known Strategic Planning methodologies, can be made an ITSP framework which has complete activity values. With complete activity values, management board which has decision of IT strategy can consider any important influences/consideration related to the formulation of ITSP. Well formulated ITSP will make easier and more effective IT strategy implementation, that will lead to the implementation success.

This research formulate problems by identification of :

- What are the values of each methodology
- How to combine ITSP methodologies to provide benefits of complete values
- How to optimize the combined activities of each methodology to produce an effective framework

This research analyze four strategy planning methodologies with different perspective, such as :

- John Ward[5] provide ITSP framework by identify information that business need, and information system can be develop to produce the information, and then IT strategy was planned to support the information system.
- Bernard Boar[2] provide ITSP framework base on seeking the opportunity of products and services which will lead to corporate advantages in hyper competition.
- Markus Venzin[6] provide strategy planning framework, not only in information technology area, but also to guide corporate, business and functional strategy.
- Chang Kim in Blue Ocean Strategy[4] provide different approach of ITSP, by create new markets (blue ocean) and leave the competition (red ocean).

2. Basic Theory

Strategy is a chosen action pattern to create competitive advantage as way to achieve certain vision of an enterprise. Information technology is technology which use to data to provide good quality information. By the term of these definitions, information technology strategy planning is the planning to use technology to provide good quality information that will lead to the corporate's competitive advantage to achieve vision in information technology.

Central Computer and Telecommunication Agency (CCTA)[7] made a well known Guidelines on Strategic Planning for Information Systems. As a guideline, it own logical activities which will be the values of the guidelines, which are :

- Business value analysis
- Create parameters of values
- Risk calculation
- Making systematic implementation planning
- Analysis of external business opportunities
- Problems and analysis restricted/boundary
- Analysis and interpretation of internal organization's business plan
- Non technical internal support of IT
- Management planning and control
- Making alternative solutions
- Basic policy and implementation
- Technology utilization analysis
- Identification and detailed description output per process
- Determination of migration

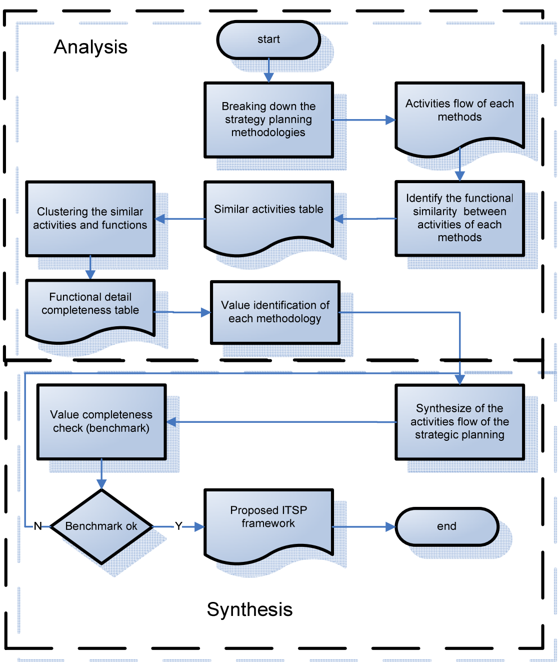


Figure 1. Research Methodology.

These values are divided in strategy planning phases, which are :

- Scope review
- Strategy formulation
- Strategy definition
- Implementation planning
- Strategy review

3. Analysis

Breakdown was conducted by identifying activities of strategy planning based on strategy planning guidance compilation according to CCTA, that is

1. Analyzing scope
2. Analyzing Strategy
3. Defining Strategy
4. Planning Implementation
5. Studying To Repeat Strategy

After the breakdown of the four methodologies, founded comparison value the following four methodologies :

3.1 Bernard Boar [2]

Boar give majority of strategy planning to

- Gain instruction and assumption from management
- Business assessment to identify requirement of business information
- Paying attention to align with business strategy
- Situational Analysis to look for various alternative strategy
- Determining objective and measured goals
- Anticipating implementation in contingency plan
- Pay attention of the effect of strategy implementation through change management plan
- Gain commitment of effect of strategy reinforcement in commitment plan
- Learning & Vigilance in implementation program & project

The goal of ITSP is to provide IT strategy that align to business strategy. The strategy has to pay attention to present position of business and IT to formulate IT strategy which can lead to the expected business position in hyper competition era, as it planning for many aspect in strategy implementation

3.2 Markus Venzin [6]

This methodologies shows:

- Identify strategy based on market-based and resource-based
- Focus on business strategy planning of differentiation, diversification, or niche market

- Planning of corporate strategy level, business and functional
- Performance of Business monitoring

The target is making a planning of company strategy, business and functional, that focus on resource and or market, by using monitoring of business performance as the trigger of strategy planning.

3.3 John Ward [5]

John Ward's strategy planning focus to

- Internal condition and eksternal business
- Internal condition and eksternal SI/TI
- Identify expected value-chain
- Identifying information and required information system needed for business
- Investment strategy
- Strategy for vendor & outsourcing
- Priority of IS/IT strategy
- Strategy details

The goal is to provide strategy planning which capable to identify business condition, management, IS/IT, and IT/IS strategy clearly and structured.

3.4 Chang Kim [4]

- Market assessment based on competition structure
- The use of strategy canvas in depicting position
- Reconstructing market definition through six path framework
- Search for alternative strategies through four action framework
- Identify of resistance of strategy implementation
- Integration execution of strategy
- Thinking pattern to sustain strategy excellence

The target is to earn a strategy plan which create value and market differ (blue ocean), through simple planning step which consider organizational factor at implementation plan

4. Synthesis

4.1 Single Point of View

Synthesis process of strategy planning was conducted base on the analysis result of identification on detail activity and function completeness of each method. Synthesized activity flow, was conducted by making a single point of view or make same approach to all identified activities. Single point of view was depicted in a few phase and sub-phase planning, constituted to reply four following question,

a. What's the scope
b. Where are we now ?
c. Where do we want to be ?
d. How do we get there?

4.2 Proposed Framework

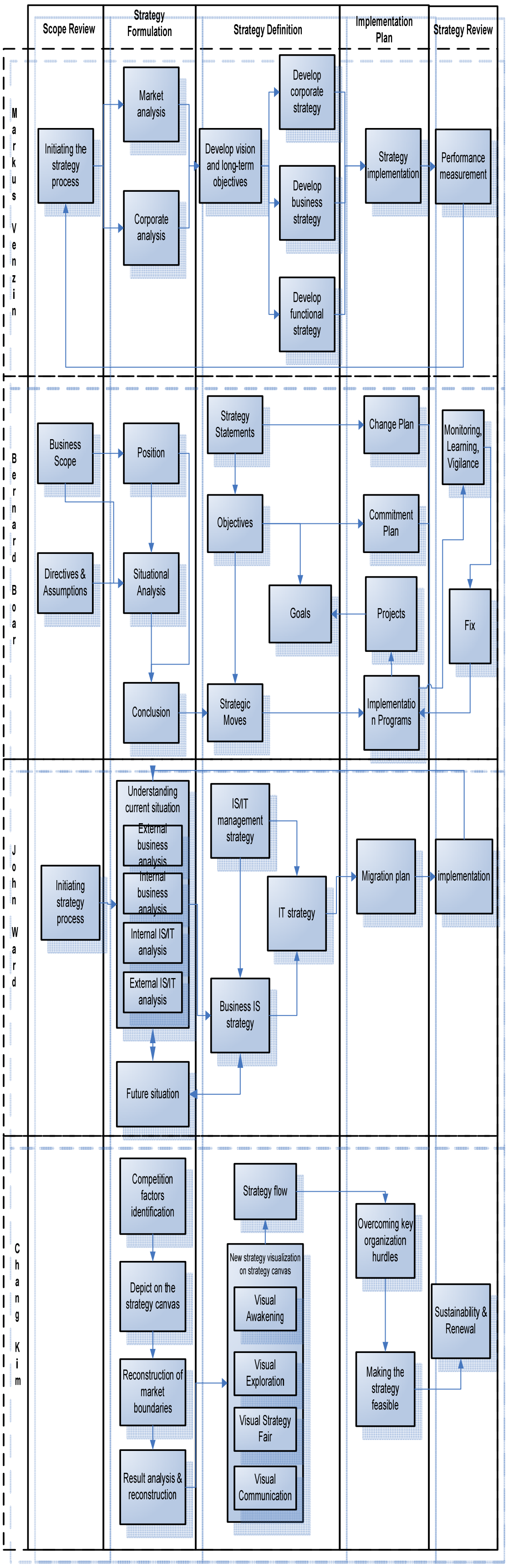


Figure 2. Activities of Each Methodology

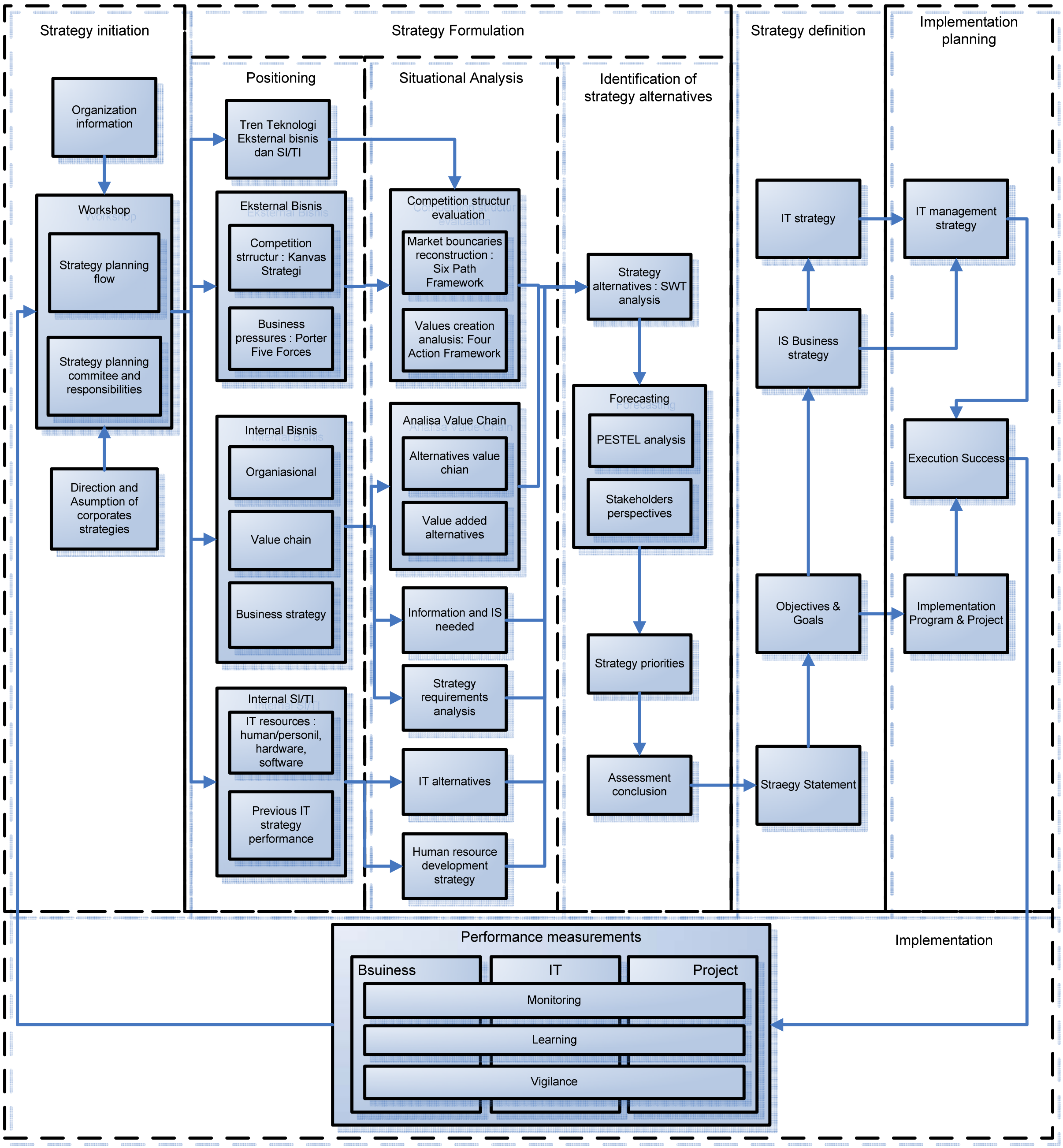


Figure 3 Proposed Synthesized Framework

4.3 Proposed Framework's Values

Values of synthesized framework seen from major activities excellence,

a. Owning instruction and assumption from management
b. Business assessment process to identify requirement of business information
c. Paying attention to align/support to business strategy
d. The use of simple strategy canvas in depicting business position
e. Internal and external situational analysis vary on market-based and resource-based, of business and IS/TI to search for various strategy alternatives
f. Look for alternative values as business inspire

g. Identifying information and required business information system
h. Identifying value-added and better reengineering process
i. Determine objective and measurable goals
j. IS strategy in details
k. IT strategy in details
l. Strategy of IS/IT management to support strategy execution/implementation
m. Overcoming resistances of strategy implementation through change management plan
n. Strategy reinforcement in commitment plan
o. Learning & vigilance in implementation's programs & projects
p. Performance monitoring of IS/IT

4.4 Activity Values Completeness Check

The proposed framework was tested by comparing it to "Guidelines on Strategic Planning for Information Systems" CCTA [7]. Benchmark result compared to CCTA, all logical activities as values in CCTA have accommodate/included completely in proposed framework.

4.5 The Result

4.5.1 Galaxy Model

Terminology:

a. All process of ITSP activities will be done to support company's vision, mission, values and objectives .
b. All activities in each phase have to follow instruction and assumption according to the organizational business information (vision, mission, values and objectives).
c. In certain activity phase, will consider any supporting activity factors which must be considered by ITSP committee.

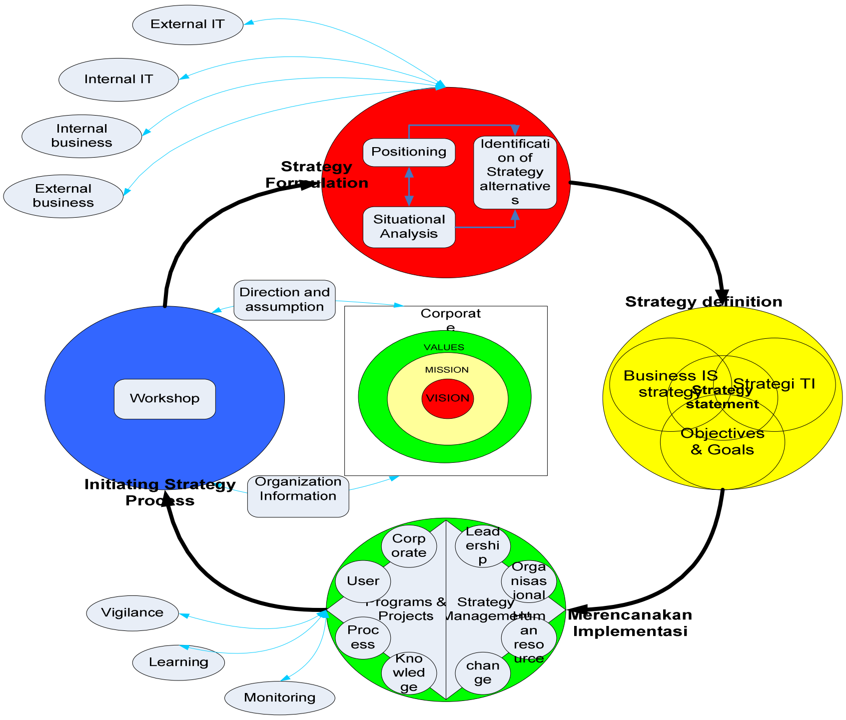


Figure 4. Galaxy Model

4.5.2 Synthesized ITSP Framework

There isn't any changes of the result of the synthesized framework, only be enhanced by activity at implementation planning phase, sub-phase planning program &implementation project, at Execution Success activity with additional duties of IS/IT project execution success, IS/IT operation execution success, business execution success (impact IT to business).

5. Conclusion and Suggestion

5.1 Conclusion

a. The proposed ITSP framework, which structured activities with complete value.
b. Structured framework shown as the framework was comprised/compiled in line sequence in each step of strategy planning.
c. Proposed framework fulfill all identified values from four combined methodology, then enhance some values which was adapted for requirement of ITSP committee.
d. This proposed ITSP framework could be implemented in many business scope by giving choice-able activity values according to the expected values in the IT strategy planning.
e. By following the proposed framework, will be able to take benefit to how company will face the competition, that is by
   - Identifying new alternative values
   - Choice of architecture and platform
   - Re-Engineering business process for better value-chain and more value-added
   - Making effective organization in supporting IT strategy
   - Program and project planning

5.2 Suggestion

Most important parts in ITSP are innovations and ability of scanning and interpret symptoms which emerge in each/every planning activity.

References

[1] Al-Hakim, Latif (2007) : *Information Quality Management*, IDEA Group Publishing

[2] Bernard Boar *The Art of Strategic Planning for Information Technology*, 2nd Ed, John Wiley & Sons, 2001

[3] Bruce E. Walters, Zaiyong Tang, IT-Enabled Strategic Management, Idea Groups Publishing, 2006

[4] Chan Kim W, Renee Mauborgne, *Blue Ocean Strategy*, Harvard Bussiness School, 2005

[5] John Ward, Joe Peppard, *Strategic Planning for Information Systems*, rd Ed, John Wiley & Sons, 2002

[6] Markus Venzin, *The Strategy Process : A Practical Handbook for Implementation in Business*

[7] Tozer, E.E. (1996), *Strategic IS/IT Planning, Professional Edition*, Butterworth-Heinemann.

MULTI LAYER PERCEPTRON CLASSIFIER ANALYSIS TO PREDICT MULTICLASS IN IT TELKOM STUDENT'S GRADE STUDY CASE

Agung Toto Wibowo[1], Husnul Khotimah[2], Warih Maharani[3]

[1,2,3] Institut Teknologi Telkom
[1]atw@ittelkom.ac.id, [2]uus_16a2@yahoo.com, [3]wrh@ittelkom.ac.id

Abstract

Artificial Neural Network (ANN) with backpropagation algorithm is widely used for classification or prediction. That is because the basic capabilities of the ANN in generalizing the given input information like a human being that can recognize specific patterns of an object based on the general pattern that has been learned and saved in the memory. A System that capable of predicting the value of specific subjects is required by students during the next semester registration. Commonly they do not know what subjects that should have been taken because the problem of interests and capabilities of the student. ANN, which can predict the value of certain subjects from a student can be used in addressing these problems. The architecture cconfigurations of an ANN determine considerably the quality of its network. Although the method that usually used is try and error, the determination of an ANN can be done based on the case that being tested. ANN Multi Layer Perceptron (MLP) with back propagation algorithm is often used for the case which the output is nonlinear. In this research, ANN architectures were tested to predict the twelve subjects in six semesters. Each subject has a grade index of A to E, so this case can be classified as a multiclass prediction. Performance will be measured using the MSE in the form of normalization and denormalization. The input is preprocessed by reducting the attribute conducted using Information Gain (IG) and then took five attributes for each subjects are needed with the highest score of IG. MSE that was obtained from the experiment are quite diverse in the twelfth such subjects. This is due to input quality aspects different. Best value for normalization and denormalization are 0.024 and 1.044. And the Average MSE value for the entire subject are 0.064 for normalization and 1.895 for denormalization.

Keywords: artificial neural network, multi layer percetron, multiclass, normalization

1. Introduction

In the academic, determining the strategy of what subjects will be taken in a semester is an important thing. It is needed by every student in order to reach a maximum results in academic achievement.

Determination of that strategy is not an easy thing to do. The solution for the strategy is by predicting the score will be obtained if a student takes a certain subject (choice of subjects). The prediction results can be used to determine what subjects should be taken or not to the next semester.

The estimation values of these subjects are also useful as a form of warning to a student if the prediction turns out to give results that are not good.

For that purpose, it needs a multiclass prediction model. Multiclass case appears because the determination of the student's score are in range of A through E. One method that can be used is the Artificial Neural Network (ANN) with a double layer perseptron or known as the Multi Layer Perceptron (MLP).

ANN is a model that capable of recognizing patterns of an object[1]. With the ability of the learning process, ANN can recognize an object without having to introduce these objects before. This capability can be used in making predictions as in the case of this final task.

The Selection MLP as ANN model based on the results of the case which the distribution of prediction values cannot linearly separable.

2. Artificial Neural Network

2.1 Biological Neural Nets

ANN is a network architecture adapted from the workings of the nervous system of human biology. Performance of human nerve itself is controlled by the brain, where it has a very complex structure and incredible ability. The structure of a biological nerve can be seen in the picture below.

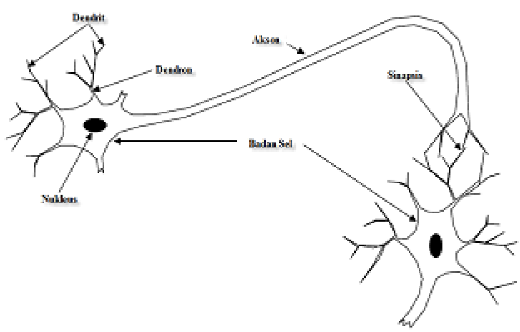


Figure 1. Biological Neural Nets

The brain consists of neurons and links called synapses. Neurons are the smallest units in the human nervous system which will form the neural network if the neurons are connected to each other. Neurons receive and process electrical signals from neighboring neurons to then be forwarded to other neighboring neurons. A neuron has four components that are structured, the dendrites, cell body, axons, and synapses. Dendrite serves as a signal receiver of several neighboring neurons and passes through a thin fiber-called dendron. Signals that have been collected in different dendrons then processed in the cell body to produce output signals. The output signal is transmitted to other neurons through a long fiber called the axon. At the end of the axon of a neuron, there are synapses regulate the flow of strong or weak signals from the recent neuron to another neuron destination.

Human neural network has a very complex structure seen from the numerous of neurons, work organization between neurons, the speed signal/information processing, and rhe energy efficiency levels required. With such complexity, the neural nets in the human brain is able to recognize patterns of objects, perform calculations, and regulate the other body organs in a very fast time.

2.2 ANN Definition

Artificial Neural Network (ANN) is an information processing system is based pemodelannya workings of the human neural network. ANN is made by imitating the activities which occur in human neural nets as to understand, store and recall information that had been learned.

Like the human brain, ANN does the learning activity to recognize objects by using certain training rule to find the connection weights. ANN learns to recognize patterns of incoming information as input into the architecture, this process is referred as "learning".

ANN is formed as a generalization of the mathematical models from biological neural networks. Some factors that determine the ANN system is :

- relationships pattern between neurons, called the network architecture
- weight determination method, called the method of training / learning / algorithms
- activation function

2.3 Multi Layer Perceptron Architecture

ANN architecture is a network models the connectivity between neurons in a system. Each architecture has a different learning algorithm. This research is only focused on the Multi-Layer Feedforward Network architecture, which is also called as Multi Layer Perceptron (MLP)

MLP is a nonlinear model of the architecture consists of a number of neuron units are arranged to form a multilayer. The complexity of the MLP is changed by adjusting the number of hidden layer neurons and the number of units in each layer.

MLP is one of the ANN architecture with its learning process are categorized into supervised learning, namely supervised learning process using a specific target.

Commonly, MLP network models using three or more layers, input layer to receive input variables used in the classification procedure, hidden layer, and output layer. The major things in MLP network is when the data from a training pattern is inserted into the input layer, the neuron units do the calculation process inside the existing layers are sequentially from the first layer until the end to obtain the output value of each layer. This output signal should indicate the appropriate class for a given input data.

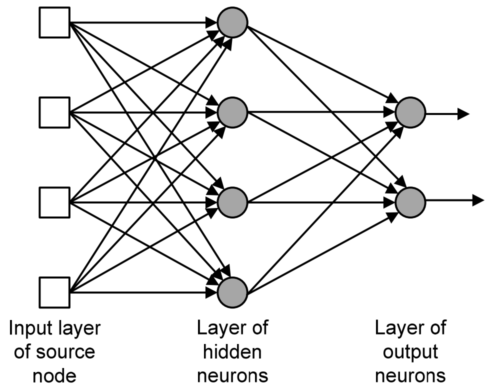


Figure 2. Multi Layer Feedforward Network

Each neuron processing in a particular layer connected with each neuron in the layer above and below through the weight that describes briefly about the network behavior. These weights value will change during the training process to obtain the best weight value that can provide a satisfactory output at the output node.

2.4 Algoritma Pembobotan Backpropagation

One of the learning processes using MLP network architecture is the back propagation algorithm.

This algorithm has two stages of calculation, i.e. the forward calculation and backward calculations. Forward calculation (from the input layer to output layer) calculates the error value from network output value and target value. The function of the error value is then propagated back to each layer in the network to improve synaptic weights value in each neuron. This process is called as backward calculation because the calculation is done sequentially start from the output layer until the first hidden layer. Weight adjustment policies in the backpropagation algorithm is derived from the principle of steepest descent approach about the search of minimum multi-valued function[2].

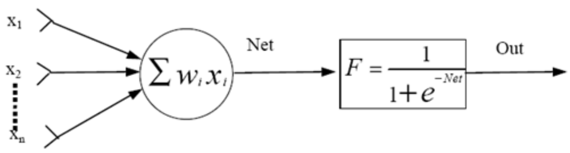


Figure 1. Representasi Neuron yang Digunakan dalam Algoritma *Backpropagation*

Neurons in the backpropagation algorithm consists of two parts (Fig. 3).

Circle $\sum wi\ xi$ represents the sum number of multiplication between the input value of each neuron in one layer with their weights. A value of i starts from 1 to n where n is the number of inputs received by a neuron in a layer.

At the checker is a function that serves to define the value of neuron output. In the backpropagation algorithm, it is known as activation function. The selection of specific mathematical function as the activation function is basically adjusted by the problems encountered, but there are some basic requirements that must be met, namely continuous, differenttialized easily, and it is a not going down function. One function that meets these requirements is the sigmoid function as shown in Figure 3. This function is very popular as an activation function in the backpropagation algorithm, because of its characteristics that continue even in the wide intervals

2.5 Learning on Backpropagation

In general, the stages backpropagation learning algorithm is described as follows

a. Initialize the network parameters

In the beginning, some parameter values will initialize to form ANN architecture. Among them is the number of hidden layers, the number of neurons per layer, termination threshold value, the value of learning rate, and the initial value of weights in each neuron.

Threshold could be the value of the mean squared error (MSE), the root mean square error (RMSE), or others. Initial value of weights of each neuron can be raise from random values at any interval.

b. Network training – forward calculation

This process is performed to calculate the output of the weights that have been initiated earlier. Usually this calculation using a particular activation function (eg sigmoid function). If using this function, then the output value of the hidden layer can be denoted as [4]:

$$z_j = \frac{1}{1+e^{-(z\_net_j)}} \qquad (1)$$

where j=1,2, ... the number of neurons in the hidden layer.

With $z\_net_j = \sum wi\ xi$ where w_i is the synaptic weights between two layers (eg input layer and hidden layer or between two hidden layers). The value of the input signal x_i is a matrix of values. Output of these layers (z_j) will be the input values for the next layer. This concept applies until the output layer.

The output value of neuron output layer (y_k) will be compared with the target value. The difference between the two produces an error rate (e.g. error or MSE)

$$E_k = t_k - y_k \qquad (2)$$

where k = 1,2, ... the number of neurons in output layer. MSE than calculated by the following equation:

$$MSE = \frac{\sum E^2}{N} \qquad (3)$$

where E_k is difference value between target value t_k and output value y_k in output layer neuron k. N is the number of input patterns.

c. Network training – backward calculation

MSE values that have been obtaine from the forward calculations are use in backward calculation. Network training process will be complete if accebtable MSE is reach.

MSE value will propagate using the backpropagation algorithm :

$$\delta_k = (t_k-y_k)f'(y_{netk}) = (t_k-y_k)\ y_k(1-y_k) \qquad (4)$$

y_net_k is input to one neuron to k in the output layer. δ_k is the error unit use in the weight change on the next layer.

Calculate the change value of weight with the value of learning rate α

$$\Delta w_{kj} = \alpha.\delta_k.z_j \qquad (5)$$

Calculate the factors of δ based on the errors hidden neurons in each hidden layer z_j.

$$\delta\_net_j = \sum_{k=1}^{m} \delta_k.w_{kj} \qquad (6)$$

m is the number of neurons in output layer.

Calculate the change value of weight in the layer

$$\Delta v_{ji} = \alpha.\delta_j.x_i \qquad (7)$$

i is the number of input neurons into the hidden layer.

The parameters above are used to improve the value of synaptic weights in each neuron in each layer, how:

$$w_{kj}(new) = w_{kj}(old) + \Delta w_{kj} \qquad (8)$$

Such steps are perform in a single training process and will be repeat until the predetermined MSE reaches.

The result of this learning process is the wighting (w) values in each neuron when the network has been work optimally.

2.6 Data Division

As a supervised learning, ANN needs a set of data. The data used in the ANN can be divided into :

1. *Training Set* : used in learning process. In MLP network, training set is used to find the value of the optimal weight
2. *Validation Set* : used to determine the stopping point in the learning process.
3. *Test Set* : used to assess the performance of the model has actually trained.

3. Experiment and Result

3.1 Data Definitions Used

In this research, we built MLP that can predict the values of subject. Predicted value subject (functioning as output) is seventh and eighth semesters (optional subjects). Data input is the value of the first semester trough six semesters.

The amount of data used in this research is 954 data. This data is taking from the students in 2000 until 2005. The data was then made the table with a row are students, and columns (attributes) are subject. Class labels in the data are 12 optional subjects.

In the data tables, we made the reduction of dimension using attribute selection (feature selection). Those attribute will be use as input in predictive models. Attribute selection using information gain (IG), where the chosen attributes are a few attributes with the highest weight IG.

For each optional subject (class label), will be train by MLP with a variety of parameters. Each optional subject will use five highest attributes from attribute selection process.

3.2 Network Architecture

Network architecture used in this research is a MLP with one hidden layer. Neurons in the input layer of five attributes derived from the results of feature extraction process using IG. Neurons in the hidden layer of three layers, obtained from the following formula:

$$\sqrt{numNeuronInput\,.\,numNeuronOutput} \quad (9)$$

Where $numNeuronInput$ represent number of neuron input, and $numNeuronOutput$ represent number of output neuron. In this research, number of output layer used a single neuron.

3.3 Performance

In this research, MLP performance measured by normalizes and denormalizes MSE value. Normalizes MSE is results obtained from the network output in a scale of 0-1 (using the binary sigmoid in the output layer). Denormalizes MSE obtained from the network output is converted to a scale of 0-4 according to the initial students data who have grade AE

3.4 Implementation

In this research, experiment conducted two different scenarios. The first scenario is done using the same architecture with the change of gamma and beta values. The second scenario is different from the architecture of gamma and beta value of the best from the first experiment. Each experiment repeated five times to get the performance.

Some of the gamma and beta values used in the first experiment were as follows :

1. gamma=2, beta=1/3;
2. gamma=2, beta=1/2;
3. gamma=3, beta=1/3;
4. gamma=4, beta=1/4; and
5. gamma=1.5, beta=0.4.

The second experiment conducted using several different architectures by using a constant value of gamma and beta (best of the first experiment). Architecture that is used is as follows :

Ars#1 :
Number of hidden layer : one.
Number of hidden neuron :
$$\sqrt{numNeuronInput\,.\,numNeuronOutput}$$

Ars#2 :
Number of hidden layer : one.
Number of hidden neuron :
$$\frac{2(numNeuronInput.\,numNeuronOutput)}{3}$$

Ars#3 :
Number of hidden layer : two.
Number of hidden neuron first hidden layer:
$$\sqrt{numNeuronInput\,.\,numNeuronOutput}$$
Number of hidden neuron second hidden layer:
$$\sqrt{numNeuronInput\,.\,numNeuronOutput} + 2$$

Ars#4 :
Number of hidden layer : two.
Number of hidden neuron first hidden layer:
$$\sqrt{numNeuronInput\,.\,numNeuronOutput}$$
Number of hidden neuron second hidden layer:
$$\frac{2(numNeuronHidden1.\,numNeuronOutput)}{3}$$

<u>Ars#5 :</u>
Number of hidden layer : two.
Number of hidden neuron first hidden layer:

$$\frac{2(numNeuronInput\ .numNeuronOutput)}{3}$$

Number of hidden neuron second hidden layer:

$$\frac{2(numNeuronHidden1.numNeuronOutput)}{3}$$

3.5 Experiment Result

Experimental results obtained MSE values and the average minimum MSE as shown in table 1 and table 2.

In table 1 the value of the best average MSE's for each class are different. For classes 1, 2, 3, 8, 9, 10, 11, and 12 the best value obtained when using the gamma and beta constants fifth. While classes 4 trough 6 provide varying results. For the average normalized MSE value, three classes gives the same result (best in the fifth beta), but for the average denormalised MSE value, classes 4, 5 and 6 obtained the best results achieved respectively at constant gamma and beta first, second, and third. Although the results obtained vary, in general can be said of gamma and beta value which gives the best result is when gamma = 1.5 and beta = 0.4

As in the first experiment, the results obtained in the second experiment also varied. Average normalizes MSE values for the 1, 2, 6, 7, 9, 10, and 12 classes obtained when using the architecture of the fifth, another class to get the best results when using the first architecture

For the first and second experiment, class 4 is the class with the best MSE value with the lowest MSE value reached 0.023.

Results are caused by varying input values given for each class varies. Can be seen in table 1 values far enough difference between the average normalized MSE in the classes 2, 3 and 4.

For the average denormalizes MSE in each class, best value is not always in the same position with the best value of the average MSE is normalized. That is because the process is not perfect denormalisasi, because the only use conversion process without taking into account other factors such as the level of accuracy the number behind the comma and the data distribution factor.

For data resume, we got the best value for normalization and denormalization are 0.024 and 1.044. And the Average MSE value for the entire subject are 0.064 for normalization and 1.895 for denormalization.

4. Conclusion

From the experimental results can be concluded that:

1. Combination gamma 1.5 and beta 0.4 values is the best combination in the experiment. This is shown from table 1, that generally the smallest error value is in this experiment.
2. From the experiment proved that an architecture is not always good for all cases. In the experiment, the best architecture in general is the fifth architecture.
3. Architecture has a stable tendency in some cases. The data indicated that each of the dominant class has the smallest error in one architecture

References

[1] en.wikipedia.org/wiki/artificial neural network

[2] Konar, A. 1999. *Artificial Intelligence and Soft Computing. Behavioral and Cognitive Modelling of the Human Brain*. CRC Press.

[3] Ludeman, L. C., 1987, *Fundamental of Digital Signal Processing,* John Wiley & Sons Inc., Singapore.

[4] Siang, Jong Jek. 2005. *JST & Pemrograman Menggunakan Matlab.* Penerbit Andi: Yogyakarta.

[5] Tomlinson,M., 1999, *"Linear Feedback Shift Register", IEEE Trans. On Communications,* Vol.8 No.3,p.391-396.

[6] <u>http://www.faqs.org/faqs/ai-faq/neural-nets/part3/section-10.html</u>

[7] Xuan , Jianhua. 2007. *Gene Selection forMulticlass Prediction by Weighted Fisher Criterion. EURASIP Journal on Bioinformatics and Systems Biology*,Volume 2007, Article ID 64628, 15 pages.

Table 1 : The Results of Experiments Using The Different Value of Gamma and Beta Parameters

Gamma&Beta Value -->			1	2	3	4	5
Class 1	Normal	Average	0.143	0.064	0.125	0.107	**0.051**
		Best	0.092	**0.042**	0.046	0.044	0.049
	Denormal	Average	3.233	1.602	2.925	2.727	**1.363**
		Best	2.145	**1.073**	1.264	1.183	1.239
Class 2	Normal	Average	0.200	0.111	0.106	0.110	**0.093**
		Best	0.137	0.100	0.103	0.104	**0.093**
	Denormal	Average	4.467	3.012	2.811	2.964	**2.712**
		Best	3.204	2.821	2.697	2.899	**2.688**
Class 3	Normal	Average	0.074	0.043	0.042	0.044	**0.037**
		Best	0.067	0.041	0.041	0.044	**0.037**
	Denormal	Average	1.821	1.265	1.266	1.325	**1.178**
		Best	1.610	1.229	1.188	1.296	**1.142**
Class 4	Normal	Average	0.029	0.024	0.024	0.024	**0.024**
		Best	0.028	0.024	0.024	0.024	**0.024**
	Denormal	Average	**1.035**	1.086	1.064	1.094	1.082
		Best	**0.989**	1.009	1.021	1.016	1.048
Class 5	Normal	Average	0.119	0.088	0.089	0.089	**0.086**
		Best	0.114	0.087	0.089	0.088	**0.086**
	Denormal	Average	3.072	**2.464**	2.516	2.486	2.549
		Best	2.920	**2.429**	2.475	2.466	2.532
Class 6	Normal	Average	0.067	0.051	0.050	0.050	**0.049**
		Best	0.064	0.049	**0.048**	0.048	0.048
	Denormal	Average	2.024	1.636	**1.575**	1.597	1.604
		Best	1.993	1.571	**1.511**	1.564	1.589
Class 7	Normal	Average	0.082	0.064	0.063	0.064	**0.063**
		Best	0.079	0.063	0.062	0.063	**0.061**
	Denormal	Average	2.241	1.941	1.939	**1.937**	1.952
		Best	2.140	1.920	1.922	**1.899**	1.908
Class 8	Normal	Average	0.113	0.084	0.083	0.084	**0.080**
		Best	0.104	0.080	0.082	0.082	**0.079**
	Denormal	Average	2.926	2.440	2.432	2.418	**2.261**
		Best	2.713	2.358	2.411	2.275	**2.243**
Class 9	Normal	Average	0.086	0.065	0.062	0.065	**0.058**
		Best	0.083	0.065	0.062	0.065	**0.058**
	Denormal	Average	2.487	2.018	1.924	2.016	**1.734**
		Best	2.429	2.018	1.913	2.016	**1.734**
Class 10	Normal	Average	0.094	0.087	0.093	0.092	**0.078**
		Best	0.093	0.082	0.092	0.092	**0.078**
	Denormal	Average	2.410	2.365	2.406	2.390	**2.098**
		Best	2.374	2.314	2.362	2.383	**2.089**
Class 11	Normal	Average	0.069	0.061	0.062	0.063	**0.060**
		Best	0.067	0.061	0.062	0.063	**0.060**
	Denormal	Average	2.035	1.814	1.922	1.920	**1.803**
		Best	2.011	1.803	1.922	1.920	**1.803**
Class 12	Normal	Average	0.141	0.122	0.124	0.125	**0.100**
		Best	0.137	0.115	0.118	0.113	**0.099**
	Denormal	Average	3.587	3.329	3.319	3.291	**2.895**
		Best	3.550	3.197	3.264	3.101	**2.791**

Tabel 2 : Research Results with a Different Architecture

Architecture -->			1	2	3	4	5
Class 1	Normal	Average	0.049	0.048	0.052	0.051	**0.046**
		Best	0.045	0.048	0.046	0.044	**0.044**
	Denormal	Average	1.296	1.302	1.432	1.386	**1.217**
		Best	1.138	1.264	**1.112**	1.193	1.147
Class 2	Normal	Average	0.093	0.093	0.095	0.097	**0.076**
		Best	0.093	0.093	0.093	0.093	**0.076**
	Denormal	Average	2.716	2.762	2.773	2.996	**2.138**
		Best	2.645	2.741	2.706	2.750	**2.099**
Class 3	Normal	Average	**0.037**	0.038	0.040	0.039	0.062
		Best	0.037	0.037	**0.035**	0.036	0.060
	Denormal	Average	**1.159**	1.221	1.256	1.273	1.803
		Best	**1.151**	1.165	1.161	1.225	1.803
Class 4	Normal	Average	**0.024**	0.024	0.024	0.024	0.116
		Best	**0.024**	0.024	0.024	0.024	0.097
	Denormal	Average	**1.063**	1.078	1.212	1.174	3.306
		Best	1.048	**1.044**	1.154	1.115	2.803
Class 5	Normal	Average	**0.086**	**0.086**	0.088	0.087	0.091
		Best	0.086	**0.086**	0.087	0.086	0.089
	Denormal	Average	2.532	**2.527**	2.615	2.620	2.700
		Best	2.518	**2.514**	2.544	2.569	2.645
Class 6	Normal	Average	0.048	0.048	0.050	0.049	**0.036**
		Best	0.048	0.048	0.049	0.049	**0.035**
	Denormal	Average	1.602	1.542	1.626	1.625	**1.153**
		Best	1.594	1.518	1.626	1.622	**1.140**
Class 7	Normal	Average	0.062	0.062	0.064	0.064	**0.024**
		Best	0.058	0.058	0.058	0.059	**0.024**
	Denormal	Average	1.944	1.972	2.074	2.004	**1.226**
		Best	1.734	1.734	1.734	1.734	**1.131**
Class 8	Normal	Average	**0.079**	0.079	0.082	0.080	0.086
		Best	**0.079**	0.079	0.079	0.079	0.086
	Denormal	Average	**2.255**	2.283	2.342	2.299	2.581
		Best	**2.241**	2.261	2.252	2.261	2.530
Class 9	Normal	Average	0.058	0.058	0.059	0.061	**0.050**
		Best	0.058	0.058	0.058	0.059	**0.049**
	Denormal	Average	1.734	1.734	1.734	1.734	**1.626**
		Best	1.734	1.734	1.734	1.734	**1.626**
Class 10	Normal	Average	0.079	0.079	0.077	0.079	**0.064**
		Best	0.078	0.078	0.076	0.076	**0.062**
	Denormal	Average	2.122	2.122	2.121	2.222	**1.986**
		Best	2.117	2.085	2.085	2.099	**1.911**
Class 11	Normal	Average	**0.060**	0.060	0.061	0.061	0.080
		Best	**0.060**	0.060	0.060	0.060	0.079
	Denormal	Average	**1.803**	1.803	1.803	1.803	2.278
		Best	**1.803**	1.803	1.803	1.803	2.252
Class 12	Normal	Average	0.100	0.099	0.131	0.115	**0.060**
		Best	0.099	0.097	0.100	0.097	**0.059**
	Denormal	Average	2.904	2.788	3.646	3.228	**1.734**
		Best	2.846	2.734	2.899	2.800	**1.734**

PARTICLE SWARM OPTIMIZATION ALGORITHM TO OPTIMIZE PROJECT RESOURCE SCHEDULING

Evi Ria Zerda[1], Suyanto[2]

Fakulty of Informatics, Institut Teknologi Telkom
[1]evie_amth@plasa.com, [2]suy@ittelkom.ac.id

Abstract

Resource scheduling is one of the project operational plannings. It identifies amount and type of resources according to the activities to be scheduled. Requirements of the resources for each time can be fluctuative. Therefore, some high fluctuative requirements during project tour must be reduced or avoided. This paper focuses on implementation of the particle swarm optimization (PSO) to optimize project resource scheduling and investigation for its performance compare to Rank-based Ant System (AS_{rank}) algorithm. Three project data (Qureshi, LOGON, and Building) with different complexities, are used to measure their performances. Computer simulation, for 30 times running, shows that PSO can reduce fluctuative requirements significantly. It reduces the fluctuation up to 50% for Qureshi project, 75% for LOGON, and 59.3% for Building. Furthermore, PSO produces much lower fluctuative requirements than AS_{rank}. For Qureshi project, PSO produces average fluctuation of 19, but AS_{rank} produces 25. For Building project, PSO produces average fluctuation of 7705, whereas AS_{rank} produces 8285. PSO and AS_{rank} have same results for LOGON project. PSO has some sensitive parameters to be carefully defined. C1 and C2 are learning factors that represent the number of potential solution particles. Those parameters are quite sensitive. From the simulation using the three project data, PSO show the best performance when C1 = 2 and C2 = 0.5. Another sensitive parameter is inertia weight, used to control the global and local exploration of particle's value. When the inertia weight is too high, PSO will not find the optimal solution because it explores new areas and local exploration is never performed. But, if inertia weight is too low, PSO will converge to a local optimum.

Keywords: particle swarm optimization, project resource scheduling, resource leveling, fluctuative requirements

1. Introduction

Resource scheduling is one of the project operational plannings. It identifies amount and type of resources according to the activities to be scheduled. Requirements of the resources for each time can be fluctuative. This fluctuation impact to budget because there are time that resources are not used but cost must be paid or increasing recruitment intensity and layoff resources which need more costs. Therefore, some high fluctuative requirements during project tour must be reduced or avoided, or in the other word the resource leveling is used to produce a solution to the problem that can minimize budget.

The particle swarm optimization algorithm was first described in 1995 by James Kennedy and Russell C. Eberhart. PSO is initialized with a population of M random particles and then searches for best position (solution or optimum) by updating generations until getting a relatively steady position or exceeding the limit of iteration number (T). In every iteration or generation, the local bests and global bests are determined through evaluating the fluctuation values of the current population of particles. The updating mechanism of a population of particle's status from the ones of the last generation during search process based on particle's new velocity and particle's previous position.

Resource leveling is delay not critical activity with use positif slack for decrease high request and fill low request. Resource leveling can be done with shift each activity that have slack. Slack from one activity can be used to calculate amount of probability that activity can be shift. If the network is not too large and there are only a few resources, the leveling process can be done manually. For larger networks and multiple resources, resource leveling becomes extremely complex, far beyond the power of manual solutions. That is why resource leveling requires to be done by AI method to find best schedule with minimum fluctuation.

There are several goals of this research. First is to analyze best parameters of the PSO algorithm, such as C1 (learning factor for particle), C2 (learning factor for swarm), inertia weight (parameter of deriving velocity to avoid stagnant of particle in optimum local). Second is to apply the PSO algorithm on the project resource scheduling problem and minimize the fluctuation. Third is to compare optimize result between Rank-based Ant System (AS_{rank}) algorithm and Particle Swarm Optimization (PSO) Algorithm for three project data (Qureshi, LOGON, and Building). Figure 1 shows the flow chart of PSO-Project Resource Scheduling system.

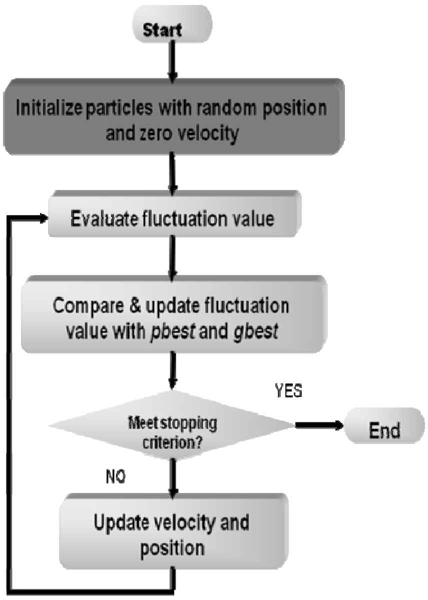


Figure 1 Flow Chart of the PSO-Project Resource Scheduling System

The paper is organized as follows. Particle swarm optimization (PSO) in section 2. Description of System In section 3. Representations of particle position vector, solution space and velocity in section 4. Fluctuation Function in section 5. Experimental Result in section 6.

2. Particle Swarm Optimization (PSO)

The particle swarm optimization algorithm was first described in 1995 by James Kennedy and Russell C. Eberhart Proceedings of the IEEE International Conference on Neural Networks, Perth, Australia [9]. Particle swarm optimization is a population based stochastic optimization technique, inspired by social behavior of bird flocking or fish schooling [8].

Particle swarm optimization algorithm simulates the behaviors of bird flocking. Suppose the following scenario: a group of birds are randomly searching food in an area. There is only one piece of food in the area being searched. All the birds do not know where the food is. But they know how far the food is in each iteration. So what's the best strategy to find the food? The effective one is to follow the bird which is nearest to the food.

Particle swarm optimization algorithm is searching method based on population and global optimization algorithm which best solution can be represented as a point or surface in a N-dimensional area [7].

Formula of velocity update (V_i) :

(1)

Formula of position update (X_i) :

(2)

3. Description of System

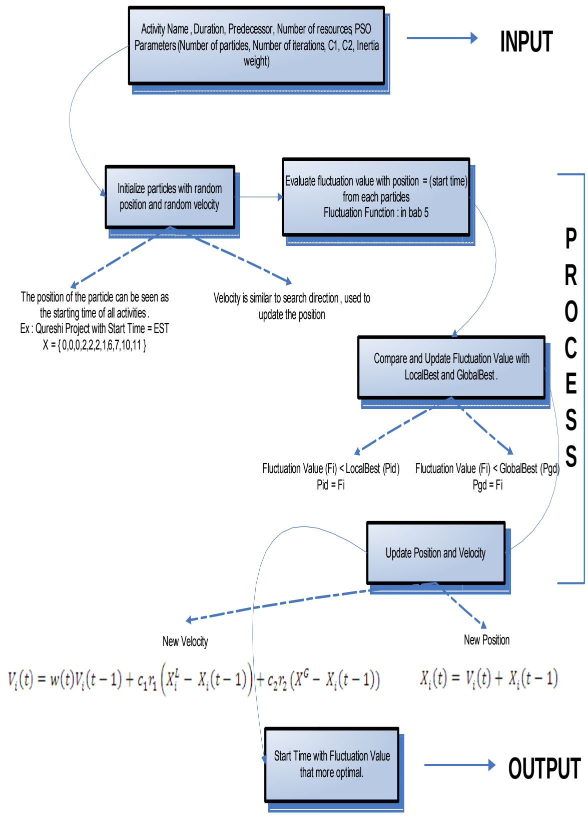


Figure 2 Description of Optimization System on Project Resource Scheduling

4. Representations of particle position vector, solution space and velocity

The goal of resources leveling is to find a network plan, of which the starting time parametric portfolio can make the consumption of resources achieve a balanced state, namely to make the variance of the resource consumed smallest. Therefore, the position of the particle can be seen as the starting time of all activities, and the solution space is a set of the starting time of all activities.

PSO is initialized with a population of M random particles and then searches for best position (solution or optimum) by updating generations until getting a relatively steady position or exceeding the limit of iteration number (T). In every iteration or generation, the local bests and global bests are determined through evaluating the fluctuation values of the current population of particles. Each particle is treated as a point in a N-dimensional space, N is number of activities from project.

The N-dimensional position for the i particle in the t iteration can be denoted as :

$$X_i(t) = \{ x_{i1}(t), x_{i2}(t), \dots, x_{iN}(t)\}. \quad (3)$$

Similarly, the velocity, also a N-dimensional vector, for the i particle in the t iteration
can be described as :

$$V_i(t) = \{ v_{i1}(t), v_{i2}(t), \dots, v_{iN}(t)\} \quad (4)$$

The following equation :

$$v_{k+1}^i = w\, v_k^i + c_1\, Rand\,()\,\big(x_{id} - x_k^i \big) + c_2\, Rand\,()\,(x_{gd} - x_k^i) \quad (5)$$

$$X_i(t) = V_i(t) + X_i(t-1) \quad (6)$$

can represent the updating mechanism of a population of particle’s status from the ones of the last generation during search process.

Equation (5) is used to calculate the particle’s new velocity according to its previous velocity and the distances of its current position from its own best experience or position and the group’s best experience or position. Then the particle flies toward a new position according to equation (6).

Particle just exploration feasible Start Time, between the Latest allowable Finish Time (LFT) and the Earliest possible Starting Time (EST) from each activity, which particle would not fly over the latest starting time. Which :

- Number of Particle $i = 1, 2, \dots, M$
- Number of Iteration $t = 1, 2, \dots, T$
- N = Number of Activity
- $X_i^L = \{ X_{i1}^L, X_{i2}^L, \dots, X_{iN}^L\}$
  Represent local best (solution or position) from i particle that related with best fluctuation value after iteration t-1.
- $X^G = \{ X_1^G, X_2^G, \dots, X_N^G\}$
  Represent global best that achieved from all population.
- c_1 and c_2 are *learning factor*.
- $r_1\ and\ r_2$ are two random vectors with each component generally a uniform random number between 0 and 1.
- The inertia weight (w) is employed to control the impact of the previous history of velocities on the current velocity, thus to influence the trade-off between global (wide-ranging) and local (nearby) exploration abilities of the "flying points".

5. Fluctuation Function

Resource usage per days that expected is constant along the average line. For example, if the resource type is worker, the costs of hiring and layoff are quite significant.

Writer use definition fluctuation from Szendroi [17]. That graphic have fluctuation are graphic non quasi concave. Difference between graphic quasi concave and graphic non quasi concave are shown in this figure below :

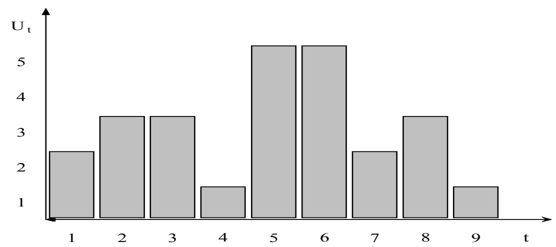


Figure 3 Graphic non quasi-concave

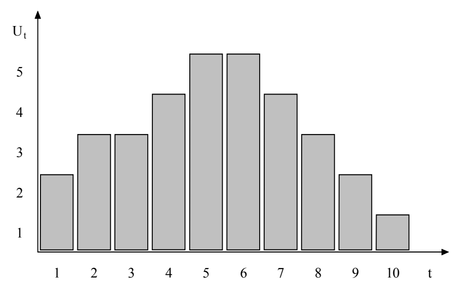


Figure 4 Graphic quasi-concave

Graphic said quasi-concave if value of t0 (high level), for each value t < t0, value of Ut never go down, but for each t > t0, value of Ut never go up.

So fluctuation in graphic of project resources are amount of resources that not use each time. The following shown in this figure below :

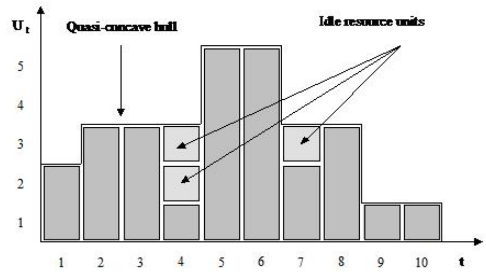


Figure 5 The quasi-concave resource profile and the idle resource units

But, just with information of quasi concave not enough to get more optimal graphic. So writer add information about high level of resources requirement. If the project requires more than one resource, then the fluctuation function are the sum of each resource. So fluctuation function are accumulation from waste of resources and high level of resources requirement, Therefore, fluctuation function can be formulated as follows :

$$Fluktuasi = \sum_{R=0}^{r}((\sum_{T=0}^{t}|x_{rt} - y_{rt}|) + \max y_r) \quad (7)$$

Where R is collection of type of resources, T is time limit of project processing, x is the quasi-concave resource and y is requirement of resource that scheduled. So $|x_{rt} - y_{rt}|$ is waste of resource, and $\max y_r$ is high level of resource requirement.

6. Experimental Result

Particle Swarm Optimization Algorithm was tested on a variety of PSO parameters to find the best parameters. Each experiment done for 30 times running.

Parameters that used are :

- *m* : number of particle and *nc* : number of maximum iteration.
  Experimental of *m* and *nc* values are **{ 60, 1000 }** dan **{ 100, 600 }**
- C1 : learning factor for particle
  Experimental of C1 values are
  C1 = { 0.5, 1, 1.5, 2, 4 }
- C2 : learning factor for swarm
  Experimental of C2 values are
  C2 = { 0.5, 1, 1.5, 2, 4 }
- W *(inertia weight)* : parameter of deriving velocity to avoid stagnant of particle in optimum local.
  Experimental of W values are
  W = { 1.2, 1.1, 1, 0.9, 0.8 }

Test done with see combination of :

- Number of particle and number of maximum iteration { 60, 1000 } with value of C1, C2 and W so there are 125 combination.
- Number of particle and number of maximum iteration { 100, 600 } with value of C1, C2 and W so there are 125 combination.

Based on experimental result for *m*=60 and *nc*=1000 so be obtained parameters C1, C2 and W are **{ 2, 0.5, 0.9 }** with average fluctuation value 27.27 but experimental result for *m*=100 and *nc*=600 so be obtained parameters C1, C2 and W are **{ 2, 0.5, 1 }** with average fluctuation value 27.27. so best parameters are used :

Table 1 Best Parameter for Logon Experiment

m	*nc*	*c1*	*c2*	*w*
60	1000	2	0.5	0.9

Reexamine of each case with best parameter which get from previous analysis result to proof that particle swarm optimization algorithm can give real optimal solution**.**

It reduces the fluctuation up to 50% for Qureshi project, 75% for LOGON, and 59.3% for Building, with minimize fluctuation, resource graphic have more lower high level and waste can be decreased.

Furthermore, PSO produces much lower fluctuative requirements than *Ant Colony Optimization* khususnya *Rank-Based Ant System* (AS_{rank}). For Qureshi project, PSO produces average fluctuation of 19, but AS_{rank} produces 25. For Building project, PSO produces average fluctuation of 7705, whereas AS_{rank} produces 8285. PSO and AS_{rank} have same results for LOGON project.

From experimental result about each project that tested, can be conclusion that particle swarm optimization algorithm can be used for get optimum solution from project resource scheduling optimize problem and produce more optimal solution if compared with Rank Based Ant System (AS_{rank}) algorithm.

References

[1] A. Carlisle and G. Dozier, 2001, *An Off-The-Shelf PSO*, In *Proceedings of the Particle Swarm OptimizationWorkshop*, pages 1–6.

[2] Clerc, M, 1999, *The swarm and the queen: towards a deterministic and adaptive particle swarm optimization*, Proceedings, 1999 ICEC, Washington, DC, pp 1951-1957.

[3] Departemen dalam negeri, 2007, *Modul 1 Pengertian Dasar Manajemen Proyek*, Lembaga Administrasi Negara.

[4] Gray, Clifford f and Erik W, 2006, *Manajemen Proyek Proses Manajerial,* Yogyakarta : Penerbit Andi.

[5] https://courses.worldcampus.psu.edu/welcome/pmangt/samplecontent/520lesson08/lesson08_01.html

[6] http://en.wikipedia.org/wiki/Particle_swarm_optimization.html

[7] http://en.wikipedia.org/wiki/Swarm_intelligence.html

[8] http://www.swarmintelligence.org/tutorials.php

[9] Hasan, Rania, 2004, *Particle Swarm Optimization : Method and Application,* Engineering Systems Division - Massachusetts Institute of Technology.

[10] Hong Zhang, Xiadong Li, Heng Li and Fulai Huang, *Particle swarm optimization for resource-constrained project scheduling, Automation in Construction*, pp. 393-404, 2004.

[11] Kennedy, J. and Eberhart, R, (1995), *Particle Swarm Optimization*, IEEE International

Conference on Neural Networks (Perth, Australia), IEEE Service Center, Piscataway, NJ, IV: 1942-1948.

[12] Luthan, Putri Lyanna A. and Syafriandi, 2005, *Aplikasi Microsoft Project untuk Penjadwalan Kerja Proyek Teknik Sipil*, Penerbit Andi. Yogyakarta.

[13] Muslim, Irfan, 113027263, *Optimasi Penjadwalan Sumber Daya Proyek Menggunakan Algoritma Genetik.* 2005, Bandung : Sekolah Tinggi Teknologi Telekomunikasi.

[14] Nicholas, John M, 1990, *Managing Business and Engineering Projects : Concepts and Implementation*, New Jersey : Prentice-Hall.

[15] Priyanto, Aruna Anggayasti, 2008, 113040076, *Penerapan Algoritma Ant Colony Optimization (ACO) pada Optimasi Penjadwalan Sumber Daya Proyek,* Bandung : Institut Teknologi Telekomunikasi.

[16] Qureshi, Liaqat Ali, 2007, *Resource Allocation & Leveling,* UET Taxila.

[17] Shi, Y. H., Eberhart, R. C., (1998), *A Modified Particle Swarm Optimizer*, IEEE International Conference on Evolutionary Computation, Anchorage, Alaska, May 4-9, 1998.

[18] Shi, Y. H., Eberhart, R. C., (1998), *Parameter Selection in Particle Swarm Optimization*, The 7th Annual Conference on Evolutionary Programming, San Diego, USA.

[19] Soeharto, Iman, 1990, *Manajemen Proyek (Dari Konseptual Sampai Operasional)*, Jakarta : Penerbit Erlangga.

[20] Spinner, M. Pete, 1992, *Elements of Project Management: Plan, Schedule and Control,* New Jersey. Prentice Hall.

[21] Szendroi, Etelka, 2006, *A Resource Leveling MILP Model For Multi-Mode Projects Based On Global Measure,* University of Pecks Pollack Mihaly Faculty of Engineering.

[22] Xing-mei, Li, Zhang Li-hui, QI Jian-xun and Zhang su-fang, 2008, *An extended particle swarm optimization algorithm based on coarse-grained and fine-grained criteria and its application*, China.

VISUALIZATION COMPARISON SELF ORGANIZING MAPPING (SOM) ANALYSIS TO LINEARITY CORRELATION BASED ON REGION POWER TRANSMISSION PATH IN JMB

Iriansyah BM Sangadji [1], Subanar [2] , Retantyo Wardoyo [3], Sri Hartati [4]

[1] Informatics, STT PLN
[2,3,4], Natural Science and Mathematics Faculty, Gadjah Mada University
[1]iriansyach@sttpln.ac.id , [3]r_wardoyo@mailcity.com, [4]shartati@ugm.ac.id

Abstract

Electricity transmission path function is connecting numbers of power plant in the large covering area. Its mean the system needs handling when the power plant must active or not. Decision is taking by consider analytic behavior and relation between power plant in each region. Knowledge about relation pattern and behavior of power need are important. Neural network (SOM) and scatterplot methods can use in this paper. Comparison of form graph between scatter plot and SOM visualization mean will be investigated. Results of this paper will showing and explain behavior relationships pattern between region 1 to region 4. Comparison method between 2 approaches will showed.

Keywords: Self Organizing Map(SOM), Linear Correlation , Region, JMB electricity transmission path.

1. Introduction

In operation electricity power handling system that the power plants , transmission path is connected in a big electricity complex system. The area of this system covering increase will correlate with integrated new numbers kind of power plants. Roles of addition or develop its system have feedback in positive or negative values to system.

To constitute a new power plants have to consider much aspects from the handling system itself to social and environment impacts. Now the question is how to make interactive between power plants unit in harmony of transmission balancing system. Balancing system has a few conditions. One condition is have to know behavioral active power from power plant and consumption in customer .

JMB (Jawa, Madura Bali) electricity transmission system is the one of complex network handling system. Its constitute from four convergent regions division power. Operating power JMB system have to based on region1 to region 4.

Figure 1. JMB electricity Transmission System

Modeling and mapping behavioral power per region areas will explain behavioral of dynamic region power.

There are many approaches and models to identify, recognize behavior pattern of electricity consume in region1 to region4. Statistically method is the most popular in using today.

This paper tray to explain compare about correlation linearity from statistically with scatter plot and neural network methods. Self Organizing Map is the one of Neural network approach. Results will show visualization of linearity as graph. Reason of linearity behavioral found is to modeling the next purpose in controllability study. It's a part of handling system

1.1 Self Organizing Map

The SOM is an unsupervised neural network, based on competitive learning, that implements a non linear smooth mapping of high dimensional input space onto low dimensional output space. The neurons of the SOM form a topologically ordered low- dimensional lattice that is an outstanding visualization tool to extract knowledge about the nature of the input space data.

The SOM is widely used as a data mining and visualization method for complex data sets. Application areas include, for instance, image processing and speech recognition, process control, economical analysis, and diagnostics in industry and in medicine.

The SOM algorithm implements its mapping in two stages. First the best matching unit (BMU) of the input vector is selected by means of a competitive process.

$$c = \arg\min_i || x - m_i(t) || ,\ i = 1, 2, ..., N \quad (1)$$

Then a cooperative step is performed, where the winning unit and its neighbors are adapted

$$m_i(t+1) = m_i(t) + \alpha(t)\, h_{ci}(t)[x(t) - m_i(t)] \quad (2)$$

1.2 Scatter Plot

A scatterplot matrix is a matrix of scatterplots where each column contains the same X axis and each row the same Y axis. A scatterplot matrix is useful for visualizing how a dataset is distributed through multiple variables. all of the scatterplots in the matrix the same way, you can see how the same clusters of points change shape from one scatterplot to another

2. Methodology

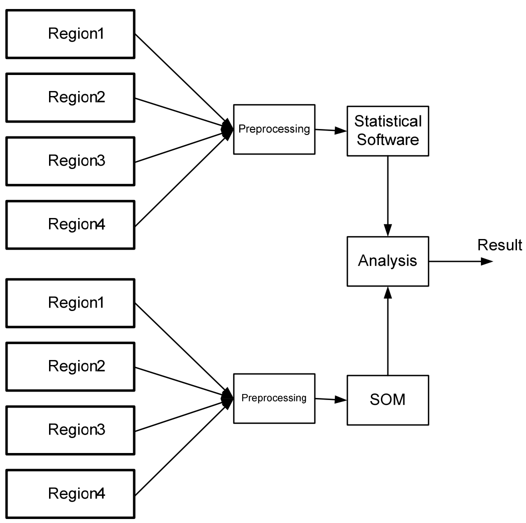
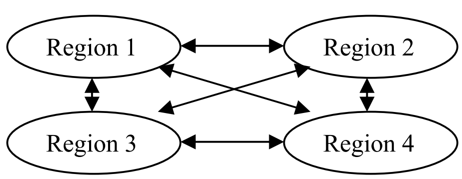


Figure 2. Analysis Methodology

2.1 Scatter Plot Vusualization

Data Prepocessing



Each regions consumption data will be correlating to another except the region itself as figure above.

Table 1. Sample Behavior Power Region in October 2006

Sum Of Power Plants			
Region I	**Region II**	**Region 3**	**Region 4**
AVG	AVG	AVG	AVG
4465	666	1969	4062
4840	678	2084	4834
4876	679	2285	4971
5101	678	2124	4954
5000	656	2318	4943
5122	655	2222	4855
5029	645	1862	4586
4993	689	1420	3948
5652	600	1894	4332
5463	560	2177	4646
5363	592	2192	4736
5359	593	2165	4702
5346	621	2022	4671
5226	612	1944	4328
5020	596	1736	3959
5553	583	2039	4417
5802	568	1644	4695
5668	576	1728	4714
5565	577	1787	4512
5340	582	1484	4370
5004	581	1237	3653
4563	595	1015	3201
4142	631	969	2827
3679	649	1000	2566
3886	640	1019	2503
3926	604	923	2957
3974	592	1034	3240
4224	590	976	3340
4472	600	1100	3138
5220	586	1341	3773
5408	608	1553	3948

Data prepared from daily operations sheet in Pusat Pengatur dan Penyaluran Beban (P3B) PT PLN (persero) from October 2006 – December 2006. Its repsenting regular days and few of holiday.

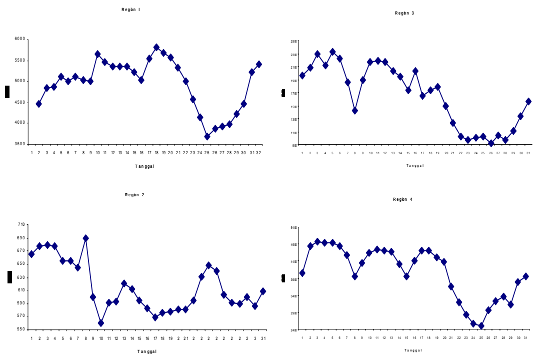


Figure 3. Region Power behavior pattern

Region 1 ↔ Region 2

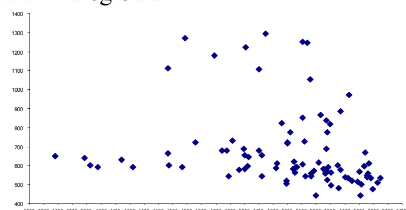


Figure 4. Scatter Plot Correlation Between Region 1 and Region 2

Figure 4 showing that there is no linear correlation between region 1 and region 2 .

Region 1 ↔ Region 3

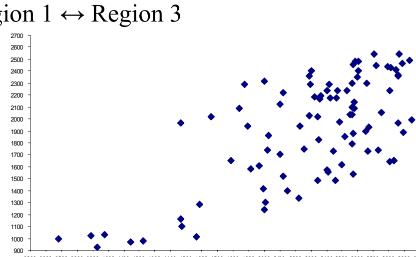


Figure 5 Scatter Plot Correlation Between Region 1 and Region 3

Figure 5 showing that there is a linear correlation between region 1 and region 3

Region 1 ↔ Region 4

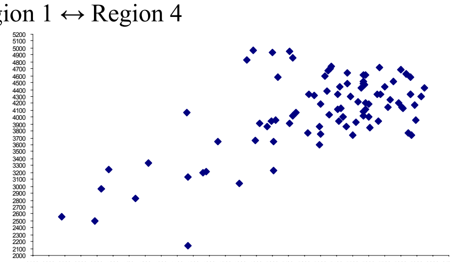


Figure 6 Scatter Plot Correlation Between Region 1 and Region 4

Figure 6 showing that there is a linear correlation between region 1 and region 4

Region 2 ↔ Region 3

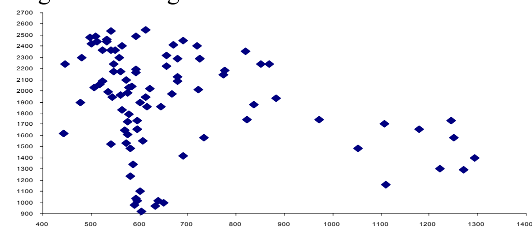


Figure 7. scatter plot correlation between Region 2 and Region 3

Figure 7 showing that there is no linear correlation between region 2 and region 3

Region 2 ↔ Region 4

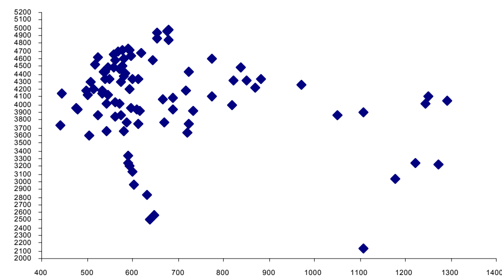


Figure 8. Scatter Plot Correlation Between Region 2 and Region 4

Figure 8 showing that there is no linear correlation between region 2 and region 4

Region 3 ↔ Region 4

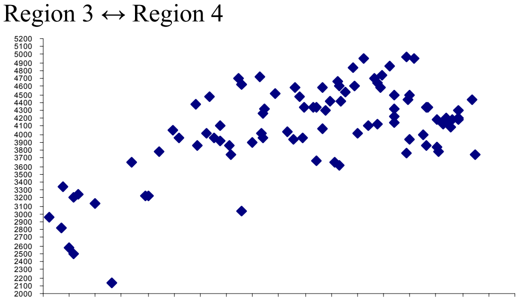


Figure 9 Scatter Plot Correlation Between Region 3 and Region 4

Figure 9 showing that there is a linear correlation between region 3 and region 4

2.2 Self Organizing Map Visualizations

Region 1 ↔ Region 2

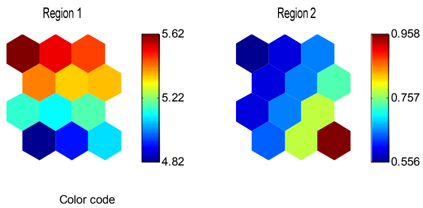


Figure 10. Correlation Between Region 1 and Region 2

Figure 10 showing that there is no linear correlation between region 1 and region 2 .

Region 1 ↔ Region 3

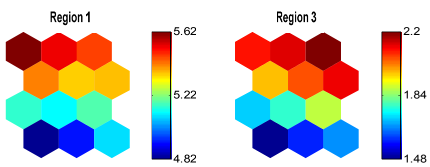


Figure 11. correlation between Region 1 and Region 3

Figure 11. Showing that There is a Linear Correlation Between Region 1 and Region 3

Region 1 ↔ Region 4

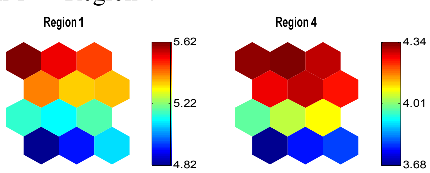


Figure 12. Correlation Between Region 1 and Region 4

Figure 12 showing that there is a linear correlation between region 1 and region 4

Region 2 ↔ Region 3

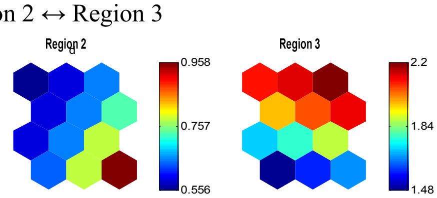


Figure 13. Correlation Between Region 2 and Region 3

Figure 13 showing that there is no linear correlation between region 2 and region 3

Region 2 ↔ Region 4

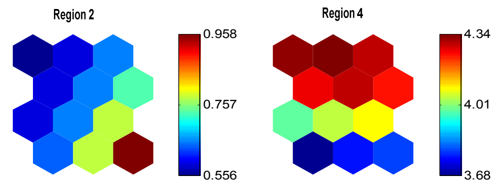


Figure 14. Correlation Between Region 2 and Region 4

Figure 14 showing that there is no linear correlation between region 2 and region 4.

Region 3 ↔ Region 4

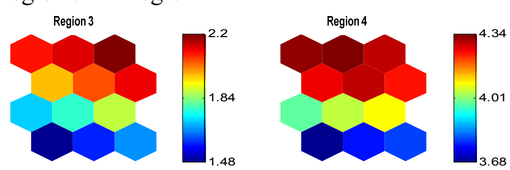


Figure 15. Correlation Between Region 3 and Region 4

Figure 15 showing that there is a linear correlation between region 3 and region 4

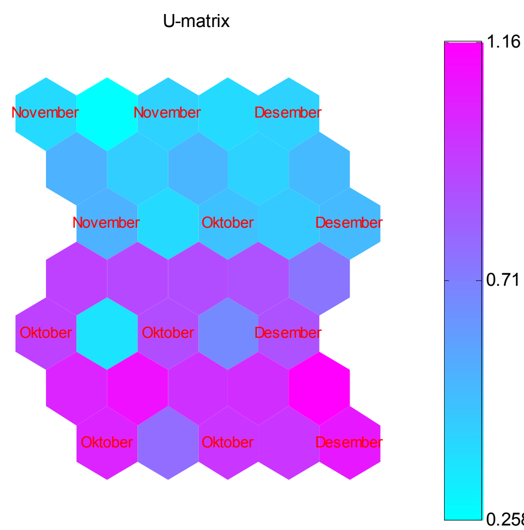


Figure 16. Mapping Between Region 1 - Region 4

3. Result

Region 1 and region 2 based on statistically or SOM results show that power behavior have no linear correlation. Its mean that each region have independent relationships in power consume.

Region 1 and region3 based on above data indicate that power behavior have a linear correlation. Its mean that each region especially region 1 needs power from region3.

Region 1 and region 4 show that region 1 need more power from region 4 because there is a linear correlations.

Region 2 and region 3 show there is no linear correlation. Its mean that each area is independent for power preparing.

Region 2 and region 4 show there is no linear correlation.

Region 3 and region 4 show there is a linear correlation and each area is dependent for power.

Figure 16 showing correlation each region to others. Its mean that region 1 have using power a lot in December. The other side region1 also have great less power in last December. Region 2 in peak power at October but not much than peak in region 1. Region 3 almost likely than region 2 in middle October. Region 4 have most less power in November.

4. Conclusion

Statistically(scatterplot) and neural networks are the same method of visualization its showed correlation linearity sets of data.

Neural networks have a colorfully visualization than scatter plot. Its can be show details the information needs.

Electricity power behavior at each region are different in peak of power and region relationship.

Power plant interactive with transmission lines path to achieve balance system with centralized operator.

Relationship happen between region 1, 3 and region

Based on data indicated that region 2 relative independent in transmission system.

Analysis result can be use to make mapping model in intelligence study area.

References:

[1] Avramovic. B, "On Neural Network Voltage Assessment", Proceedings of INNS/EPRI Summer workshop of Neural Network Computing for the Electric Power Industry, Stanford,CA, Aug 1992.

[2] Bala. BK, "Principles of System Dynamics", www.albany.edu/cp/sds/sdcourse/ , 1999, Aug 2006.

[3] Blackburn, "Protective Relaying Principles and Aplications", Mercell Decker Inc, 1987.

[4] Bose. Anjan., Tomsovic. Kevin, "Automated Operating Procedures for Transfer Limits", www.pserc.wisc.edu/ecow/get/publicatio/reports/2001reports/ , 2001, Agustus 2006.

[5] Bouthiba, "Fault Location in EHF Transmission Lines Using Artificial Neural Network", www.doi.wiley.com/10.1002/etep.5/, Agustus 2006.

[6] Cegrell T, "*Power System Control Technology*", Prentice Hall, 1986.

[7] El-Arraudi, *et al*, "Comprehensive Transmission Distance Protection Setting Using Intelligence Analysis of Events and Consequences", IEEE Transction of Power Delivery, Vol 20 No 3, Juli 2005, p 1817.

[8] El-Keib, Ma, "Aplication of Neural Network in Voltage Stability Assessment", IEEE Transactions of Power System, Vol 10 No 4, November 1995, p 1890 – 1896.

[9] Fausett. Laurene, "Fundamentals of Neural Network: Architecture algorithm and Aplication", Prentice Hall, 1994.

[10] Haykin. Simon, "Neural Network A Comprehensive Faundation", Prentice Hall, 1994.

[11] Han and Kamber, "Data Mining; Concepts and Tecniques", Morgan Kauffman Publisher, 2001.

[12] He. H, Cheng.S, *et al*, "Home Network Power Line Communication Signal Processing Based On Wavelet Packet Analysis", IEEE Transaction On Power Delivery, Vol 20 No 3, July 2005, p 1879

[13] Klir G, Yuan,Bo., 1995, "Fuzzy Set And Fuzzy Logic Theory and Application", Prentice Hall.

[14] Kusumadewi,Sri et al, 2004, " Aplikasi Logika Fuzzy Untuk Pendukung Keputusan", Graha Ilmu, Jogjakarta.

[15] Marsudi, Djiteng., 2005, "Operasi Sistem Tenaga Listrik", Graha Ilmu, Jogjakarta.

[16] Sudibyo, 2005, "Operasi Sistem Tenaga Listrik dan Perangkat pendukungnya", Majalah Teknologi dan Energi Vol 5 No 1 Januari 2005 Hal :77 – 91, STT PLN

[17] Riza, 2005, "Otomatisasi Transmisi dan Distribusi untuk Meningkatkan Kualitas Tenaga Listrik", Majalah Teknologi dan Energi Vol 5 No 1 Januari 2005 Hal :92 – 102, STT PLN

RECONSTRUCTION OF PHYLOGENETIC TREE USING ANT COLONY OPTIMIZATION

Ulfa Hasana[1], Suyanto[2]

Faculty o Informatics, Institut Teknologi Telkom
[1]ulfa.hasana@yahoo.com, [2]suy@ittelkom.ac.id

Abstract

Reconstruction of phylogenetic tree is an important issue in genetic science. It is a fundamental aspect to understand the structure, function, and genetic distances of the DNA of among species. The best phylogenetic tree is the one with minimum score of distances among species. This paper addresses optimization of phylogenetic tree reconstruction using Ant Colony Optimization (ACO). ACO is applied using a fully connected graph constructed using the distance matrix among species. ACO uses the distance matrix to get the best path of the graph. Afterward, the phylogenetic tree is built using pheromone's score from nodes in the path. Some observations are performed to find the optimal ACO's parameters. Afterward, the optimal parameters are used to reconstruct phylogenetic trees. Four datasets, Reptile (9 species), Mammal (16 species), Bird (21 species), and Laurashiteria (29 species), are used to see the performance of ACO. Experiments to the four datasets show that ACO gives very good solutions with minimal score of distances among species. For Reptile dataset, ACO gives a phylogenetic tree with score of 0.650212069. The tree states that *molesink kakapao*, *arenaria grturtle*, *caiman epturtle*, and *haematapus iguana*, have short genetic distance (*sister taxa*). Whereas *morepork* has genetic distance that closer to *bifurcating* and alligator is on the outgroup tree. Furthermore, ACO gives a phylogenetic tree with score of 1.46832902 for Mammal dataset, 2.19170688 for Bird dataset, and 1.956873924 for Laurashiteria dataset.

Keywords: ant colony optimization, phylogenetic tree, genetic distances

1. Introduction

Phylogenetic trees are developed in order to help unveil the genetic distances among species. With analyze a set of amino's species, we can use it to know the relationship among species. Reconstruction of phylogenetic tree is an important issue in genetic science, like many others, it is still an open subject for research.

This paper addresses optimization of phylogenetic trees reconstruction using Ant Colony Optimization (ACO). ACO uses the behavior of ants, how a colony of ants could reach food resource using the shortest path. In reconstruction of phylogenetic tree, ACO focuses on how a colony of ants could unveil the genetic distances between two species and it will continue to the next species until become a path. From this path, we can construct the phylogenetic tree of the dataset.

The main goals of this research are: 1) Build a system that implements a reconstruction of phylogenetic tree; 2) Analyze parameters of ACO algorithm for minimizing tree's score; and 3) Analyze the optimization of ACO algorithm in reconstruction of phylogenetic tree using different types of datasets.

2. Phylogenetic Tree

Optimization is a process to reach an ideal result (effective reachable value). In mathematics, optimization is a study trying to find a minimum or maximum value from a real function. An optimal value is a reachable value from a process and it is the best value from the existing solutions. This optimal value can be found with two methods [9]. The first method is conventional method, try all possibilities and choose the best one. This method is ineffective because need much times [9]. The second method is by using an equation or picture in order to get an optimal value effectively because it faster and accurate and we use this second method to get optimal solution on this paper. Optimal solution is the best solution of a case from the existing solution. In this paper, an optimal solution means a phylogenetic tree that has a minimum score tree from the whole research.

Phylogenetic trees are developed in order to help unveil the genetic distances among species, from the internal species to external species [1]. Reconstruction of a phylogenetic tree becomes a fundamental aspect in biology computation since DNA and amino sequence are easy find. The main goal of phylogenetic tree reconstruction is to optimize the tree score from the root to the whole tree structure. The best score means high probability from the tree that indicates the genetic relationship between species [7].

A phylogenetic tree can be considered as a binary tree whose leaf nodes represent the species to be analyzed and inner nodes the ancestral species from which the current species have evolved. Phylogenetic trees may or may not have a root (see Figure 1) that indicates the oldest ancestor. Usually, a rooted tree better represents the genetic distances between species [1].

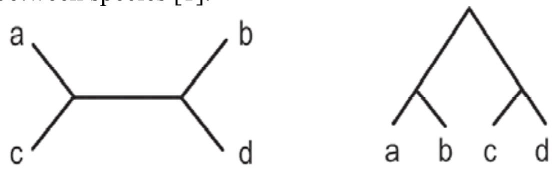


Figure 1 Unroot and Root tree

There are some of phylogenetic tree branching [2]:

1. Poliphyly
   Relationship between bifurcating with sister-taxa from their descendant taxa's family.
2. Paraphyly
   Relationship between sister-taxa with single-taxa of ascendant from it's ascendant.
3. Polytomy
   Relationship between three single taxa.
4. Bifurcating
   Relationship between sister taxa with single taxa as their two descendant.

To define how ant colony optimization (ACO) applied to reconstruction of phylogenetic trees, we used a fully connected graph, constructed using the distance matrix among species (Figure 2). In this graph, nodes represent the species and edges represent the evolutionary distances between species [12].

Cat	ATTTGCGGTA
Dog	ATCTGCGATA
Rat	ATTGCCGTTT
Cow	TTCGCTGTTT

	Cat	Dog	Rat	Cow
Cat	0	0.2	0.4	0.7
Dog	0.2	0	0.5	0.6
Rat	0.4	0.5	0	0.3
Cow	0.7	0.6	0.3	0

Figure 2 Visualization of matrix distances

Ant algorithm is adopted from the behavior of ant colony called ant system (Dorigo, 1996). Naturally, ant colony could find the shortest path from nest to food resources. Ant colony could find shortest route between nest and food resources depend on footprint on the path. If a path that chosen by most of ants, the footprint will be clearer. It means the path that chosen by fewest ants, will be no ants that choose it in a time[10]. Figure 3 shows the destination of the ants to find the shortest path.

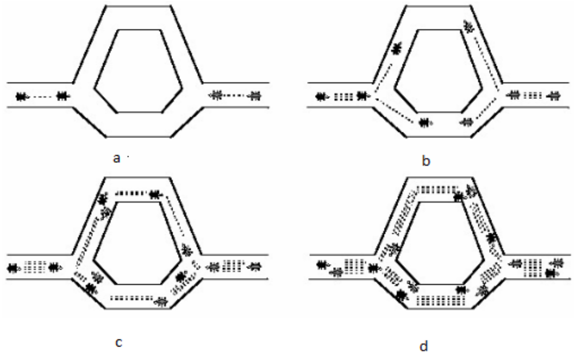


Figure 3 The destination of ants to find food resources

There are two group of ants will try to find food resources (see Figure 3 (a)). A group (called L) is group come from the left side (nest) and the other group (called R) is group come from the right side (food resources). The two of groups divide their group in two. The first part come from upon path, and the other come from below path (see Figure 3 (b)). All of ants walk in the same velocity and leave pheromone or footprint (see Figure 3 (c)). The leaved pheromone of upon path more evaporate than the below path because the number of ants that come to that path fewer than the below path because the distance is longer. Finally, the other ants will decide to choose the below path because the leaved pheromone is bigger than the upon path. (see Figure 3 (d)) [10].

3. Analysis and Design

The main goal of the application of phylogenetic tree reconstruction is implements ACO algorithm in order to help reconstruct phylogenetic tree from the different types of dataset and optimize the score tree using the combination of parameters.

To define how ACO concept applied to phylogenetic tree reconstruction, the steps are:

1. Species is construct into fully connected graph that connected by genetic distance from one species to the other species in dataset [6].
2. The genetic distance was compiled by the distance matrix from the DNA/amino species [7].
3. Initially, ants start in a randomly selected node. Then, they travel across the structured graph, and at each node, a transition function (Equation 1) determines its direction. This equation represents the probability that the *k*-th ant, being at node *i*, goes to node *j* in its next step [7].

$$P_k(i,j)=\frac{[\tau(i,j)]^{\alpha}\cdot\left[{}^{1}\!/_{d(i,j)}\right]^{-\beta}}{\sum_{J_i^k}\left([\tau(i,u)]^{\alpha}\cdot\left[{}^{1}\!/_{d(i,j)}\right]^{-\beta}\right)} \qquad (1)$$

where $P_k(i,j)$ is the probability of transition between node i and j, τ is the pheromone trail between two nodes, $d(i,j)$ is the evolutionary distance between nodes i and j, J is the set of nodes connected to node i and already visited by the k-th ant, and α and β are arbitrary constants. Equation 1 is composed of two terms:

- The first is based on the genetic distance between species i and j.
- The second is based on the accumulated experience - the pheromone trail. This trail is represented as a matrix (like that for the distance between species),whose values are dynamically changed by the algorithm, and determined according to the paths chosen by ants. Therefore, τ (i,j) represents the attractiveness of node j, while the ant is at node i. Therefore, the objective of a given ant is to find a path in the graph that maximizes the transition probabilities, thus obtaining a sequence of species that produces the smallest genetic distance.

4. Equation for genetic distances is represented [7]:

$$d(i,j)=\begin{cases} d(i,u)+[d(i,u)-d(j,u)].\eta, d(i,u)>=d(j,u) \\ d(j,u)+[d(j,u)-d(i,u)].\mu, d(j,u)>d(i,u) \end{cases} \quad (2)$$

Different with a traditional ACO, where moves are made between nodes, our system creates an intermediary node between the two previously selected ones. This node will represent the ancestral species of the other two, and it will not be in the list of nodes (species) to be set in the tree. Where u is a node that does not belong to the set of nodes connected to node i and already visited by the k-th ant. $dnu(i,j)$ is the distance between the new node n and node u, based on the previous distances between (i,u) and (u,j). $d(i,u)$ is the distance between nodes i and u, and η is a scale constant that defines the distance between the new node n and its descendents i and j. Equation 1 and 2 are repeated until all nodes belong to the list of already visited nodes, and then a path is constructed. The score of this path is given by the sum of the transition probabilities of the adjacent nodes of the path [7].

5. Paths constructed by the ants are used for updating the pheromone trail. An increment of the pheromone trail is made at all nodes belonging to at least one path, created in an execution cycle. This key point avoids fast convergence to a local maximum. The pheromone trail matrix is updated according to Equation 3 [7]:

$$\tau(i,j)=\rho.\tau(i,j)+(1-\rho)\Delta\tau(i,j) \quad (3)$$

where 〉is the rate of evaporation of the pheromone that reduces the persistence of the environment to the ants. In this system, the rate of increment of pheromone, ⊗|(i,j), was modified to allow an increment proportional to all the obtained paths, given by the division of the current path and the best path, as shown in Equation 4:

$$\Delta\tau(i,j)=\begin{cases} \sum_{t=0}^{k} S_{c(t)}.(S_{best})^{-1}, (i,j)\in c(t) \\ 0, otherwise \end{cases} \quad (4)$$

where k is the number of ants, $c(t)$ is the path constructed by an ant up to time t, $Sc(t)$ is score of the path $c(t)$, and S_{best} is the score of the best path found up to now. Using this procedure, ants travel through the graph, and at the end of a predefined number of cycles, it is possible to reconstruct the tree using the best path found. Therefore, we can conclude parameters of ACO are:

- Pheromone trail (τij) and the increment (Δτij) must be initialed before the cycles start. τij is used in probability equation that will be travelled by this algorithm. Δτij will be initialed after 1 cycle.
- Constant of ants (α)
- Constant of visibility (β)
- Constant of distance between the new node n and its descendents (η)
- Constant of ant evaporation (ρ)
- Numbers of ants (k)
- Genetic distance d(i,j)
- Number of iteration

4. Testing

Testing is used to determine the effectiveness of phylogenetic tree reconstruction system with ACO algorithm. To do this, we must define the goal of testing. The scenario of testing, the parameters, and the cases that will be used. The result of testing will be analyzed to conclude the functional parameters for ACO algorithm.

The goals of this testing are:

- Analyzed ACO's parameters
- Analyzed the phylogenetic tree reconstruction

The scenarios of the testing are:

- Number of ants and the cycles are **{60, 100}.**
- Alpha: Learning factor of ants. **α = {0.5, 1, 2}.**
- Beta: Learning factor of genetic distances. **β = {0.5, 1, 2}.**
- Rho: Constant of ant evaporation. **ρ = {0.3, 0.6}.**
- Myu: Distance between the new node n and its descendents **η = {0.3, 0.5}.**
- The first pheromone: $\tau_0(i,j)$ **= {0.0001, 0.5, 2}**

We will use all 108 combinations of the parameters for testing. For the datasets, we use reptile dataset as a case to determine the best parameters and we used those parameters for other datasets.

Testing parameters for this case are:

1. Tree Score
   Tree score is score from the root to the whole tree structure as the result of phylogenetic tree reconstruction. The higher value means the worst solution, the lower values means the best solution.
2. Path Score
   Path score is a path form ACO process from one species to the whole species. The higher value means the worst solution, the lower values means the best solution.

Datasets that we used for this testing are:

1. Reptile datasets[14]
   This dataset is consist of 9 species
2. Mammals datasets[14]
   This dataset is consist of 16 species
3. Laurashiteria datasets[14]
   This dataset is consist of 29 species
4. Bird datasets[14]
   This dataset is consist of 9 species

For seeing the behavior of the system as the effect of parameters changing, we used those 108 combination of parameters with 30 trials and use the mean result. This testing used reptile dataset and the result of this testing shown on figure 4:

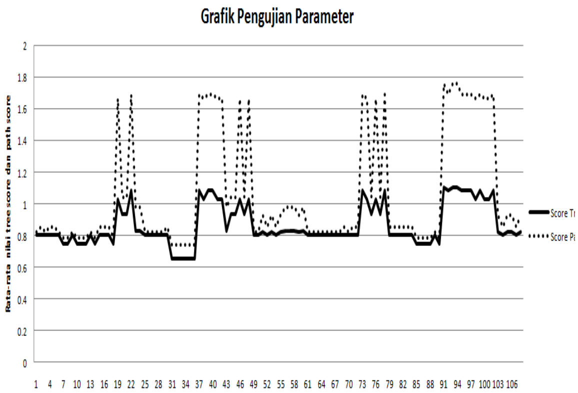


Figure 4 Graphic of parameters testing

From this testing, we could determine the best parameters Alpha and Beta **{0.5, 1, 2}** with tree score **{0.650212069}** and path score **{0.73785252}**

Parameter α controls the exploration of the search space, by weighting the importance of the pheromone trail in the decision of an ant when it arrives at a branch. The algorithm is sensitive to high values of this parameter, the higher value means the ants will be reach the solution faster, but if too big, ants will lead to a fast convergence to a local optimum [13]. The value zero (0) means ant did not communicate to each other [13]. So, from the oldest research decide that alpha will be better if the values are between 0.5 and 2 [7]. From this testing, we can determine that these parameters are so sensitive and give best result when the value is 1 and smaller than beta values.

Parameter β defines the relative importance of the distance between species in the transitions between nodes(genetic distances). In practice, we observed that it has to be higher than α and give best result when the value is 2. But, values that are too high make the algorithm converge to a tree that groups species sequentially. The algorithm also sensitive to this parameter. When the values is zero , solution will be not optimal because the important of heuristic values unused [7].

The genetic distance between an ancestor and two descendent species are controlled by parameter η and it has values between 0.1 and 0.9. For this paper, we used 0.3 and 0.5 .0.3 means the genetic distances between ascendant and descendant is 0.3 and with the other descendant 0.7 0.5 means the distance between ascendant and their descendant are the same and we observed that the best result will be produced when the values is 0.5.

The pheromone trail evaporation is controlled by the parameter ρ, which is influenced by the number of ants (κ) and the number of cycles. Experimentally, we observed that the higher values will give higher value to pheromone trail and this parameter is insensitive to this algorithm, but the value is still important to reach optimal solution.

We observed that the first pheromone will be better if the values is small but not zero but this parameter is in sensitive to this algorithm.

We observed that score tree fro reptile dataset is **0.650212069.** The tree states that *molesink kakapao*, *arenaria grturtle*, *caiman epturtle*, and *haematapus iguana*, have short genetic distance (*sister taxa*). Whereas *morepork* has genetic distance that closer to *bifurcating* and alligator is on the outgroup tree.

Furthermore, Laurasitheria dataset gives tree score of **1.956873924.** The tree has 8 sister-taxa, 3 bifurcating and 3 paraphyly. For the bird datasets, the value is **2.19170688**. It has 7 sister-taxa, 6 bifurcating, and 5 paraphyly. For Mammals dataset, the value is **1.46832902** with 7 sister-taxa, 3 bifurcating, 3 polyphyly, 1 paraphyly, and 1 outgroup.

We observed from 30 trials have given consistent result. It's means that it gives the optimal values. We also could determine that ACO algorithm could give optimal solution for this case because the search could balance the exploration and exploitation factor. Exploitation means the best solution space is using by the next search. eksploration means that the algorithm could reach another solution space that never touch.

5. Conclusion

1. ACO algorithm could give optimal solution for reconstruction of phylogenetic tree.
2. The best parameter for this case is:

Myu	Alfa	Beta	Rho	Phero awal	Iterasi	Jumant
0.5	1	2.0	>0	>0	100	60

3. ACO algorithm that using heuristic distance and pheromone trail importance could reach optimal solution faster.

References

[1] Allan, Wilson. 2003. Overview of Phylogenetic methods and applications. *http://www.genetics.org/cgi/search?qbe=genetics;155/2/765&journalcode=genetics&minscore=5000*. Downloaded on Dec 27th 2008

[2] Dorigo, Marco dan Colorni, Alberto dan Maniezzo, Vittorio. 1996 The Ant System: Optimization By A Colony Of Cooperating Agents. *http://iridia.ulb.ac.be/pub/mdorigo/journals/IJ.10-SMC96.pdf*. Downloaded on January 04th 2009

[3] Dorigo, Marco dan Stuzle Thomas. 2000. The Ant Colony Optimization Metaheuristics: Algorithms, Applications and Advances. *http://iridia.ulb.ac.be/pub/mdorigo/journals/IJ.10-ACO00.pdf*. Downloaded on January 4th 2009.

[4] Hariyanto, Bambang Ir, MT. 2007. "*Esensi-esensi Bahasa Pemrograman Java*". Informatika Bandung.

[5] Hartati, Sri, et al., 2006.*Pemrograman GUI Swing Java dengan Netbeans* 5. Penerbit Andi.

[6] Holland, Barbara. 2006. *Distance Based Methods for estimating phylogenetic trees. http://elesys.fsaintek.unair.ac.id/admin/makalah/Makalah0607-93.pdf.* Downloaded on January 18th 2009.

[7] Mauricio Perretto,Heitor Silverio Lopes. 2005. *Reconstruction of phylogenetic trees using the ant colony optimization paradigm*. *http://www.funpecrp.com.br/gmr/year2005/vol3-4//pdf/wob09.pdf* _downloaded on Dec 27 2008

[8] Porter, Sandra. 2007. *Phylogenetic trees*. *http://www.biomedcentral.com/content/pdf/1471-2105-7-S4-S1.pdf.* Downloaded on January 15th 2009

[9] Priyanto, Aruna Anggayasti. 2008. Tugas Akhir – *Penerapan Algoritma Ant Colony Optimization (ACO) pada Optimasi Penjadwalan Sumber Daya Proyek*. Bandung: Institut Teknologi Telkom.

[10] Shin Ando and Hitoshi Iba.2006.*Ant Algorithm for Construction of Evolutionary Tree*. *http://www.iba.t.utokyo.ac.jp/papers/2002/andoWCCI2002Ant.pdf*. Downloaded on January 15th 2009

[11] Soenartono, Adisoemarto.2004. *Penentuan Hierarki Dalam Klasifikasi Plasma Nutfah dan Pengaturan Namanya*. *http://www.papua.go.id/bkpbapedalda/penentuan_hierarki_dalam_klasifi.htm.* Downloaded on January 18th 2009

[12] Suyanto, ST, MSc. 2007. Artificial Intelligence Searching, Reasoning, Planning And Learning. Informatika Bandung

[13] Suyanto, ST, MSc. 2008. Evolutionary Computation. Informatika Bandung

[14] UCI.1985. *Download dataset*. *http://archive.ics.uci.edu/ml/datasets/Molecular+Biology+(Protein+Secondary+Structte*. Downloaded on Dec 27th 2009.

CORPORATE INFORMATION FACTORY PLANNING IN TELECOMMUNICATION OPERATOR

Kusuma Ayu Laksitowening

Informatics Faculty, Institut Teknologi Telkom
kal@ittelkom.ac.id

Abstract

Begins with the growth of players in telecommuncation business, rapid development of technology, new services trend, increasing market demand, and high potention market in Indonesia, making the business competition in this sector became highly competitive. To face this condition, the enterprise needs a creative and dynamic movement to survive the competition. But the enterprise's creative and dynamic movement only is not enough, it also needs strategic decisions that match the market condition. Critical decisions depend on the availability of the strategic information that match in the enterprise. The needs of strategic information are more related on how to build a healthy organization and the ability in surviving the competition. This can be achieved by implementing CIF. Corporate Information Factory (CIF) is a paradigm that specificaly pointed to provide a vital strategic information to the enterprise. CIF provides information quickly and in a format that greatly enhances the decision making process.

Key words: Strategic information, Corporate Information Factory, Telecommunication.

1. Background

Rapid development of telecommunication technology, deregulation, and globalization combine to pressure telecommunications operators and their ability to respond to these changes. In Indonesian Telecommunications industry, deregulation has brought a very exciting competition. Competition has created alternative product offerings as well as price sensitivity. Meanwhile revenues and margins in the core voice telephony business are in systemic decline.

The telecommunication operators need right tools to support the decisions making. Making better decisions faster can make the difference between surviving and thriving in an increasingly competitive market. This can be achieved by implementing Corporate Information Factory (CIF). CIF allows enterprise to exploit the potential of detailed information previously locked in legacy systems or summarized in distributed data marts and hence inaccessible to the business user.

2. Corporate Information Factory

2.1 Information Ecosystems and CIF

The economies gained three decades by automating manual business processes are no longer enough to gain a competitive advantages in today's marketplace. To compete, businesses need to build a new set of capabilities that deliver business intelligense and business management solutions[4].

Today, companies start to implement IT to survive in the competition. IT division is being bombarded with a growing number of targeted information architecture, technologies, methodologies, and tools, especially in data processing and information administration. But implementing many data processing tools in the same time can be confusing and not efficient.

An information ecosystem is needed to orchestrate the use of various information technologies and constructs, and to foster communication and cooperative exchange of work, data, process, and knowledge as part of a symbolic relationship. An information ecosystem is a system with different components, each serving a community directly while working in concert with other components to produce a cohesive, balanced information environment. Corporate information factory (CIF) is the physical embodiment of the notion of an information ecosystem. The CIF is at the same time generic in its structure and is unique to each company and organization as it is shaped by culture, politics, economics, and technology [4].

The Corporate Information Factory (CIF) is a logical architecture whose purpose is to deliver business intelligence and business management capabilities driven by data provided from business operations [3].

2.1 The Components

CIF is built from some componets. The different components of the CIF create a foundation for information delivery and decision-making activities that can occur anywhere in the CIF. Many of these activities are in the form of decision-support systems(DSS) that provide the end user with easy-to-use, intuitively simple tools to distill information from data [4]. The key components of the CIF are shown in Figure 1.

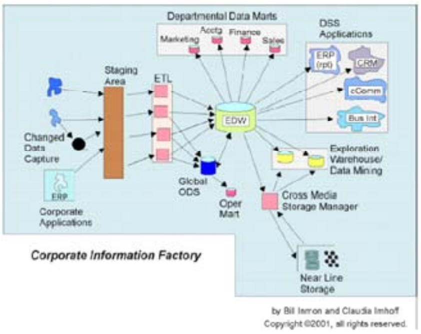


Figure 1. The CIF Structure.

3. Strategic Planning in CIF Development

The first step in building the corporate information factory is to understand what competencies drive the business. Business competencies represent areas of proficiency needed to support the business processes as they will exist in tomorrow's business landscape. These competencies are usually composed of a combination of people, processes and system.

The next step is to align the different componets of the CIF to these core competencies. So the company can see how these competencies are being positioned to support such business processes. This overlay of competencies to business processes represents the company's business vision.

However, the development of the CIF should be driven by strategic actions aimed at tactical business needs crucial to the survival of the business ecosystems that it support. With each strategic action, the CIF evolves –in form and function- to deliver incremental value to the business.

4. Development Phase

4.1 Company profile Identification

This phase includes the identification of business profile and company strategy. The purposes of these activities are to get a holistic view of the business.

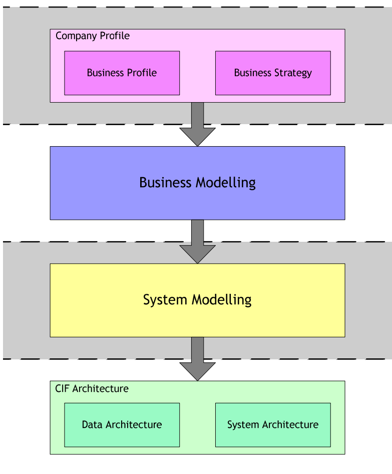


Figure 2. Development Phase

4.2 Business Modelling

In this phases, identification of the value chain and business processes are conducted to provide a comprehensive description about all business activities in the company. It is important to see the opportunity where CIF can support the critical activities by providing data or knowledge.

Business Area

The typical telecommunication company needs CIF to cover three business areas:

- CRM & *Marketing*,
- *Service & Network*, and
- *Billing*.

Knowledge Needed

CIF on telecommunication is running to provide comprehensive information that can be accessed to support the following analysis and reporting processes:

- Campaign Analysis
- Campaign Report
- Sales Analysis
- Sales Report
- Customer Acquisition Analysis
- Churn Analysis
- Churn Report
- Customer Behaviour Analysis
- Customer Complaint Analysis
- Service Order Processing Analysis
- Service Order Processing Report
- Customer Profitability Analysis
- Product Profitability Analysis
- Customer Life-time-value Analysis
- Network Usage Analysis
- Service Usage Analysis
- Revenue Analysis
- Product Report
- Network Traffic & Performance Report

4.3 System Modelling

The processes of this phase are conducted to get clear description about the background and the motivation of CIF development. In this phase, the functionalities of the CIF are identified.

CIF Development Goal

- To provide comprehensive data that could be accessed by the whole enterprise and could be used to create report, analyze and support decision making.
- To track enterprise business activities.

Process Chart

The Process Chart is the primary diagram to model process flow logic within System. One or more Process Charts are used to describe the process flow in the organization.

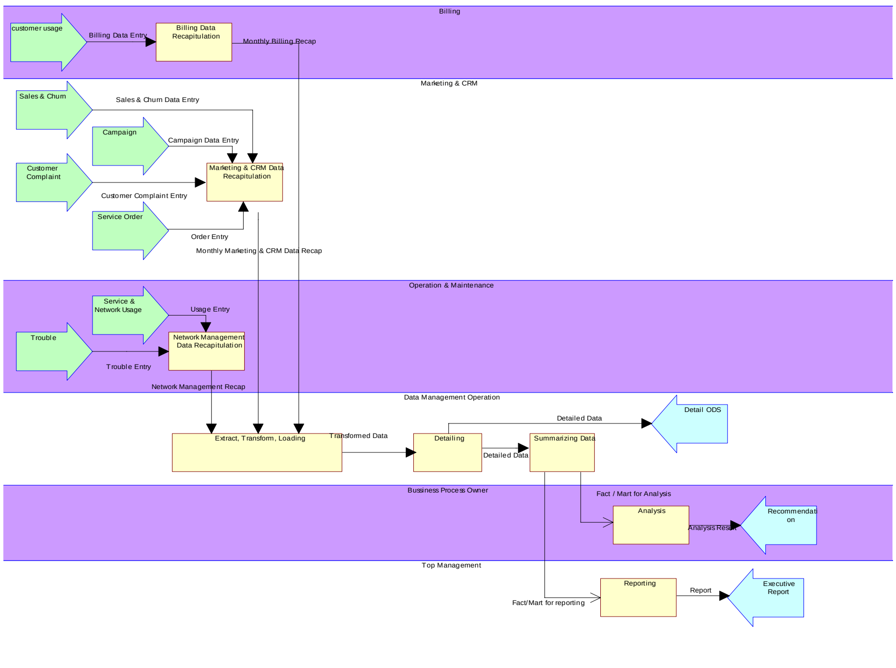


Figure 3. Example of Process Chart in Telecommunication Operator

4.4 CIF Architecture Design

This phase includes data and system architecture design. These processes are conducted to get detailed description about data being managed by the CIF and information about the environment where the CIF would be placed.

Data Mart and Data Warehouse Design

Data marts are analytical data stores designed to focus on specific business functions for a specific community within an organization. Data marts are often derived from subsets of data in a data warehouse, though in the *bottom-up* data warehouse design methodology the data warehouse is created from the union of organizational data marts. The data marts and data warehouse design are derived from the needs of business area covered by the CIF and the requirement analysis results

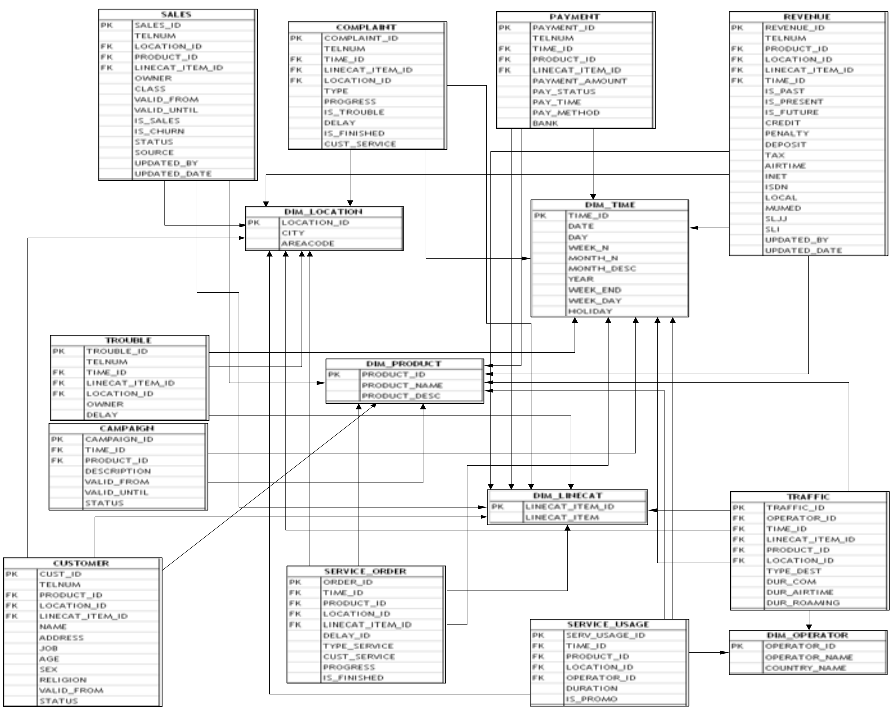


Figure 4. Example of Data Warehaouse Design in Telecommunication Operator

5. Implementation Issue and Success Keys

CIF implementation faces some important obstacles. Therefore, the enterprise should concern to some issues, such as:

- data inconsistency,
- decentralized data,
- design and architecture issue, and
- leadership issue.

Success of CIF design implementation depends on:

- Participation and support from management in CIF implementation
- Availability of human resources who is competent in IT and CIF
- Socialization
- Applied technology
- Leadership on implementation

6. Conclusion

- Strategic planning on CIF development would result integrated and effective system, and resource efficiency.
- Data warehouse design activity became easier since it was made based on requirement analysis result.

References

[1] Imhoff, Claudia (1999), *The Corporate Information Factory*, http://www.dmreview.com/ accessed on June 9 2006.

[2] Inmon, W.H., Imhoff, C., Sousa, R. (1998), *Corporate Information Factory*, Wiley Computer Publishing.

[3] Inmon, W.H., Terdeman, R H., (2001), *The Evolution of The Corporate Information Factory*, www.billinmon.com accessed on June 7 2006.

[4] Inmon, W.H., (2005), *Building the Data warehouse*, Fourth Edition,Wiley.

[5] Matisson, R., (1997), *Data Warehousing and Data Mining for Telecommunication*, Artech House.

[6] Ponniah, P., (2001), *Data Warehousing Fundamentals: A Comprehensive Guide for IT Professionals*, John Wiley & Sons, Inc.

[7] ________________(2003), *Telecommunications Data warehouse*, IBM

DESIGN AND IMPLEMENTATION OF COMMUNICATION BETWEEN VIRTUAL WORLD AND REAL WORLD BASED ON CROQUET IN 3D VIRTUAL WORLD

Leonardy Setyawan[1], Jin Chun Kim[2], Byung Ho Kim[3], Ihan Martoyo[4], Grit Schuster[5]

[1,4]Dept. of Electrical Engineering, Pelita Harapan University
[2,3]Dept. of Computer Engineering, Kyungsung University
[5]Free Interaction Designer (MA), Berlin, Germany
[1]leonardy_setyawan@uph.edu, [2]jckim@ks.ac.kr, [3]bhkim@ks.ac.kr, [4]ihan_martoyo@uph.edu, [5]mail@gritschuster.de

Abstract

This work is a continuation of the previous works to dismantle the barrier between virtual world and real world for the ease of communication. This system performs a two way (*from real world to virtual world and from virtual world to real world*) communication instead of the general one way (*from real world to virtual world*) communication. Croquet, a software system, is used to manage the 3D virtual world and the communication between the virtual world and microcontrollers. Movements in the real world will be captured and showed in 3D virtual world (Croquet virtual world). On the other hand, any action in the Croquet virtual world will cause the data from virtual world to be sent into the real world. By doing this, the user will be able to notice physically that their avatars are having an encounter with another avatar. This work uses electronic devices, microcontrollers, and Croquet. The electronic devices that are being used are flex sensors for taking the input data. A vibration motor and LED (light emitting diode) are used as the output parts. Microcontrollers are used to manage the communication between the real world and the computer. The uses of Croquet, microcontrollers, and electronic devices have made the dismantling of the barrier between the 3D virtual world and the real world possible. This research has increased the attractiveness of a virtual world and can still be improved by adding different kinds of output parts to the system.

1. Introduction

In the last couple of years, there have been attempting to dismantle the barrier between the virtual world and the real world for the ease of communication. With the advanced improvement in computer technology and other supporting technology, improvement in the field of communication has become widely known and used.

One of the recent improvements that have attracted a lot of researchers is Croquet. Croquet and other electronic devices can be used in improving the communication system. By using croquet and other electronic devices, dismantling this barrier is no longer unreachable. Croquet is a program used to implement the 3D virtual world and to collaborate with other users in 3D virtual world. To establish this system, supporting electronic devices are needed so that the 3D virtual world and the real world can conveniently communicate with one another. These electronic devices are used as sensors and as actuators.

Sensors will be used to collect data from the real world and actuators will be used to respond to the real world. This paper will focus on active vitality which means that all changes caused by the user will be monitored directly both in the virtual world and in the real world. Vitality is something that can only be recognized by constant changes, similar to some phenomena from nature like moving. By implementing active vitality, all personal expressions of the user such as speaking, pointing, grabbing, and gesturing can be performed in the virtual world. This research is also focused on the effect of things happening in the real world cause by things happened in the virtual world. By doing this the user in the real world can notice the effect of something done by the avatar in the virtual world.

The communication between the 3D virtual world and the real world will be managed by a computer and a microcontroller using serial communication. All of these subsystems will then be combined into a system that can be used for dismantling the barrier between the virtual world and the real world and thus making communication more convenience.

The inputs in this system are generated by a flex sensor, a computer keyboard, and a computer mouse. The flex sensor is one of the inputs which are attached to the puppet’s lip. The movement of this flex sensor on the puppet’s lip will generate an electronic signal. This electronic signal will be processed by the microcontroller and will then be forwarded to the computer via serial communication.

The computer within Croquet application will display the 3D virtual world. In the 3D virtual world, there will be an avatar which represents the user. The avatar has a similar shape with the puppet in the real world. This avatar is able to communicate with the other avatars in the 3D virtual world. When the two avatars make a physical contact in the 3D virtual world, the Croquet application will send serial data to the microcontroller and this data will be forwarded to the actuator as an output of this system. The actuator in this system is a vibration motor.

When the contact takes place, the vibration motor will be activated. This vibration will be felt by the user. So, the user will be able to notice that they have an encounter with another user and may want to respond toward it.

2. Related Work

Several theories are needed to realize the idea of dismantling the barrier between the 3D virtual world and the real world. These theories include squeak language programming, Croquet as the platform of this system, and the microcontroller as the connector between the computer and the real world.

- *Squeak*

Squeak is a modern, open source, fully-featured implementation of the Smalltalk programming language. Squeak is highly portable – even its virtual machine is written entirely in Smalltalk, making it easy to debug, analyze, and change. Squeak is the vehicle for a wide range of innovative projects from multimedia applications and educational platforms to commercial web development worlds. There are three parts that are needed to run squeak, consisting of four files which can be seen in Figure 1.

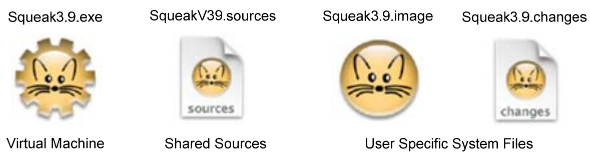


Figure 1. The Squeak Files

The virtual machine (VM) is the only part of the system that is different for each operating system and processor. Pre-compiled virtual machines are available for all the major computing worlds. The source file contains the source code for all the parts of Squeak that do not change very frequently. In Figure 1 it is called SqueakV39.sources. Note that the file SqueakV39.sources is only for versions 3.9 and later of Squeak. The current system image is a snapshot of a running Squeak system, frozen in time. It consists of two files: an .image file, which contains the state of all of the objects in the system (including classes and methods, since they are objects too), and a .changes file, which contains a log of all of the changes to the source code of the system. Squeak, like most modern Smalltalk dialects, adopts a syntax very close to that of Smalltalk-80. The syntax is designed so that the program text can be read aloud as though it were a kind of Pidgin English.

- *Croquet*

Croquet is a new approach for developing and delivering collaborative interactive media applications. Every part of the system is designed around enabling real-time, identical interactions between groups of users. The architecture of Croquet actually makes it quite easy to develop collaborative applications without having to spend a lot of effort and expertise in understanding how replicated applications work. There are a number of simple patterns and rules to remember, but otherwise, it is quite simple to quickly develop very powerful systems.

TeaTime and Islands are the basis for Croquet's replicated computation and synchronization. They are designed to support multi-user applications that can be scaled to massive numbers of concurrently interacting users in a shared virtual space. Croquet’s treatment of distributed computation assumes a truly large scale distributed computing platform, consisting of heterogeneous computing devices distributed throughout a planet-scale communications network. Applications are expected to span machines and involve many users. In contrast to the more traditional architectures, Croquet incorporates replication of computation (both objects and activity), and the idea of active shared subspaces in its basic interpreter model. More traditional distributed systems replicate data, but try very hard not to replicate computation. It is often easier and more efficient to send the request for the computation to the data, rather than the other way around. Consequently, Croquet is defined so that the replication of computations is just as easy as replication of data alone.

- *Microcontroller*

A microcontroller (also microcontroller unit, MCU or μC) is a small computer on a single integrated circuit consisting of a relatively simple CPU combined with support functions such as a crystal oscillator, timers, watchdog, serial and analog, I/O etc. Program memory in the form of NOR flash or OTP ROM is also often included on the chip, as well as a typically small read/write memory.

Microcontrollers are designed for small applications. Thus, in contrast to the microprocessors used in personal computers and other high-performance applications, simplicity is emphasized. Some microcontrollers may operate at the clock frequencies as low as 32KHz, as this is adequate for many typical applications, enabling low power consumption (milliwatts or microwatts). They will generally have the ability to retain functionality while waiting for an event such as a button pressed or other interrupts; power consumption while sleeping (CPU clock and most peripherals off) may be just nanowatts, making many of them well suited for long lasting battery applications. Microcontrollers are used in automatically controlled products and devices, such as automobile engine control systems, remote controls, office machines, appliances, power tools, and toys. By reducing the size and cost compared to a design that uses a separate microprocessor, memory, and input/output devices, microcontrollers make it economical to digitally control even more devices and processes.

3. Design and Implementation

There are many works that have been done before and are related to this work. Mostly are only designed to take the input from the real world and present it in the 3D virtual world. This work is the extension of those works because an output part is added in the system so that the communication takes place in both ways, from the real world to the 3D virtual world and from the 3D virtual world to the real world. In order to do this, the 3D virtual world has to be synchronized with the real world. An efficient and fast way of collecting data from the real world and visualizing them in the 3D virtual world are needed in a way to sync virtual world with reality. Activities that the user does will be captured by sensors and forwarded by the microcontroller to the 3D virtual world. Data from the 3D virtual world will then be forwarded to the microcontroller. Finally, the microcontroller will forward the data to actuators so that the actual response can be noticed directly by the user.

A. Information Detection

Information detection from the real world is an important part to enter the system. Nowadays, tiny little sensor products combined with small microcontrollers are able to change the life and the way to deal with things around us. For example, a temperature sensor that can detect the heat at a certain room, when combined with a microcontroller will make a subsystem that can be used for communicating with the computer. So, the sensor is the gate for data to be processed in a system. These data consist many kind of information that is needed to make a full system. Sensors are electronic devices, which are often used for receiving data about the conditions of the world; such as temperature, lighting condition, pressure, and distance; and converting these analogous parameters into electrical signals.

Electrical signals are the necessary input for a microcontroller which will process it and produce digital information as an output. The types of sensors differ regarding to their complexity and their kind of output. Starting with a switch producing either 0V or 5V, the complexity for measuring temperature or distance is even higher. The more complicated sensors are color sensors or GPS sensors which provide data like longitude and latitude for positioning information. Figure 2 shows several types of sensors: digital temperature sensor (upper left), RGB color sensor (upper right), flex sensor (lower left), and ultrasonic range sensor (lower right).

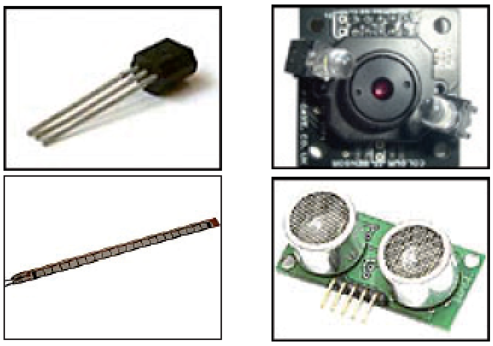


Figure 2. Several Types of Sensors

In this work, the real world is designed to interact within three dimensional applications. However, interacting within three dimensional applications by using traditional means like a keyboard or a mouse is limited. By using microcontroller and sensors, more specific and dedicated input and manipulation devices can be created. By using a flex sensor, simple motions of the user‘s hand can be interpreted and used as input. For example, by detecting the distance between the user‘s thumb and index finger, a measuring gesture of the user can be detected and used to manipulate sizes in virtual space.

In this system the movement of the flex sensor will be interpreted as the movement of the avatar’s lip in croquet. The circuit diagram of the flex sensor can be seen in Figure 3. The movement of the avatar in the 3D virtual world is controlled by a keyboard or a mouse.

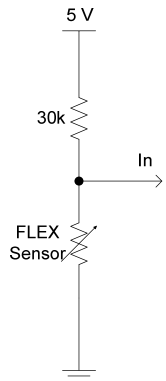


Figure 3. Circuit Diagram of FLEX Sensor

B. Actuator

Actuator is an electronic device which is used to show the response of the system if an input is sent to the system. There are many kinds of actuators that can be used to show different responses. The most common electronic devices used as an actuator are LED (Light Emitting Diode), speaker, motor, and vibration motor. These common electronic devices are usually used in everyday life. Figure 4 shows several types of actuators (From the left is speaker, vibration motor, and LED).

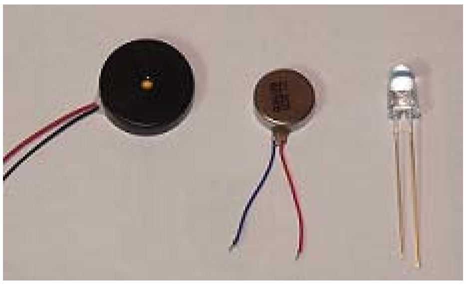


Figure 4. Several Types of Actuators

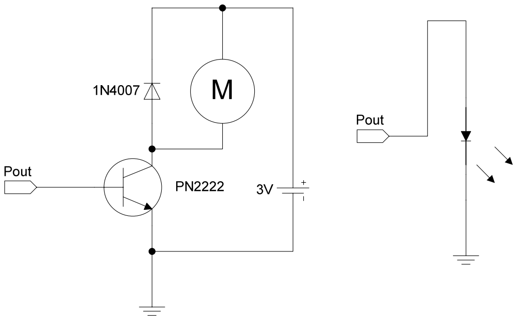


Figure 5. Vibration Motor and LED Circuit

This system uses three kinds of actuators: LED, vibration motor, and speaker. LED and vibration are controlled by the microcontroller while the speaker is controlled by the computer because the computer's speaker is being used. Actuators are used so that the user can feel the significant situation which happens in the 3D virtual world. By doing this, the user will feel no more separation between the 3D virtual world and the real world.

C. Microcontroller

The communication between the computer and the real world is managed by a microcontroller. The microcontroller will receive data from the sensors and forward it to the computer. Serial communication is used for the communication between the microcontroller and the computer. The microcontroller also controls the actuator. The microcontroller will receive data from the computer and forward it to the actuator. The microcontroller used in this system is Arduino microcontroller as seen in Figure 6. Arduino is a controller that can be plugged into a computer and thus enabling the computers to sense the physical world around you and give more output possibilities than your desktop computer. It is an open-source physical computing platform based on a simple microcontroller board, and a development environment for writing software for the board. Arduino can be used to develop interactive objects, taking inputs from a variety of switches or sensors, and controlling a variety of lights, motors, and other physical outputs. Arduino projects can be stand-alone projects, or communicate with other software running on your computer (e.g. Flash, Processing, MaxMSP.) The boards can be assembled by hand or purchased preassembled; the open-source IDE can be downloaded for free. The Arduino programming language is an implementation of Wiring, a similar physical computing platform, which is based on the Processing multimedia programming world.

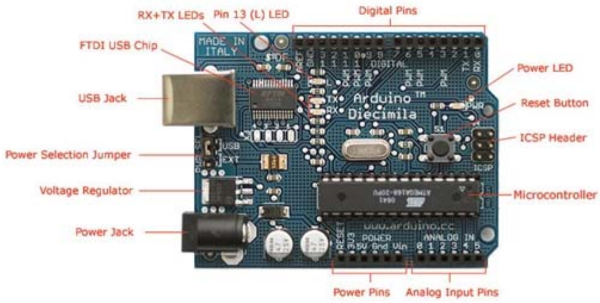


Figure 6. Details of Arduino Microcontroller

Figure 7 shows the circuit diagram of the system. The system is controlled by the Arduino microcontroller and the computer itself. The Arduino microcontroller controls the electronic devices and the communication to the computer while the computer, by using Croquet, controls the 3D virtual world.

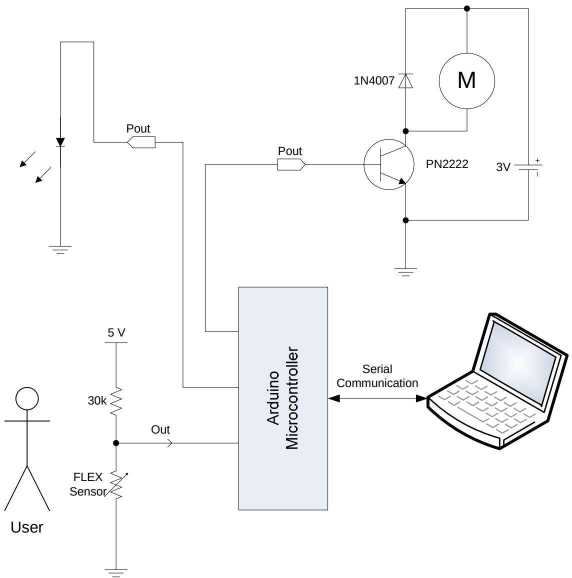


Figure 7. Circuit Diagram of the System

D. Croquet

Using Croquet as the technological basis was decided for the 3D virtual collaboration world. Croquet, as seen in Figure 8 is an open source software platform for creating collaborative multi-user online applications with its network architecture designed for collaboration, communication, resource sharing, and synchronous computation among multiple users. It runs exactly identical on several operating systems like Macintosh, Windows, or Linux. By using Croquet, one can author and change the virtual 3D worlds, visualizations, and simulations.

Croquet runs inside Squeak, a software development environment for live software construction using the object-oriented programming language Smalltalk. Morphic, its direct-manipulation User Interface construction kit, works with graphical objects called Morphs and replaces the original Model View Controller graphics toolkit of Smalltalk-80. Squeak implements a simple interface to the serial ports of the underlying platform. By using the class #SerialPort, a port specified by its number and specific parameters like the baud rate can be opened. Communication through this interface works in both directions: A string or a byte array from the port can be read or sent out. The microcontroller continuously sends data to and receives data from the computer through the serial connection.

The Arduino microcontroller is used for reading, processing, sending analog sensor data to, and receiving data from the computer. In Squeak, these data can be read from an opened serial port. At present, Croquet uses Morphic to communicate with Squeak.

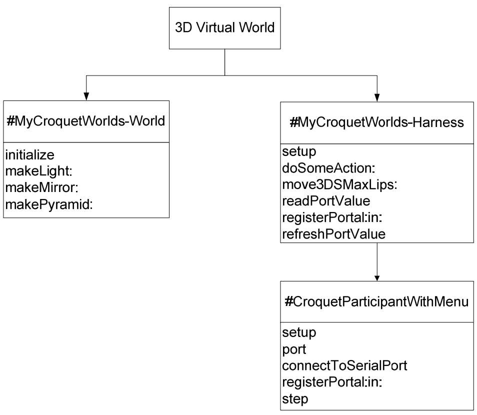


Figure 8. Class Diagram of the Application

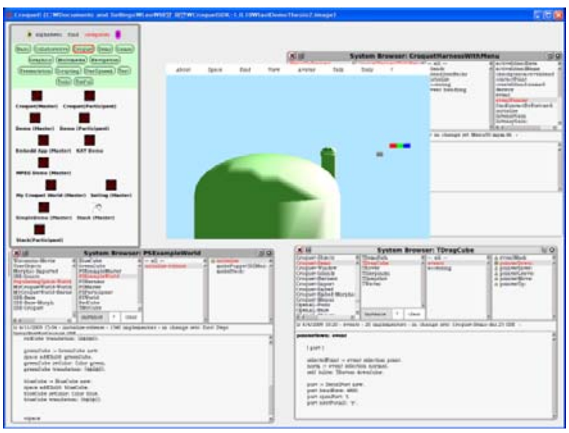


Figure 9. Croquet Application of the System

The #CroquetParticipantWithMenu runs as a single Morphic window within Squeak and serves the view into the 3D Croquet world. The actual state of the virtual 3D world seen through this window is rendered periodically by the #CroquetHarnessWithMenu that is the primary interface between the user and the Croquet world. All of the UI (User Interface) and presentation layers around Croquet go directly through the Croquet Harness which is responsible for underlying infrastructure and event management. The block diagram of the system can be seen in Figure 9 below.

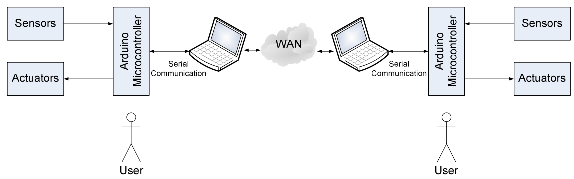


Figure 10. Block Diagram of the Entire System

E. *System Procedure*

The input data is taken from the flex sensor, the keyboard key, and the mouse. The flex sensor is attached in the puppet's lip. When the puppet's lip is moved the flex sensor will be bent. The degree of the flex sensor bending in the puppet's lip will affect the avatar lip in the 3D virtual world. The movement of avatar in the 3D virtual world is controlled by the keyboard key or the mouse. The program is expected to improve the vitality of the application. The idea is when the avatar meets another avatar in the 3D virtual world; the system will send serial data to Arduino. Arduino will process the serial data and make the vibration motor attached in the puppet vibrating. By this, the user can feel that the other avatar is near them and may want to interact with them. Besides that, there are three cubes (red, green, and blue) displayed in Croquet. Each of these cubes represents the LED in the real world. When the avatar clicks on the cube, the LED status in the real world will be toggled. The flow chart of the system procedure can be seen in Figure 10.

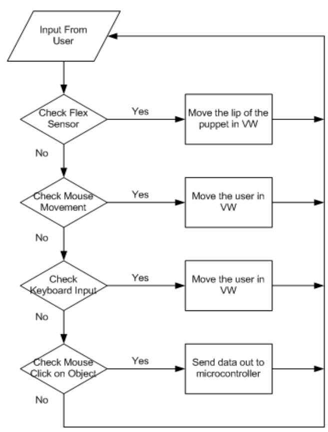


Figure 11. Flow Chart of the System

4. Result and Analysis

The puppet represents the user in the 3D virtual world. The user controls the movement of the puppet's mouth by using the flex sensor attached in the real puppet. The joy and comfort of direct manipulation can be used to animate virtual puppets. The motion between the thumb and the index finger is translated into the opening and closing of the puppet‘s mouth, thus creating a familiar way of interaction and animation. Direct manipulation is an excellent example for active vitality. The glove puppet is a joyful input device. The sensor is bent to open or close the puppet‘s mouth. The values of the sensor control a virtual puppet's lip. The movement of the puppet in the 3D virtual world based on Croquet is handled by the right mouse click and the keyboard. In the 3D virtual world, there are cubes. These cubes are used to send data to the real world. If a puppet click on one of this cube, it will send serial data to the Arduino microcontroller and then the Arduino microcontroller will process that data.

The result will be shown at the actuator connected to the Arduino microcontroller. So, anything significant that happened in the 3D virtual world would also happen in the real world. Thus, the user will be able to really feel it happening just like in the real world. Objects can be added in the 3D virtual world to make the barrier between the 3D virtual world and the real world becomes less significant.

By doing this, the user can feel the 3D virtual world as a real world. The system procedure has been realized but not completely. The "touching between the users" part is still cannot be realized. To resolve this problem, a grey cube is used as the substitution. So, when the user clicks on this cube, the vibration motor will be vibrating. The rest of the described procedure has been applied in this system. Dismantling the barrier of the communication between the virtual world and the real world is possible. All things that happened in the virtual world can also be felt in the real world. Croquet as the tool for collaboration can fully help the users to experience virtual world in the real world. By combining it with the Arduino microcontroller, this system has succeeded in realizing that idea.

This system has several advantages for the real life, especially in the working place, educational program, and 3D game. By using this system, the communication barrier that has been mentioned earlier can be removed so that users can almost feel a real communication with another user. In applying this system, human gestures, gazes, and expressions can be shown entirely for the convenience of the user. The main application that is mostly suitable and desirable with this system can be seen in 3 major parts that will be explained briefly.

Meeting is an important time where all the members need to gather around in one room to discuss an important issue or to evaluate any development periodically. This is the most advantageous part of the application of this system. A general meeting would spend a lot of time not only on the presentation or debate but also on the transportation, especially in a country where traffic jam is a matter of everyday's life. It is not only more interesting because the members are in different places, but it will also save a lot of time and energy, not to mention the advantages caused by the less carbon dioxide (CO_2) produced by car or other means of transportation. The application of this system will accommodate a tele-meeting which capable of arranging a virtual meeting using internet as a medium. The member will be able to feel as if all members are in the same place at the same time even though they can be miles apart.

E-learning is also an interesting proposal that can take advantages from this application. This will allow the system to be applied as a real education system in the real world. The problem that some educator or educational field has these days is mostly cause by the lack of facilities to hold a class, and the willingness to provide a good education to those who by some problem cannot attend a regular school. With this new technology, teachers and students can do the teaching and learning process as usual without actually being at the same place or meeting with each other. Of course this program cannot replace education systems because socializing is also a part of education. This solution is more directed to those who do not have the opportunity to enter a regular school whether it is because of a constant change in place to live, a security reason, sickness or disability, or simply a lack of time.

The 3D game is a new exciting development that has captured a lot of teenager's heart nowadays. The application on this field will mostly be useful to those who are interested in making a whole new level of gaming which can be played like the real life. For this application, this system must be improved to the next level of application so that the 3D display can be used as an input device also.

5. Conclusion

By using Arduino microcontrollers, Croquet, and electronic devices, dismantling the barrier between the 3D virtual world and the real world is possible to be realized. The vitality within a virtual world can be improved by adding different kinds of output parts to the system. To make this system works, the virtual world has to be synchronized with the reality. With syncing, the system does not have to copy the visual appearance of reality, but more so to create properties and behaviors corresponding to the phenomena of the real world. The effect of the response made by the 3D virtual world can cause the user to feel no difference between the 3D virtual world and the real world because all the significant things happened in the 3D virtual world also happen in the real world.

For the next research, this system can be enhanced by changing and adding the type of the output. By doing this, the users can feel the virtual world even more in the real world. One specific example is the implementation of a 3D display, which will allows the user to experience the virtual world directly in the real world not only virtually but also physically. Furthermore, as the technology develops, the users will no longer be capable to distinguish between the virtual world and the real world.

Acknowledgement

This paper was supported by Department of Computer Engineering in Kyungsung University with a tremendous help from Grit Schuster.

References

[1] Black, A. P., et. al., *Squeak by Example*, 2008.

[2] Smith, D. A., A. Raab, D. Reed, and A. Kay, *Croquet Programming,* Qwac, Inc., 2006.

[3] Smith, D. A., A. Kay, A. Raab, and D. P. Reed, *"Croquet – A Collaboration System Architecture"*, In: *Proceedings of the First International Conference on Creating, Connecting and Collaborating through Computing (c5)*, 2003.

[4] Smith, D. A., A. Raab, D. P. Reed, and A. Kay, *"Croquet: A Menagerie of New User Interfaces"*, In: *Proceedings of the Second International Conference on Creating, Connecting and Collaborating through Computing*, 2004.

[5] Smith, D. A., A. Raab, D. P. Reed, and A. Kay, *Croquet User Manual v.0.0.1*. Web site: http://www.opencroquet.org, 2004.

[6] Smith, D. A., A. Raab, D. Reed, and A. Kay, *"Croquet: The User Manual, Draft Revision 0.1"*. Viewpoint Research Institute, 2002.

[7] O'Sullivan, D. and T. Igoe, *Physical Computing: Sensing and Controlling the Physical World with Computers*. Thomson Course Technology PTR, 2004.

[8] Schuster, G., C Strothotte, and C. Zwick, *"Syncing Croquet with the Real World"*. In: *Proceedings of the Fifth International Conference on Creating, Connecting, and Collaborating through Computing*, 2007.

[9] Tanaka, K., *"Tea-Time Museum: Croquet As a Browsing and Searching Environment for Digital Libraries". Proceeding of the Conference on Creating, Connecting and Collaborating through Computing*, pp.1-6, 2003.

[10] Borders, M., *"In All Possible Worlds: A radical new technology has arrived and the world will never be the same"*, 2004.

[11] Schmidt, M., *"Tutorial for Building a Simple Croquet World"*, 2008.

[12] Sari, R. F. and Muliawan, *"Interactive object and collision detection algorithm implementation on a virtual museum based on croquet"*. In: *Proceedings of the 4th International Conference on Innovations in Information Technology*, pages 685–689, 2007.

[13] Kumar, S., et.al, *"Second Life and the New Generation of Virtual Worlds"*. IEEE, 2008.

[14] http://www.arduino.cc accessed in May, 2008.

[15] [15] http://xaverse.blogspot.com accessed in June, 2008.

[16] http://www.opencroquet.org/index.php/Tutorials accessed in Feb, 2008.

[17] http://gl.ict.usc.edu/Research/3DTeleconferencing/ accessed in March, 2008.

[18] http://wiki.squeak.org/squeak/792 accessed in June, 2008.

[19] http://www.dmu.com/crb/crb0.html accessed in June, 2008.

[20] http://www.dmu.com/crb2/crbb0.html accessed in June, 2008.

[21] http://www.dmu.com/croquet2/ncr0.html accessed in June, 2008.

[22] http://www.dmu.com/croquet/cr0.html accessed in June, 2008.

PERFORMANCE TEST OF STATISTICAL TRANSLATION MACHINE AT TRANSLATING ENGLISH TO INDONESIAN LANGUAGE

Mohamad Ramdhani[1], Arry Akhmad Arman[2]

[1]Faculty of electrical and communication, Institut Teknologi Telkom
[2]STEI Institute Technology Bandung
[1]mra@stttelkom.ac.id, [2]aa_arman@rocketmail.com

Abstract

Statistical translation machine works by finding the biggest probability or possibility of a sentence to be translated in target language when the source language sentence is given. This translation machine is composed using modelling of statistic from Bayes theorem, that consist of language model, translation model and decoder. Based on testing, there are a some results. First, that statistical machine translation can remember all sentences which have been trained previously, second, it can be trained to repeat in innovating its own knowledge, third, it can translate sentences which have not yet been introduced by using sample sentence examples which have been studied , and four, as the data are more intensively trained, the machine translation will progressively give even better performance.

Keywords: Statistical machine translation, probability, Bayes theorem

1. Introduction

Language can be differentiated into natural language and artificial language. One of application form in processing natural language that more developed is natural translation machine. It can translate sentences from one language to another natural language.

Statistical translation machine is a translation technique based on statics by using instruction model from one group of sentences (*corpus*) as trained data.

The research of Translation machine which use statics hasn't developed yet, especially in Indonesia, whereas in developed country,the research has been developed well.Furthermore, it has many used, although the usage is still in research scope .This research has purpose to test the succes and accuracy of statistical translation machine from english to Indonesian language.

2. Statistical Translation Machine

The basic idea from statistical translation machine is finding the biggest possibility of translation sentences of target language ,when it's given the sentences of source language[1]. Statistical translation machine has some other advantageous, some of them are: 1) flexibility in overcome the ambiguity problem, 2) can be related to the idiom which is available in trained data, 3) Do not need a special linguist and needs minimal effort only, 4) can be made into a new prototype system quickly, according to trained data which is available and furthermore it minimize the budget that should be paid.

The weakness in relationing to syntax language, where statistical translation machine can not decide what the relation of sytax in a language clearly.

A *string* source language 'f' can be translated to string target language 'e' by many methods. In statistical translation machine , can be viewed that every *string* target language 'e' is a various possibility from *string* source language'f'. When it's given *string* source language 'f', the task of machine is finding the *string* target language 'e' from all possibility in string target language for the best *P (e|f)* based on Bayes theorem written[1]:

$$\hat{e} = \arg\max_{e} P(e)P(f \mid e) \qquad (1)$$

Brown and his friends called the equation above with fundamental equation from statistical translation machine or known as *"noisy channel model"*. Where the target language is signed by 'e' and in source language by 'f'. From equation (1), we get a model that P(e) is a model language, P(f|e) as model translation , and **e^** as *decoder*. Here the system block of statistical translation machine :

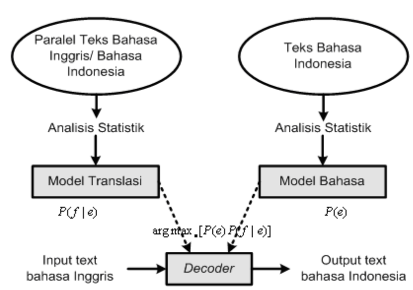


Figure 1.System Block of Statistical Translation Machine.

2.1 Language Model

Language model in statistical translation machine shows is a sentence is a target of a good target language sentence. Language model makes sure that the output from target language has run well by guide method in choosing word and its order. Model language is a probability value from target language which provides a set of word. It is potential to be translated.

Model language uses n-gram standard of language model from SRI International Speech Technology and Research. The probability of model language is written in equation below:

$$P(w_1, w_2, ..., w_{n-1}, w_n) = P(w_1^{\,n}) = \prod_{k=1}^{n} P(w_k \mid w_1^{\,n-1}) \quad (2)$$

w is the word order and n is the number of words from those sentences.

2.2 Phrase based Translation Model

This translation model based on *word alignment* concept . *Word alignment* is a mapping between source word and target word in a set paralel sentence. The probability of *word alignment* is accounted by using formula [8]:

$$P(f \mid e) = \sum_a P(a, f \mid e) = \frac{\varepsilon}{(l+1)^m} \sum_{a1=0}^{l} ... \sum_{am=0}^{l} \prod_{j=1}^{m} P_t(f_j \mid e_{aj})$$

$$P(f \mid e) = \prod_{j=1}^{m} \sum_{i=0}^{l} P_t(f_j \mid e_i) \quad (3)$$

Where l is a number of word in source language and m is a number of word in target language.

Phrase based translation is built from *phrase alignment* where this *alignment* is developed from previous *word alignment* . The fundamental idea from phrase based translation is segmentation from source sentence ,and after that's translated to each phrase. Finally, a sentence of target language is formed [3]. Probability of phrase alignment[4]:

$$\phi(\bar{e}, \bar{f}) = \frac{count\ (\bar{e}, \bar{f})}{\sum_f count\ (\bar{e}, \bar{f})} \quad (4)$$

2.3 Decoder

The task of decoder is using knowledge and its function density for finding the best order from source transformation sentence to the target sentence or an effective and efficient searching technique for finding the target sentence. For searching unlimited and unorganized searching space, heouristic is needed. Heouristic which is used in decoder is algorithm *beam search.*

3. The Experiment

3.1 Test of Remembrance Capability

This test has goal to see the capability of statistical translation machine to remember the sentences that have ever been trained previously. The test is done by giving sentences as input which have been trained previously.

Based on the result of test, can be taken a conclusion that statistical translation machine is be able to remember all sentences which have ever been trained previously. The value which is gotten by using software , help the mesurement . The BLUE value is about 10000, means that all structured sentences and its words are right. Since the training process until the output result needs time about 6 seconds.

Table 1. Table of A Pair Trained Sentences for Testing Remembrance Capability.

Kalimat Latih (Bahasa Inggris)	Kalimat Latih (Bahasa Inggris)
I need a book	Saya membaca sebuah buku
You go to school	Anda pergi ke sekolah
She buy an umbrella	Dia membeli sebuah payung
He must work hard	Dia harus bekerja keras
They have a car	Mereka mempunyai sebuah mobil
We will go there	Kami akan pergi ke sana
I see a bridge	Saya melihat sebuah jembatan
You give some money	Anda memberi sedikit uang
She has two children	Dia mempunyai dua anak
They can play footbal	Mereka dapat bermain sepakbola

3.2 Test of Repeating Study Ability

This test has purpose to see the abillity of statistical translation machine to repeat the study. Repeating study is usually needed when there are some new sentences which have never been known previously. The test is done by giving the sentences as input which are same with previous trained sentences , where the trained sentences consist of five last sentences and five new sentences.

According to the result of test, it shows that statistical translation machine can be trained again for innovating the own knowledge . The value is gotten by using software ,to help measurement and gotten BLUE value is about 10000. Since the training process till the output result needs time during 10 seconds.

3.3 Test of Concluding The New Knowledge

This test has goal to see the capability of statistical translation machine to conclude the new knowledge, as the result, untrained sentences can be translated. The test is done by giving sentences as input which have ever been trained previously into two part, they are a) all previous trined words , and b) all untrained words.

Based on the result the test, it's viewed that statistical translation machine can translate the sentences which have never been introduced when it refers to the examples of sentence which have ever been studied . To have a test of untrained words, seems that statistical translation machine isn't be able to translate the previous untrained words.

Table 2. The Result Test Table of All Trained Words.

Kalimat Latih (Bahasa Inggris)	Kalimat Latih (Bahasa Inggris)
They need a book*	*Mereka* membaca sebuah buku
You go to *bridge**	Anda pergi ke *jembatan*
She *give* an umbrella*	Dia *memberi* sebuah payung
He *will* work hard*	Dia *akan* bekerja keras
They have *an umbrella**	Mereka mempunyai *sebuah payung*
You *have* some money*	Anda *memiliki* sedikit uang
We see *a book**	*Kami* melihat *sebuah buku*
Children give some money*	*Anak* memberi sedikit uang
You have two children*	*Anda mempunyai* dua anak
He can play football*	*Dia* dapat bermain sepakbola

3.4 The Test of Impact The Number of Words in An Ordered Translation Sentence.

The test is done by two methods , they are a) by the number of same word for each sentence, and b) by the number of different word for each sentence.

From the test, it can be taken a point that the number of words in each sentences do not affect the ordered translation from source language to target language . It takes 7 seconds from the training process till the output result.

3.5 Test of Impact Crossed Translation at The Number of Same Words.

In doing test, it has purpose to see how the impact when source sentence which its translation word crossed the words in target sentence is available. Based on that reason, it is done a strategy test in two methods, they are :a)when there is a sentence with two crossed words, and b) when there is two sentences which each sentence has two crossed translation words. The testing is done by untrained sentence that consist of one crossed translation word.

Table 3. The Result Table When Crossed Sentence is Available

Kalimat Latih (Bahasa Inggris)	Kalimat Latih (Bahasa Inggris)
I need *your book*	Saya membaca *buku kamu*
You go to school	Anda pergi ke sekolah
She buy an umbrella	Dia membeli sebuah payung
He must work hard	Dia harus bekerja keras
They have *your* car*	Mereka mempunyai *buku* mobil
We will go there	Kami akan pergi ke sana
I see a bridge	Saya melihat sebuah jembatan
You give some money	Anda memberi sedikit uang
She has two children	Dia mempunyai dua anak
They can play footbal	Mereka dapat bermain sepakbola

The value which helped by software measurement ,we get BLEU = 0.92 66. During the training process until the output result, it needs 6 seconds. According to the test, there is a mistake in word " your car" becomes "buku mobil" which should be " mobil kamu", this is caused the trained sentence has ordered *alignment* principle. At previous trained sentence, word " your" is translated " kamu" and " book" is translated " buku".

Based on repeating test, we can get a conclusion that by adding the trained data, we can get better translation performance , or more words are trained, the translation of statistical translation machine would be better.During training process until the output result needs 7 seconds.

For having a test of two sentences in crossed translation ,shows that more trained sentences which has crossed translation means more trained sentence that should be repaired or incresing the performance of the result.

During training process until the output result needs 7 seconds. To test two crossed translation sentences, it's showed more trained sentences which has crossed translation means more trained sentence that should be repaired or incresing the result performance .During training process until the output result needs 7 seconds.

3.6 Test of Impact The Number of Different Translation Word.

Test is done by giving input sentence which have never been trained with all previous trained word. The value BLEU = 0.8306 is gotten by helping from software in measurement . During the training process until output result needs 5 seconds.By adding trained sentence " I am " and" Student", the result of translation in this test can be repaired. The conclusion for translation which has more number of different words needs additional sentence at trained sentence to get the good one.

3.7 Test of Time Translation Process.

This test is done to know time translation process since training process until the output from statistical translation machine.

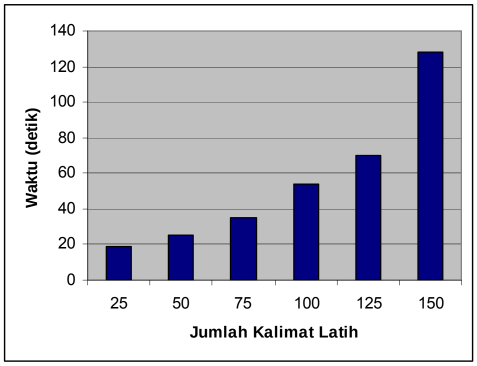


Figure 2.Graph of Translation Processing Time.

The scenario of test is done by dividing trained sentence into 25, 50,75, 100,125, 150 sentences by each trained sentence is 25 same sentences.

From the result of test, we get graph of time translation process to increasing a number of trained sentence, when there are more trained sentences, it would increase time processing too.

4 Conclusion and Suggestion

4.1 Conclusion

Based on the test results which has been done in this research, we can take some points.They are :

1. Statistical translation machine is be able to remember all sentences which have been trained previously.
2. Statistical translation machine can be trained repeatly for renewing or innovating own knowledge.
3. Statistical translation machine can translate the sentences which have never been introduced yet, according to examples of sentences that have ever been studied.
4. Statistical translation machine can not translate the untrained sentences, so that all words in source language and target language should have been trained previously.
5. To order the translation from source language to target language, the number of words for each application does not affect the translation.
6. In crossing translation, it needs additional trained sentences to increase the translation performance in a result.
7. Adding additional trained sentences is needed to get the good performance of translation in sentences which have different number of word translation.
8. To get better translation result, it needs additional data in trained sentences.
9. Time translation processing is longer as the additional trained sentence.

4.2 Suggestion

As the follow up from this research ,some suggestions which can be done are:

1. Decide maximum iteration value in translation model for shorting the data training process.
2. Decide value *back off* in language model to increaseperformance of target language.

References

[1] Brown, P.F., Pietra, S.A.D., Pietra, V.J.D., Mercer, R.L. (1993), *The mathematics of statistical machine translation: parameter estimation* Available : http://acl.ldc.upenn.edu/J/J93/J93-2003.pdf [23 November 2006].

[2] Knight, Kevin., (1999), *A Statistical MT Tutorial Workbook*, Available: http://www.isi.edu/naturallanguage/mt/wkbk.rtf, [25 November 2006].

[3] Koehn, Philip. (2006), *Statistical machine translation: the basic, the novel, and the speculative*.Available:http://www.iccs.inf.ed.ac.uk/~pkoehn/publications/tutorial2006.pdf [11 Desember 2006].

[4] Koehn, P., F. J. Och, and D. Marcu. (2003), *Statistical phrase-based translation*.Available : http://www.iccs.inf.ed.ac.uk/~pkoehn/publications/phrase2003.pdf[11 Desember 2006].

[5] Kuhn, Jonas. (2006), *Statistical machine translation and other parallel corpus-based techniques : training of statistical models*. Avaialble:http://www.ling.uni-postdam.de/~jonask/teaching/StatMUe-WS06-07/3-StatMT-Training.pdf[24 Desember 2006].

[6] Mosteller, F., Rourke, R. E. K., Thomas, G. B. (1988), *Peluang Dengan Statistika Terapannya*, Penerbit ITB Bandung, 201 – 208.

[7] Papineni, K., Roukos, S., Ward, T., and Zhu, W.J. (2001), *BLEU : a method for automatic evaluation of machine translation*. Available: http://www1.cs.columbia.edu/nlp/sgd/bleu.pdf , [22 Januari 2007].

[8] Zens, R., Ney, H., (2004), *Improvements in phrase-based statistical machine translation*, http://tt2.atosorigin.es/docs/Zens+Ney_ImprovementsInPhraseBasedSMT_HLT_NAACL_2004.pdf, 5 Januari 2007.

DEVELOPMENT OF SOCIAL SECURITY CARD CASE ONMONITORING MALNUTRITION PATIENTS IN NTT AREA

Muhammad Ashar[1], Didik Dwi Prasetya[2], Slamet Wibawano[3]

[1,2,3]Department of Electrical Engineering State University of Malang
[1]ashar@um.ac.id, [2]didik@um.ac.id, [3]slametwi@yahoo.com

Abstract

Malnutrition will affect the high maternal mortality rate, infants, and toddlers, and low life expectancy. In addition, the impact of malnutrition seen that the low participation of schools, low education, and slow economic growth. At least some of the fundamental problems identified in this study is the weak effectiveness of prevention of malnutrition, Lack of cooperation and coordination between the various parties involved, the biased approach-curative emergency short-term without a strategic program-oriented long-term preventive, ignored potential and resources of local communities, and systems administration services manually. The purpose of research is to improve the ability to analyze the condition of malnutrition in NTT area. To be able to set priorities to address the problems, able to choose appropriate interventions and cost effective according to local needs and able to build functioning institutions and nutrition through the use of health information technology and smartcard ID, are able to monitor and evaluate food development by mapping a source of food and nutritional standards of food processing in accordance with the existing potential in the food NTT.

Keywords: Malnutrition, Social Security Card, Smart ID

1. Introduction

Improved nutritional status of communities by emphasizing efforts NTT preventive and promotive health and nutrition services to the poor in order to reduce the number of people with malnutrition. Through the utilization of information technology in empowering potential of community, institutional, and government that ensures the integration policies, programs and activities between central and local government health sector especially malnutrition. While nutrition has not been made systematically in a model of administrative services with operational oversight and support plan non-governmental funding sources such as fund local governmental organizations, nationally and internationally.

So that specific goals in building the administrative model of SSC-based community services. (Social Security Card) at the NTT pursued as a model of technology capable health-care system in:

- Increasing the ability of communities and individuals in health and nutrition services equitable, affordable and quality and cost-effective.
- Improve access to family health and nutrition information to form a conscious behavior of food and nutrition and healthy living based smartcard ID.
- Increasing the supply of food security through increased participation of food producers and implementation of effective supervision and efficient through geographic information systems.
- Improve coordination of food handling and nutrition problems in an integrated, both from posyandu, health centers, hospitals and health departments online support transparency of funds nationally and internationally through a reporting system based on SSA (Social Security Administration)

2. Design and Modelling System

Draft national strategy of research that will be made a priority in the development and improvement of malnutrition through the administrative system services include:

Administration section model of efficient and effective services

Modeling administration malnutrition community service aims to how the system is capable of arranging food programs and nutrition are coordinated with cross-sectoral cooperation in the form of digital information.

Appropriate technology to access system services malnutrition

The use of technology appropriate access system will make direct community empowerment through the provision of ID number for access by the severely malnourished, administrative personnel, and professional personnel in order to maximize the effectiveness of nutrition programs nutrition improvement.

Increased monitoring of malnutrition improvement

With the database administration design services through an information system using smart

ID, easier monitoring system with administrative staff utilizing both national and local levels in conducting a recapitulation of malnutrition data for the purposes of nutrition research and development with reliable data and information.

Distribution of fast and accurate service

Distribution of food, medicines and vaccines at the service centers such as Posyandu, health center or clinic coordinated through the control of the center to the regions rapidly through the information system is equipped with a system of decision-making connected with health management Information System a regional and local agriculture.

Transparency source of food and funds

Food sources of information and funds are arranged in the short term and long term elements involving the community, institutional and government in fostering for the allocation of funds for food and nutrition in an integrated society through social security (social security).

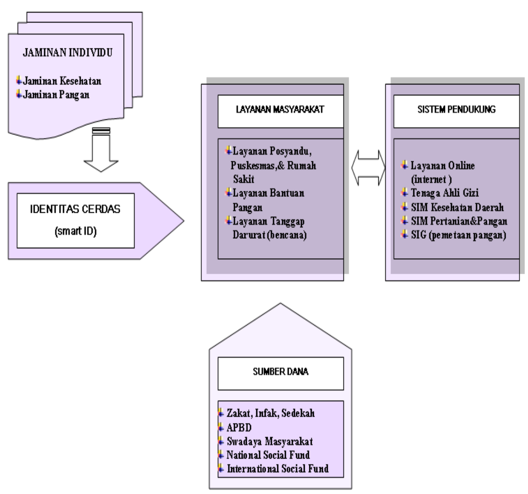


Figure 1. Public Administration Service Model using the SSC

Classification model of administration with SSC / SSA include:

1. Smart identity (Smart ID) cards are used as individual guarantees contained in the severely malnourished in accessing community services and support systems with the goal of all service activities that have been given to severely malnourished can be recorded and displayed in an online database. This is the monitoring by the community, institutions, and governments with the support systems through information systems.

2. Individual guarantees can be obtained through the charging system of an online form with the provision of ID number of patients with malnutrition with legalized by the mother of the patient (parents).

3. Funding Sources for obtaining access to services by the severely malnourished and transparency made the funding allocation is designed in accordance with the short-term needs and long term.

4. Supporting systems covering how an online system that was built to support information technology services through the health management information systems areas, agriculture & food, as well as geographic information system that can visualize the mapping of food resources, distribution of food and medicine to the needs of communities across the district, the location service malnutrition, and the fluctuations of malnutrition district area respectively.

Security Social Administration

Fig 2. SSA information systems online

SSA process are :

- A number of administrators working in a team that supports the creation of nutrition services activities
- Implementation of IT-based electronic system is in the field of infrastructure development to improve nutrition and food
- The design of the new system with the modification and integration of electronic medical, online transactions using a computer network automation systems.
- The design and implementation of inter-test improvement of health services
- Participation intensively from the government in implementing SSA
- Support from the community, non-governmental agencies.

Methodology

1.Observation and documentation
2.Design Software & Hardware
3 Data Validation System

Administrative Servicesi Model with SSC

Modeling administrative services web based on SMART ID

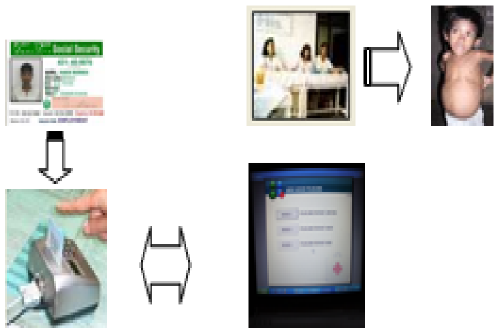


Figure 3. Process Flow Malnutrition Health

3. Results of Research

Development of NTT malnutrition Health Information System with Web-based monitoring included: Form SSA , SSC cards, Integrated Service Post, the monitoring of malnutrition with patient documentation.

Examples of form-based monitoring malnutrition Health Information System as follows:

Figure 3. SSA online Form

Figure 4. SSC online form

Figure 5. alnutrition Evidence News

Figure 6. Patient Malnutrition

4. **Summary**

Smartcard ID application of the Malnutrition status was made in order to gain access layanana online. The system works based on the patient's visit is scheduled on the card to health. Each patient got smartcard as SSC. Monitoring will be conducted by qualified nutrition in a number of Posyandu. The system can display the data documentation in the form:
Condition (nutritional status), Early rehabilitation, Nutrition Food intake patterns, Nutrition Supplement intake patterns, Pattern of drug intake, Examination results (physically and Lab), Food Sources, Source of Funds

References

[1] Ecosoc Right, 2007, "Busung Lapar NTT Berakar dari Soal Struktural", Antara News, Jakarta)

[2] Nicole Ad, 2007, "New paradigm for public health" , health Journal California.

[3] Richard P Adler , 1994," Electronic Delevery of Social Security service, Contractor paper prepared for the off ice of technology Asssessment

[4] Torney, 2007," Social Security Numbers and Why Your Baby Needs One", USA

[5] Lou dan Y Aron,2003,, "Recommendations to the Social Security Administration on the Design of the Mental Health Treatment Study" , New Freedom Commission on Mental Health

ANALYSIS OF INDONESIAN NEWS DOCUMENT CLASSIFICATION USING CENTROID BASED CLASSIFIER METHOD

Neneng Siti Fatimah[1], Moch. Arif Bijaksana[2], Yuliant Sibaroni[3]

[1,2,3]Institut Teknologi Telkom

[1]neng_sifa@yahoo.com, [2]mab@ittelkom.ac.id, [3]yls@ittelkom.ac.id

Abstract

Today, amount of information, such as news articles, available on the web are growing fastly. Large number of information causes the user get into trouble in finding important information. One of *Data Mining* task, *Text Categorization*, which is the task of assigning documents to pre-specified classes (categories) of documents can be used as the solution to organizing news documents. One of t*ext Categorization* method is *Centroid Based Classifier*. *Centroid Based* method represent a document as a vector. This method create the *centroid vector* for each set of documents belonging to the same class. That centroid vector will be used as model to classify documents using *cosinus similarity*. In this paper, performance of *Centroid Based Classifier* method is compared with performance of k-NN and Naïve Bayes. Accuracy and f- measure are parameters that used to compare the performance of those methods. Beside of that, we also analyze outlier to improve the accuracy of *Centroid Based classifier*. The experiments show that *Centroid Based Classifier* give much better performance than k-NN and Naïve Bayes, and outlier handling can improve the accuracy of *Centroid Based Classifier Method*.

Keywords : ***text categorization, centroid based classifier, centroid vector, akurasi, f- measure.***

1. Introduction

Nowadays the growth of existing information within the web continues to increase, as well as news articles published on the web. A large amount of information cause the difficulty in finding information or news that we want, then much information will be useless. Text Categorization, is a method to classify documents into a particular topic or category can be used as a solution to organize news documents.

To organize Indonesian news documents, we use the Centroid Based Classifier method where this is one of simple classification methods whom has better performance than other algorithms such as Naive Bayes, k-NN, and C4.5 [2]. The basic idea of Centroid Based Classifier is a document represented as a vector and the method classify a document based on cosine similarity.

One of problems that appear is outlier. When a document has far distance from it's centroid class cause the performance of classification accuracy be lower. This document referred as an outlier. We have to handle these outlier if we want a good model.

Problem Identification

The problem identification is:

1. How to classify a news document using *Centroid Based Classifier* method.
2. How the performance of this Centroid Based Classifier.
3. How to handle outlier to improve performance of the centroid Based classifier

2. Basic theory

2.1 *Text Categorization*

Text Categorization (TC) isa method to classify documents (news) into one or more predefined classes. Categorization text main goal is to find a model to categorize natural language text [2]. Then The model will be used to determine the class of a new document

Given a group of training documents D = (d_1, ..., d_i) with known class C = (c_1, ..., c_i) and the new document q, the text Categorization will determine the class of the new document. Several methods such as Text Categorization: k-Nearest Neighbor, Naive Bayes, Support Vector Machine, Decision Tree, Neural Networks, Boosting. there are three stages in text categorization: preprocessing, training phase and testing phase

2.1.1 Preprocessing

At this stage, we are preparing document to be ready for next stage. There are 3 sub-stages in preprocessing i.e. [1]:

- Feature extraction
- Feature Selection
- Document Representation

2.1.2 Training Phase

In the training phase, we develops classifier model. This classifier model is called Vector Space Model. Vector space model is performed from training data.

2.1.3 Testing Phase

In the testing phase, we will classify a document using vector Space Model. Each document in testing data will be compared with all centroid class using the cosine similarity. Accuracy of this model is measured to evaluate if this model suitable to classify new data.

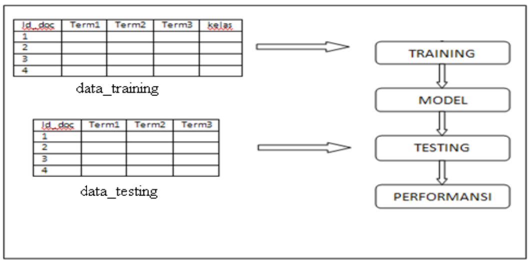


Figure 1. Text Categorization Process

2.2 Information Retrieval

In Information retrieval(IR), a document described by a set of word representations called term index. Not all terms have the same interest in describing the contents of the document. To determine the most important term in describing the contents of the document, weighting process is applied to each index term. There are three IR models used to represent the documents [5]. These models are: Boolean Model, Probabilistic Model, and the Vector Model. In this research, we use Vector Space Model

2.3 Vector Space Model

Vector Space Model is a model represents documents in a set of Index Terms which weighted based on the their interests[2].

, is weight of term j-th at document i-th.

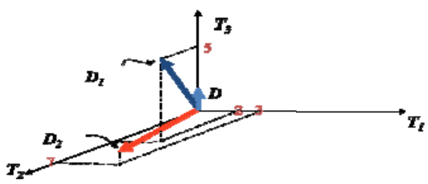


Figure 1. Document Representation

The similarities between the document vector d and q can be calculated using cosine of the angle between the two vectors i.e:

2.4 Weighting

- TF (Term Frequency)

f (d, t) is the occurrence frequency of term t-th in document d-th

- IDF (Inverse Document Frequency)

N is the total number of documents
df (t) the number of documents containing term t.

- TFIDF

TFIDF is terms that often appear in the document but rarely appears on the set of documents provides a high weight value.

2.5 Centroid Based Classifier

Centroid-based classifier is a simple linear algorithm with strong performance but not yet widely studied and analyzed [7].

2.5.1 How Centroid Based Classifier Work

In centroid-based algorithm, the document is represented using the vector space model, where the vector elements associated with the weight term. In this model, each document d is considered as vectors in term space. Each document is represented as TFIDF vectors.

Value of is normalized so that .

Given a group of documents S in training data and the vector representations of them. For each class consisting of a group of documents, we define the centroid vector. The centroid values represent the prototype vector centroid set of documents in a class. Centroid vector is defined as .

In the testing stage, each document in testing data is compared with the centroid vector of each class. The similarity between testing data and the centroid vector is calculated using the cosine similarity:

Due to each document in testing data have been normalized, the cosine similarity becomes:

Document i-th will be grouped in class j-th when it have biggest cosine value.

(2.5)

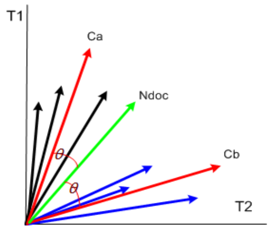


Figure 3. Ilustration of The Cosine Similarity

The calculation complexity of training stage is in proportion to amount of training data and amount of term in training data, meanwhile the calculation complexity of testing stage is in proportion to amount of class and amount of term in testing data

2.5.2 Calculation of The Centroid Value

There some method to calculate centroid value in training stage, i.e.: [2]:

- Centroid Rocchio
  Previously, Rocchio method is used to optimize query from related *feedback* IR, then this method can be adapted for *text categorization*[13].
  There are two parameter in this method, i.e. : β and γ. The recommended value for β and γ [13] are β =16 and γ = 4.

$$\vec{c}_j = \beta \cdot \sum_{\vec{d}_i \in D_{Cj}} \vec{d}_i - \gamma \cdot \sum_{\vec{d}_i \notin D_{Cj}} \vec{d}_i$$

- Centroid Average
  In this method, each class centroid c is represented as average of all document vector in class c

$$\vec{c}_j = \frac{1}{|D_{Cj}|} \cdot \sum_{\vec{d}_i \in D_{Cj}} \vec{d}_i$$

- Centroid Sum
  In this method, each class centroid c is represented as sum of all document vector in class c..

$$\vec{c}_j = \sum_{\vec{d}_i \in D_{Cj}} \vec{d}_i$$

2.6 Outlier Handling

One of factor cause decrease of classifier method is outlier. In this method, outlier is a document with far distance (specified constraint) from it's centroid.

To handle the outlier, we clean training data previously from outlier before calculate centroid. We remove document which detected as outlier. The outlier threshold is determined with some criteria. We use two kind of threshold i.e. fix threshold and relative threshold.

The fix threshold use fix constant threshold for every document whereas relative threshold use different threshold for each class. Statistics that used in relative threshold is standard deviation (σ).

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (D_{cj} - \bar{X})^2}{n}}$$

$\bar{X}$ = centroid of class i-th
n = amount of document training.
We can adapt $(D_{cj} - \bar{x})$ as $\cos(\vec{d}_i, \vec{c}_j)$, then value of threshold ε :

$$\varepsilon = \sqrt{\frac{\sum_{i=1}^{n} (\cos(\vec{d}_i, \vec{c}_j))^2}{n}}$$

In relative threshold, data k-th is said outlier at class j-th when

$$d_k < \bar{x} - k\varepsilon.$$

Meanwhile, in fix threshold data k-th is said outlier at class j-th when

$$\cos(\vec{d}_i, \vec{c}_j) < \varepsilon$$

2.6 Measurement Parameter

Mostly, evaluation of text classification method using accuracy and *F-measure* parameters i.e:

$$\text{accuracy} = \frac{\sum \text{true classification}}{\sum \text{classification document}}$$

$$precision\ A = \frac{\sum \text{true classification document with result A}}{\sum \text{classification document with result A}}$$

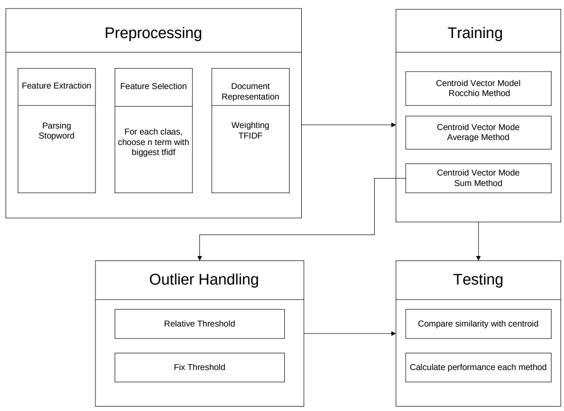
$$recall\ A = \frac{\sum \text{true classification document with actual class A}}{\sum \text{classification document with actual class A}}$$

$$Fmeasure = \frac{2(recall)(precision)}{recall + precision}$$

The accuracy and F-measure are measures between 0 and 1. The classifier model is model with highest accuracy and F-measure value.

3. Process Diagram

The diagram below shows steps in our research.



Picture 4. Process Diagram

The first process is preprocessing which contains Feature extraction,Feature Selection and

Document Representation process. The next process is training process where classifier methods is performed. There two kind of data to perform classifier model are data with outlier and data without outlier. After we get classifier model, we determine performance of classifier model in testing process.

4. Experiment

4.1 Dataset

Data set that used in this research is news data from website such Okezone. com, Jawapos. com, Kompas. com, Pikiran Rakyat. com, Liputan6. com, Antara. co. id. This Data was taken between March until August 2008. *Preprocessing* is taken to data set so that data set has matrix form (.csv).

4.2 Experiment Scenario

There are some scenario in experiment i.e.:

1. Analysis and Calculation Centroid of all Methods
   In this experiment, We calculate centroid with all method and analysis the result.
2. Accuracy Comparison between sum method and Rocchio Method.
   We divide data randomly in two part: training data (80%) and testing data(20%). The classifier model is developed from training data and accuracy is calculated from testing data. This experiment is repeated in five time. The accuracy of each method is average of all accuracy.
3. Performance comparison among Centroid Method with k-NN and Naïve Bayes method.
   We compare these methods for all data set and performance calculation k-NN and Naïve Bayes is obtained using Weka tool with k value = 5 (k-NN). Performance is measured from accuracy and f-measure value.
4. Outlier handling using relative threshold and fix threshold
   We compare performance of centroid method for data with outlier and data without outlier. In this experiment, we only use the best centroid method from early experiment. We make 5 experiments for each threshold with many different thresholds. At the last, we compare performance both of threshold handling methods.

4.3 Experiment Result

1 Analysis and Calculation Centroid of all Methods
From the analysis of centroid calculation that we have done, we get result that the average method and the sum method give same model, so we only use two centroid methods for next experiment i.e Rocchio method and Sum method.

2 Accuracy Comparison between sum method and Rocchio Method.

Table 1. Accuracy of Sum and Rocchio Methods

Data-set	Sum Method	Rocchio Method
okezone	**94.20%**	92.60%
jawapos	**98.30%**	97.87%
kompas	**94.62%**	93.76%
pikiranrakyat	90.00%	**90.50%**
liputan6	**96.00%**	93.00%
Antara	**90.94%**	87.19%

Sum Method has better accuracy than Rocchio Method

3 Performance comparison among Centroid Method with k-NN and Naïve Bayes method.

Based on early experiment, the sum method give better accuracy than Rocchio method so at the next experiment we use Sum method as the best centroid method.

Table 2. Accuracy of Sum, k-NN and Naïve B

Data-set	Centroid (Sum)	k-NN	NB
Okezone	**94.20%**	70.40%	92.80%
Jawapos	**98.30%**	67.23%	97.44%
Kompas	**94.62%**	65.37%	91.18%
Pikiran Rakyat	**90.00%**	67.00%	86.00%
liputan6	**96.00%**	71.00%	91.00%
Antara	**90.94%**	63.13%	87.76%

Table 3. F-Measure of Sum, k-NN and Naïve B

Dataset	Sum	k-NN	NB
Okezone	**94.22%**	71.16%	92.76%
Jawapos	**98.21%**	66.37%	97.21%
Kompas	**94.53%**	66.87%	91.40%
Pikiran Rakyat	**90.36%**	65.21%	86.02%
Liputan6	**95.63%**	70.49%	90.39%
Antara	**90.54%**	64.78%	87.84%

From table above, we see that Centroid Model has better performance than others method.

4. Outlier handling using relative threshold (RT) and fix threshold (FT)

The experiment shows that outlier handling can increase accuracy until 3% for both of threshold method but it is not valid for all experiment. The performance both methods is same relatively.

The fix threshold gives best accuracy for ε=0,15 meanwhile the relative threshold give the best accuracy for k=0,25.

5. Conclusion

1. Sum method and average method give same Classifier model. The difference between those model is only for model configuration time.

2. Sum centroid sum give better performance than Rocchio Centroid Method.
3. *Centroid Based Classifier* method give much better perfomnace in text classification than k-NN and Naïve Bayes Method.
4. Outlier handling in *Centroid Based Classifier* method can increase accuracy until 3%.
5. The fix threshold gives best accuracy for ε=0,15 meanwhile the relative threshold give the best accuracy for k=0,25..

6. Reference

[1] Briicher, Heide. Knolmayer, Gerhard. Mittermayer, Marc-Andre. Document Classification Methods for Organizing Explicit Knowledge. University of Bern. Switzerland.

[2] Cardoso-Cachopo, Ana. and Oliveira, Arlindo. 2006. *Empirical Evaluation of Centriod-based Model for Single-label Text Categorization*. Instituto Superior Tecnico. Portugal.

[3] Cardoso-Cachopo, Ana. and Oliveira, Arlindo. 2007. *Semi-Supervised Single Label Text Categorization Using Centroid Based Classifiers*. Portugal.

[4] Feldman, Ronen dan James Sanger. 2007. *The Text Mining Handbook.* New York: Cambridge University Press

[5] Chipaila. Kisesa. Amuel.2006. *Information Retrieval Usung Relevance Feedback (Rocchio Technique).*

[6] Feldman, Ronen dan James Sanger. 2007. *The Text Mining Handbook.* New York: Cambridge University Press.

[7] Han, Eui-Hong and Karypis, G. 2000. *Centroid – Based Document Classification : Analysis & Experimental Rresult*. Department of Computer Science Minessota University. Minneapolis.

[8] Han, Yong. Abraham, Ajith. And Shin, Kwangcheol. 2006. *Enhanced Centroid-Based Classification Technique by Filtering Outliers.*Chung –Ang University. Korea.

[9] Jiawei, Han. dan Micheline, Kamber. 2001. *Data Mining : Concepts and Techniques*", Intelligence Database Systems Research Lab, School of Computing Science, SImo Fraser University.

[10] Sebastiani, Fabrizio. *A Tutorial on Automated Text Categorisation.* Istituto di Elaborazione dell'Informazione Consiglio Nazionale delle Ricerche. Italy.

[11] Schutze, Hinrich. Raghavan, Prabhakar. Manning, Christopher. 2007. *An Introduction to Information Retrieval*. Cambridge University Press. England.

[12] Tokunaga, Takenobu. *and* Iwayama, Makoto. 1994. *Text Categorization based on Weighted Inverse Document Frequency*. Department of Computer Science. Tokyo Institute of Technology. Japan.

[13] Thorsten, Joachims.1997. A Probabilistic Analysis of the Rocchio Algorithm with tfidf for Text Categorization. Universitat Dortmund. Germany.

[14] Uden, Mark. *Rocchio : relevance Feedback in Learning Classification Algorithms*. Department of Computing Science. University of Nijmegen. Netherlands.

[15] WEKA Explorer User Guide for Version 3.5.6

SIMILARITY MEASUREMENT IN DIGITAL MUSIC FILE BASE ON CHROMA-BASED REPRESENTATION

Zulhendra[1], Fazmah Arif Yulianto[2], Tjokorda Agung Budi W[3]

[1,2,3]Faculty of Informatics Institut Teknologi Telkom
[1]**jp.reloaded@gmail.com**, [2]**faz@ittelkom.ac.id**, [3]**cok@ittelkom.ac.id**

Abstract

In the image and video file, there is a thumbnail to see the preview of the file, but in the audio or music file, we cannot get the preview of the music file. Therefore, we need a tool that can make the preview that represents the content of digital music file. In music, the part, which represents the whole music, is the chorus, because it is the core of the music and chorus is the most repeated part in music. Chroma-based representation is a representation of the human perception of pitch. In chroma-based representation, the pitch from music will be map into 12 pitch of chromatic scale. With chroma-based representation, we hope that we can detect the repeated part of the music or a part that have similarity with other part in music. Correlation can be use to detect the repeated pattern from the chroma representation of the music. The software will detect whether a part of the music is repeated or not based on the result of the correlation calculation from the two different part of music's chromas. If the correlation result could detect that there is a repeated part of the music based on the music's chroma, the conclusion is chroma can be use to search the similarity among parts of the music. The result of this research is chroma generally can be use to detect the repeated part of the music that can be consider as a preview of the music file. However, to make a sample from the music there is some more process to do.

Keywords*:* chroma-based representation, correlation, repeated pattern, chorus.

1. Introduction

Development of internet with high-speed bandwidth make the distribution of multimedia file became widely and easy. Multimedia files usually need a large space in storage, especially the high quality multimedia file. Therefore, the required storage medium for storing large multimedia files are increase, along with the need to access those files quickly.

In the picture file or video, we can get a preview or thumbnails of the file, but this feature is not available on the music file. The lack of thumbnails feature on music file causes difficulty to do searching for the music file that we like on the internet. Usually we must download the entire music file, listened the music file and sometimes it does not match with what we want to search.

In this research, we make a mechanism to make a pseudo-preview from the music file by finding the chorus part of the music file. Chorus can be considering, as a core of the music by assumption that chorus is the most repeated part in music.

2. Chroma-Representation

Music consist of a sounds that organize in time domain, we call this sound as tone. The tone usually determined by the frequency of the sound, the higher frequencies creates a high-level tone. The differences between the tone frequencies called the interval. To determine the level of the tone can use the equation 1.

$$p = 69 + 12 * \log_2\left(\frac{f}{440}\right) \qquad (1)$$

The symbol p is the number of tone levels and f is the *fundamental frequency* tone. *Fundamental frequency* is the frequency of the tone or in a melody is the lowest frequency of these melodies.

Perceptual structure from the music can be representing using two-dimensional vector; this vector has s a helix form. Vertical dimension of the vector represent the level of the tone while the angular dimension represent the chroma. Roger N Shepard first proposed this, a psychologist from Stanford University in 1964 in his book entitled *"Circularity in Judgments of relative pitch"*. The detail can found in the figures1.

Based on figure 1 we can see that the angular dimension contain the sequence of chroma in chromatic scale and the vertical dimension is a frequency values of the tone.

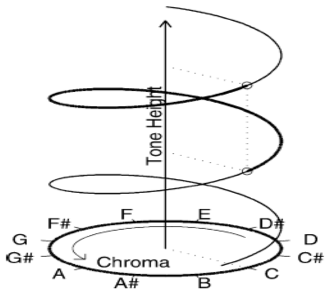


Figure 1Perceptual Structure of the music

In this research chroma classification from the music signal can be determined using equation 2.

$$c = \left\lfloor 12 * \log_2 \left(\frac{f}{261{,}63} \right) \right\rfloor \quad (2)$$

From the equation 2, the grade levels of chroma depend on the frequency of the musical signal. The value 261.63 became a divider factor to determine the classification level of chroma octave. Based of tone frequency table the value 261.63 is the frequency of tone C in the fourth octave. So that the resulting chroma classes in our reserach will start from C - C # - D - D # - E - F - F # - G - G # - A - A # - B.

2.1 Chromagram

Chromagram is a graph depicting the chroma interpretation of the song that in form of chroma. Chromagram that used in this research will used color scale to visualize the minimum and maximum value of the chroma. Chroma colors in each chromagram described in the chroma signal strength so that the stronger the signal means that the higher the value chromanya, and the higher the value chromanya the higher the color scale. Figure 2 illustrate a chromagram.

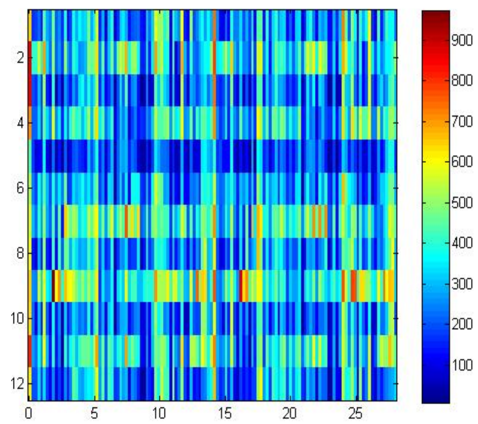


Figure 2 Example of Chromagram

2.2 Beat Tracking Algorithm

Beat tracking algorithm is the algorithm used to detect where or when the beat or beats that occur along the track. This algorithm will result a range-time where the beats throughout the song.

Beat tracking algorithm used in this final task is based on dynamic programming [2]. Dynamic programming used to optimize the algorithm to determine the starting point with the right beats and keeping the consistency of the interval between each beat. Equation 3 used to determine the beat position in this algorithm.

$$C(\{t_i\}) = \sum_{i=1}^{N} O(t_i) + \alpha \sum_{i=2}^{N} F(t_i - t_{i-1}, \tau_p) \quad (.3)$$

Symbol (t_i) is a sequence of beats produced by the beat tracking algorithm. O(t_i) is the starting point obtained from the beat of the song, α is a balancing constant, and F (Δt, τ p) is a function that measures the consistency between the estimated interval inter-beat (Δt) with the ideal length for each beat for beat (τ p) based on the predetermined tempo.

2.3 Similarity Measurement

Correlation

Correlation is measurement that often used in probability theory and statistics to show how large the correlation strength and direction of the linear relationship between variables.

The correlation between two variables produces a correlation coefficient. To calculate the correlation coefficient between two variables commonly used methods of correlation coefficient Pearson product-moment, where to find the correlation coefficient value is the value by dividing the covariance between two variables by multiplying the standard deviation. Equation 4 describes the correlation coefficient Pearson product-moment.

$$corr(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sigma X \sigma Y}$$
$$corr(X,Y) = \frac{E((X-\mu)(Y-\upsilon))}{\sigma X \sigma Y} \quad (4)$$

Where:

$E((X-\mu)(Y-\upsilon))$ is the covariance of X and Y.

σX and σY are the standard deviation of X and Y.

Correlation coefficient values range from -1 to +1, closer to -1 or +1 value of the correlation coefficient the higher the correlation between these two variables. Positive or negative values of the correlation coefficient shows the correlation of two variables, if approached +1 then these two variables have a diversity of data that direction, while if -1 then approached two of these variables has a variety of data in the opposite direction. The definition of this correlation is similar to the definition of the value of covariance; the covariance value is only able to show the variance from both of variable, whereas the correlation value of the correlation coefficient also showed large and the direction of the relationship between two variables.

3. System Overview

The systems that already built in our research is to analyze if the chroma-based representation can be used to determined the relation between ref part and full part of the music.

The full part and ref part of the music that will be test mapped into the chroma-based representation before looking whether or not there is a repetition of the parts of the song, based on assumption that ref part usually occurs repeatedly in the music.

The flow of process of this system consist of:

1. Opening music files (full part and ref part) in wav format.
2. Execute the beat-tracking algorithm to do segmentation process.
3. Map each segment from step 2 into a matrix chroma.
4. Calculate the value of the correlation between the two chroma matrix mapping the parts of the song.
5. Analyze a recurrence or not between two parts of a song based on the correlation coefficient threshold value and the value of the percentage similarity between two parts of the song.

The general process can be view in figure 3.

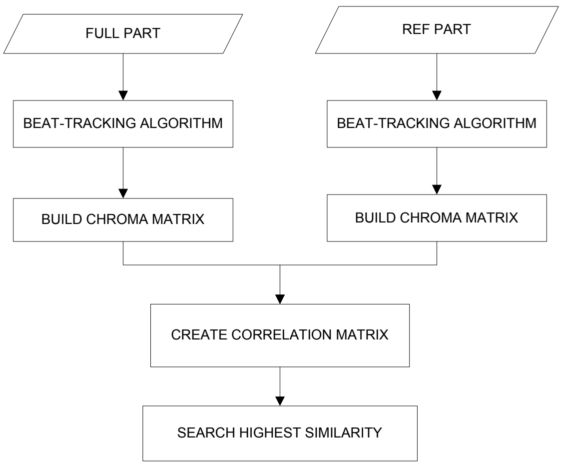


Figure 3 General Process of the System

The method, of searching similarity value between two of music file, which we purpose in our research is base-on dot-matrix matching mechanism in the correlation matrix.

The basic idea is to find the highest diagonal length in correlation matrix that qualifies the threshold value and make a similarity measurement value depend on the length of highest diagonal. Measurement will be on percentage value that can be found by divide the count of correlation value on correlation matrix that pass the threshold with the real length of diagonal.

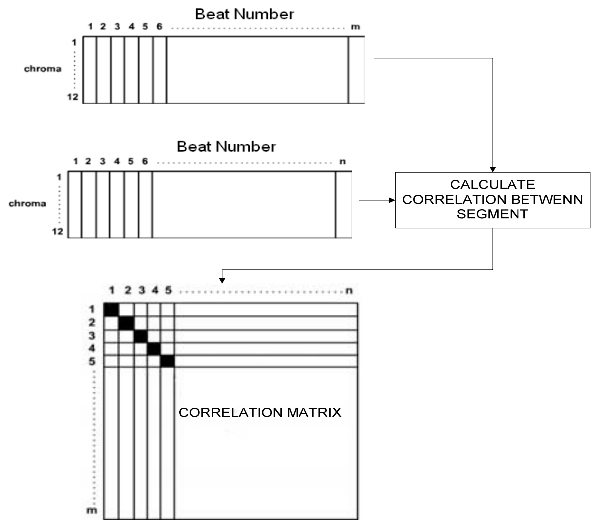


Figure 4 Process to build Correlation Matrix

4. Result

We have test our propose method in a music file with Title is Believe, by Cher. First off, all we manually extract 3-ref part from the full part of music. Then we try to compare 3 ref part using our purpose method, we expect that the method can be result some value to indicate a high similarity value.

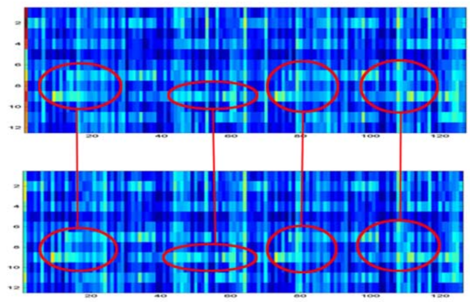


Figure 5 Similarity between ref 1 and ref 2 base on visual evaluation

Based on visual evaluation on chromagram of ref1 and ref2 from the music file Cher-Believe, can be seen that in some part of chronogram both have similar form of chroma-value. This shows that from the subjective evaluation, these two ref part have a common chromagram, now what we need is an objective evaluation to measure this common part.

Therefore, the correlation method used to indicate the similarity level from both chromagram ref parts, with the percentage of similarity parameters obtained from the diagonal matrix formed on correlation and using the search threshold value of the correlation coefficient.

In the first attempt we use threshold level at 0.95 for the correlation matrix value, the correlation value that pass the threshold indicate by red color while blue color indicate the value that did not pass the threshold value.

The correlation matrix from music file Cher-Believe illustrated in figure 6. Using threshold level at 0.95, we get 3 long diagonal view from the correlation matrix. Lower left diagonal shows that the high correlation between ref 1 again ref 2 occurs in the last part of ref 1 and in the beginning of ref 2. An upper right diagonal show the high correlation of the correlation was occurs in the beginning of ref 1 with the last part of ref 2. The longest diagonal in the middle of the matrix shows the correlation along the ref 1 and ref 2.

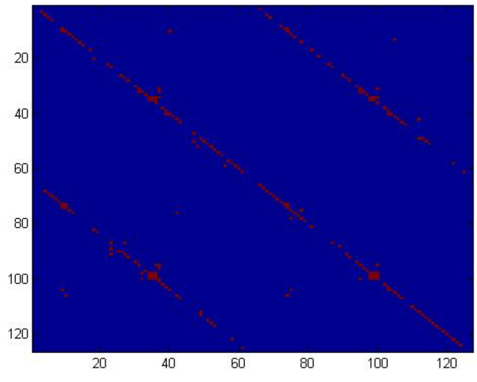


Figure 6 Correlation Matrix of ref 1 and ref 2

From the results of testing on music file Cher - Believe by counting the red color of the longest diagonal that occurs in the correlation matrix and divide it with the length of diagonal we find the percentage of similarity by a diagonal matrix correlation of ref 1 and ref 2 was 70.635% with 95% confidence level. Figure 7 inform the more detail result of similarity value from music file Cher-Believe.

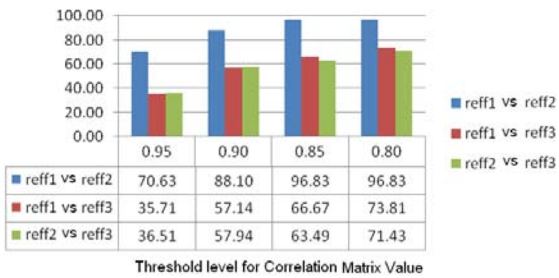


Figure 7 Detail result of the experiment

These results would be different if applied to other test data. In this research, we used 15 test data of music file with 3 different genres, namely pop, rock, and acoustic.

Like our first experiment, we extract the ref part manually and test it with our purpose method. We use different threshold level based on the characteristic of the genre. In rock genre we limit the minimal threshold at 0.80 considering the noise level in this genre is high, while the other genre we use minimal threshold at 0.70.

The result is our purpose methods are able to get high percentage level from 12 test data.

5. Conclusion

The results of our research showed that the percentage of the value of the correlation threshold on a high level of trust can still produce high percentage value that indicate the similarity of two music file. The conclusion is the chroma-base representation can be use to make a similarity measurement to validate the preview of the music file.

6. Feature Work

The method that we purpose are only able to make a similarity measurement from two music file, one can be consider as the full part of the music file while the other are the preview /ref part, in this research we make the preview part manually. The next research can be targeted to make an automatic preview of the music file base on chroma-base representation using sampling mechanism.

References

[1] Bartsch, Mark A and Wakefield, Gregory H., 2005, “Audio Thumbnailing of Popular Music Using Chroma-Based Representations”, IEEE : Transactions on Multimedia Vol. 7 No. 1

[2] Ellis, Daniel P.W., 2007, “Beat Tracking by Dynamic Programming”, LabROSA, Columbia University, New York.

[3] Ellis, Daniel P.W., 2007, “Identifying ‘Cover Song’ with Chroma Features and Dynamics Programming Beat Tracking”, LabROSA, Columbia University, New York.

[4] Lartillot, Olivier, 2008 , “ MIRtoolbox 1.1 User’s Manual”, University of Jyväskylä, Finland.

[5] Santoso, Singgih., “Structural Equation Modelling, Konsep dan Aplikasi dengan AMOS”, Elex Media Komputindo, Jakarta, 2006.

[6] Rummel, R.J., 1976, “Understanding Correlation”, University of Hawaii.

[7] Schmidt-Jones, Catherine., “Amplitude and Dynamics”, http://cnx.org/content/m12372/latest/, access time 3 Februari 2009.

DESIGN AND SOFTWARE IMPLEMENTATION FOR WIRELESS AUTOMATION IN PLC BASED USING JAVA PLATFORM

Efri Suhartono[1], Bambang Riyanto[2]

[1,] Electrical Engineering, Institut Teknologi Telkom
[2] School of Electrical and Information Technology Bandung Institute of Technology
[1]efr@ittelkom.ac.id, [2]briyanto@lskk.ee.itb.ac.id

Abstract

The role of control systems which is easily controlled either by the plant or can be done anywhere has become a vital need and has been needed in the automation system .Mobility and flexibility that characterizes the wireless technology is an alternative that can be utilized for the purposes of telecontrol and telemonitor. In this research,it has been carried out to design automation system design by using J2ME (Java 2 Micro Edition) that runs on handheld devices (emulators / hp) in user side, and application server used J2EE Platform and desktop application using J2SE. J2ME is one of the Java platform for the micro version, can be applied to mobile devices devices. Some of the benefits of java in the implementation of wireless applications, the java as a secure programming language. Java language is strong language, and portable. J2ME provisioning application requires only that a number of IP or HTTP address for accessing the web server or browser wml, while for the WAP Gateway for transcoding needs. However, J2ME can interact with Web servers via WAP or WML browsers. The design is started from the PLC to PC server communication is done by using the format C-Command communications so they can directly access to the PLC device. The data was taken and then processed and made an application server (servlet) which is placed on apache tomcat .. Application run on the handheld device to access the server, then used to control the PLC device. Beside that, application can be run from web browser that access JSP from server side. The time that required to access the server was fast, so memory usage is still below the heap memory, and the number of bytes required is still small.

Keywords: PLCs (Programmable Logic Control), J2SE, J2EE, J2ME (Java 2 Micro Edition), C-Command, servlet, JSP apache tomcat.

1. Introduction

The development of technology was rapid, will lead the industrial world environment in order to meet production targets, quality and continuity and have a rapid response to changes in industrial systems. The role of control systems are easy to operate either close to or control plant that can be done anywhere has become a vital requirement for the condition of the status of an object / plant can be monitored and controlled.

Mobility and flexibility that can be characteristic of wireless technology is an alternative that can be used for the purposes of controlling, monitoring.

Java 2 Micro Edition (J2ME) is one of the Java technology that allows mobile users can access and interact with information and services wireless applications quickly and easily, using mobile devices such as mobile devices and PDAs that have the characteristics of each [10].

Physical architecture is needed is to use the device PLC (Programmable Logic Controller), PC as the control PLC and applications J2ME (Java 2 Micro Edition) for Mobile devices. Side J2ME client for wireless applications, J2EE (Java 2 Enterprise Edition) for the server device, and access to communications protocol (interface) side of the PLC through the Command Interpreter J2ME enabled devices can access the PLC. In the automation system to move a data set of plant conditions and the status of current conditions, developed by environmental mobile devices. characteristics of mobile devices different causes mobile user can not access the wireless application from an existing network providers (eg WAP) in a multiplatform environment.

2. Basic Teory

2.1 PLC devices

PLC (Programmable Logic Controller) is a device (device) comprising control CPU (Central Processing Unit), Memory Unit, Unit Input / Output and Power Supply Unit (Power Supply) and is programmable (programmable)[7].

A PLC will accept input data (digital or analog), which comes from input devices (sensors, switches, push button, etc.) through the input unit.

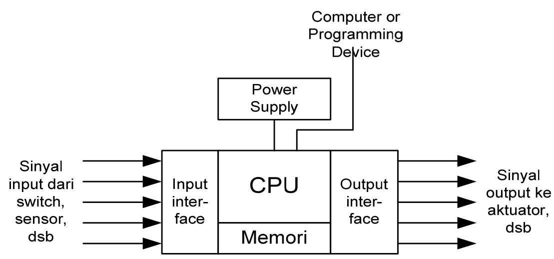


Figure 1. Block Diagram of PLC [8]

2.2 The Main Components of The PLC

CPU (*Central Processing Unit*)

CPU of a PLC usually uses microcontroller or microprocessor. depending on the manufacturer of the manufacturer.

Memory

The types of memory commonly used, among others:

a. ROM (Read-Only Memory)
b. RAM (Random Access Memory)
c. EPROM (Erasable Programmable Read Only Memory)
d. EEPROM (Electrically Erasable Programmable Read OnlyMemory)

Unit Input /Output

A unit that is connected directly to the input devices and output, such as: sensors, switches, buttons, motors, cylinders, display, etc. Power Supply (Power Supply)

2.3 PLC Programming

PLC programming is a process of determining the sequence of commands that will solve a specific algorithm associated with the task of control . it can used other language form of programming language that is [8]:

o Instruction List
o Relay or Ladder diagram
o Function diagram
o Structured text language

2.3.1 Omron CS1GH CPU44 H

PLC is a PLC type of medium. PLC general specifications of this type are:

1. Capacity of I / O 1280 bits.
2. Program memory (the total number of instruction steps in the program eg. LD and OUT takes 1st step, but MOV (021) need 3 step instructions) 30K.
3. 64K data memory words.
4. External memory data 32K words x 3 banks.
5. Instruction length of 1 to 7 steps per instructions.
6. Execution time 0.02 min (basic instructions), 0.06 µs min (special instructions)

2.3.2 CX Programmer V 3.2 Omron

This program is needed to configure the user memory in the PLC hardware algorithm that has the function of the process required, then program the input and output numbering both the discrete signals and process the analog process.

2.3.3 PLC Communication Format of Instruction CS1GH-CPU44H

This type of PLC may receive instruction formats of communication with:

1. C-mode commands
   This instruction is used for the communication format of Host Links (serialcommunication).
2. Finds Commands

a. This instruction format is used for a variety of networks (eg Controller Links, Ethernet, etc.) and serial communications (Host Links).
b. Maximum PLC that can be handled by the PC when using the C-Commands are 32 pieces, whereas if you use a maximum of fins 4064 Commands fruit. Examples of instruction format C-Command [9]:

C-Frame Commands

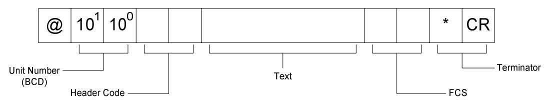


Figure 2. C-Commands Frame

***C-Commands Response* Frame**

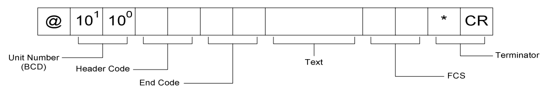


Figure 3. C-Commands Response Frame

- @ : Markers beginning of a command.
- *Unit Number* : The numbering PLC units are connected to the PC, its value in BCD (0-31).
- *Header Code* : a special character which includes command.
- *End Code* : Parameters that indicate the results of command execution.
- *Text* : It is the response parameters.
- FCS : Includes the value of the calculation of FCS (Frame Check Sequence).
- *Terminator* : It is a combination of characters "*" and CR (Carriage Return) / Chr (13), and used as a command end marker.

Table 1.Examples of Instruction Format C Command [9]

Tipe	Header Code	Name	Function
Read memory I / O	RR	Read area CIO	Reading of CIO word
	RD	Read area DM	Reading area of DM word is sent by the DM Word.
write memori I/O	WR	Write area CIO	Writing area-specific data words with the word in appointed CIO.
	WL	Write area LR	Writing LR area-specific data words with thewordinLR sent.
Forced set/ Reset	KS	Forced set	Forcing set bits set to 1 on the intended
	KR	Forced reset	Forcing set bits set to 1 on the intended

Instruction Format **FINS (FINS Command Frame)** [9] :

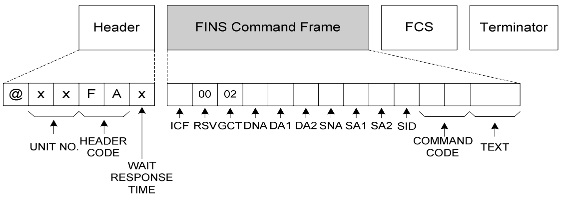


Figure 4. FINS Commands Frame

FINS Response Frame

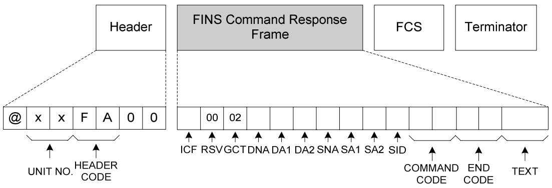
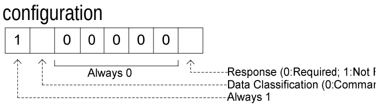


Figure 5. FINS Commands Response Frame

Caption 4 and 5 :

- @ : Markers beginning of a command.
- *Unit Number* : The numbering PLC units are connected directly to the PC.
- *Header Code* : In the C-Commands header contains the command code
- *Wait response time* : Set the length of a PLC to give response to the PC.
- ICF : Information Control Field, with the following configuration



- RSV : Backup, worth 00.
- GCT : Gateway Count, the number of gateways that pass. When sending a command value 02.
- DNA : Addressing the target tissue (00h - 7FH).
- DA1 : Addressing the target node (00_H – 20_H).
- DA2 : Addressing the target unit
- SNA : Addressing the source network (00h - 7FH).
- SA1 : Addressing the source node (00h - 20H).
- SA2 : addressing unit source
- SID : *Service* ID
- FCS : *Frame Check Sequence*
- *Terminator* : the end of the command ("*"+ CR).

2.3.4 FCS *(Frame Check Sequence)*

Used in PLC and the PC to check the command frame and the frame response is received, FCS value obtained by performing XOR operation (Exclusive OR) at each command or response characters (except "*" and CR).

Architecture of the current control system can be described as follows in Figure 6.

2.4 Information Technology in Industrial Automation

With technological developments and increasingly intense competition in the demanding industrial environment, industrial system must be efficient, effective and integrated in order to have the quantity and quality of the optimal production. Examples of applications in the industrial world in Figure 7.[24]

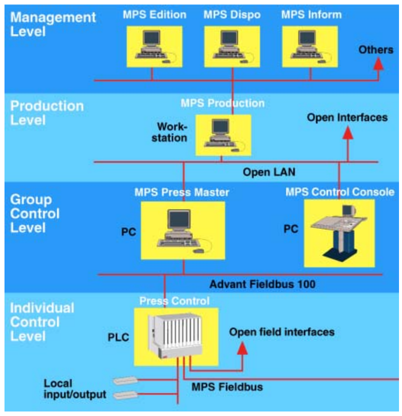


Figure 6. Control System Architecture

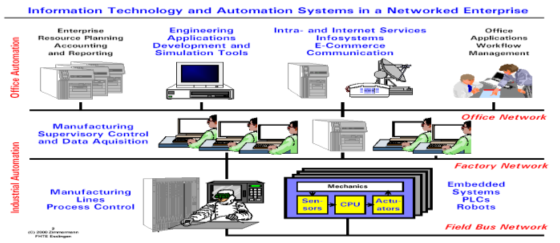


Figure 7. Information Technology and Industrial Automation

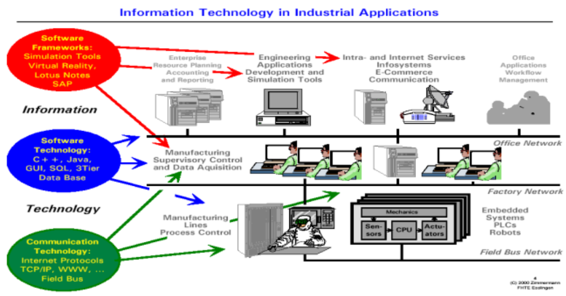


Figure 8. Application of Information Technology in The Industry [24]

Internet service can be access via wireline or wireless. Via the wireless control can be done via SMS, WAP or java applications (J2ME). Examples of e-control architecture is applied by Alstom [26]:

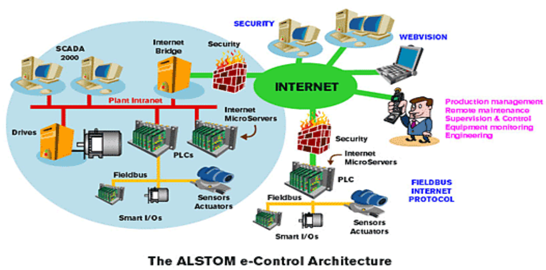


Figure 9. E-Control Architecture at ALSTOM

2.5 Trend Industrial Automation System

Interoperability at the level of technical / devices to the level of enterprise resource planning (ERP) should be able to easily communicate with each other and between devices and software from various vendors:

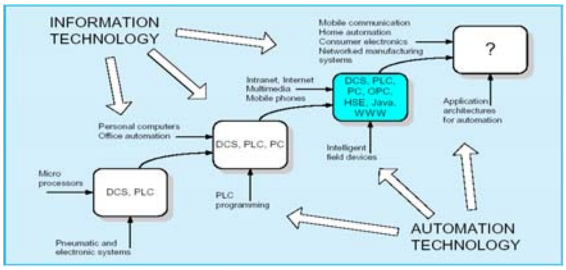


Figure 10. New Generation of Control Systems

In industrial automation, these radio systems are used to transfer specific data for long periods of time (not real time). Wireless system solutions to industrial control systems has become critical things such as a portable user interface, remote monitoring and service message. [26]

2.6 Java 2 Micro Edition (J2ME)

J2ME is one part of Java 2 technology, Sun developed Microsistem for java programs can run on mobile devices or handheld devices like mobile phones, Palm, PDAs, and Pocket PC.

Figure 11. Work Environment Java Technology

Other categories in the work environment is the Java 2:

- Java2 Standard Edition (J2SE)
- Java 2 Enterprise Edition (J2EE), this category to run and develop java applications in enterprise environments, with added functionality such as EJB (Enterprise Java Bean), Java CORBA, servlet and JSP, and XML (Extensible Markup Language).

J2ME consists of components as follows:

- Java Virtual Machine (JVM)
- Java API (Application Programming Interface)

Other Tools for the development of Java applications, such emulator Java Phone,

2.7 J2ME Configuration

J2ME Configuration mendefinisikan lingkungan kerja J2ME *runtime*. Ada dua kategori J2ME Configuration yakni:

- ➢ CLDC (Configuration Limited Device Configuration)
- ➢ CDC (Connected Device Configuration)

2.8 WEB Component

Java Servlet

Servlet is an application used to run a server-side java-based application (JSP / Servlet) to a server. Servlet can be incorporated in the Java web server or as a separate container such as Tomcat or Jetty. MIDlet connections to the database server through a servlet, is actually a J2ME integration with J2EE.

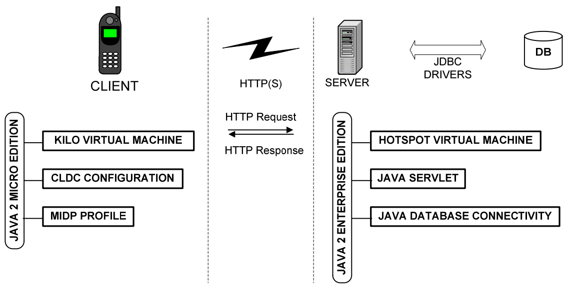


Figure 12. J2ME/J2EE Integration [4]

Java Server Pages (JSP)

JSP allows the components are combined and working in a web page and then sent to the client. A JSP can consist of HTML elements, java code and Java Bean components.

Java Database Connectivity (JDBC)

JDBC is the application programming interface that provides universal data access for the Java programming language. Technology JDBC drivers can be divided into 4 categories[13]:

1. JDBC-ODBC Bridge plus ODBC Driver.
2. Native-API Partly java technology-enabled driver
3. Pure Java driver for database middleware (JDBC-Net)
4. Native-Protocol Pure Driver

With a driver written in java programming language, all the information necessary to make the connection defined by the JDBC URL or a DataSource object is registered with a JNDI (Java Naming and Directory Interface).

2.9 Development of J2ME Service

J2ME Application provisioning requires only that a number of IP or HTTP address to access the web server or a wml browser. interaction between the java application with WAP or i-Mode can be modeled in the image below.

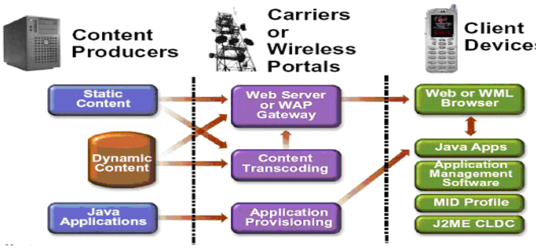


Figure 13. Interaction Java Application with WAP or i-Mode [17]

3. Requirement SystemAnalisys

3.1 User Requirement

Implementation is used for discrete control of the loads for the conditions on / off. How load conditions at the same time monitoring can be done (from the plant / load), perform control functions from the user so that particular moment load conditions change according to the desired. Some of the things needed by the user for it as follows:

a. Data processing facilities
b. Programming environment
c. Execution environment

Retrieval and delivery mechanisms of data, divided into six parts namely:

1. Process control mechanisms in the plant. This function is performed locally by OMRON PLC CPU 1GH CS 44H.
2. PC communication mechanism with the PLC using a special instruction formats that are recognized by the PLC hardware.
3. The mechanism of data communication network provider.

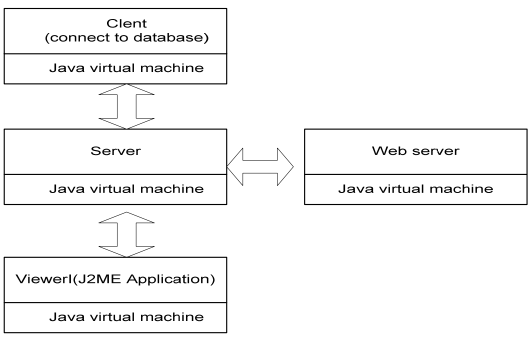


Figure 14. The Mechanism of Data Communication Network Provider

4. Mechanism of mobile information providers.
5. Storage and retrieval of data.
6. Visualization process

From the system specification that has several functions and operational requirements, the system will be made can be divided into 4 subsystems, namely:

a. Local control subsystem
b. Subsystems PC / Server
c. Database subsystem
d. User subsystem

4. System Design

Model as a whole system to be built.

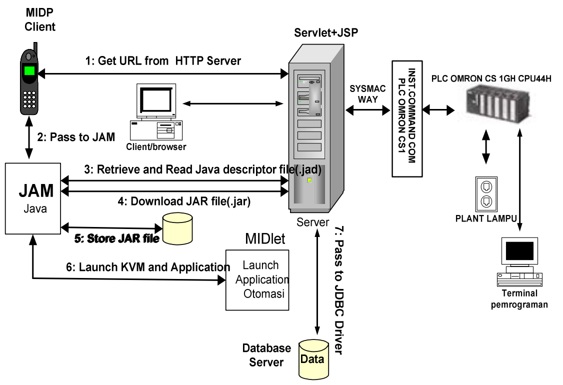


Figure 15. Automation Control System Architecture Using The J2ME Platform

In summary, the integration will be built between the J2ME client side with the J2EE (servlet, JDBC) server side can be seen in the picture below.

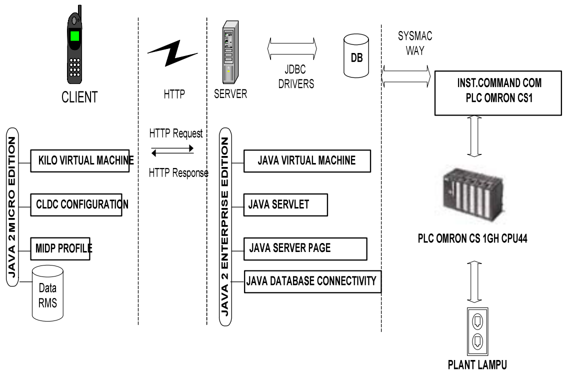


Figure 16. Integration of J2ME-J2EE System-Controllers

System architecture will be built using the architecture of three layers namely:

1. presentation layer or user interface
2. middle layer
3. data management layer

4.1 System Requirements

The requirement systems designed include:

- Applications can be run on mobile devices emulators.
- Users can control the change in plant (run / load) is desired.
- Users can monitor plant data that occurs.
- Users can clear the status of the CPU unit occurs in the PLC.
- Users can find out when the emergency conditions on the device.

The diagram below will explain the relationship between the cases and actors and activities in it:

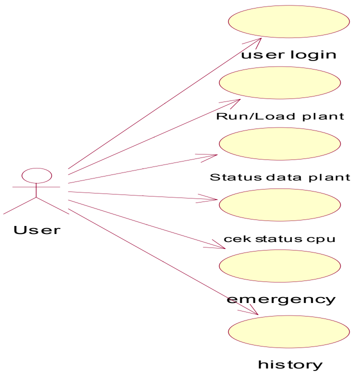


Figure 17. Use Case Diagram

4.2 Design of Scenario Process

The scenario process can be described in sequence diagrams. Sequence diagrams describe interactions between objects in and around the system (including the user, displays and so on) in the form of message is described with respect to time.

Collaboration diagrams represent interactions between objects as well as sequence diagrams, but more emphasis on the role of each object.

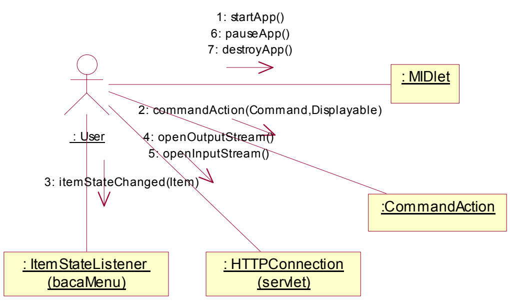


Figure 18. MIDlet Collaboration Diagram

Sequence diagrams that occur in the http servlet can be seen in the picture below.

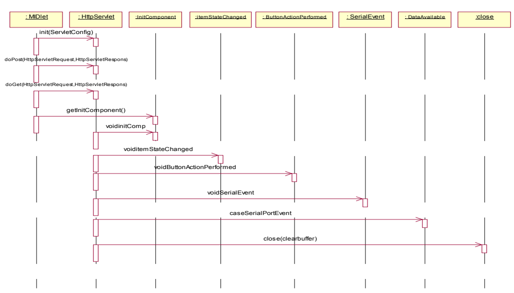


Figure 19. ***Diagram Sequence HttpServlet***

Http servlet Collaboration diagrams can be seen in the picture below.

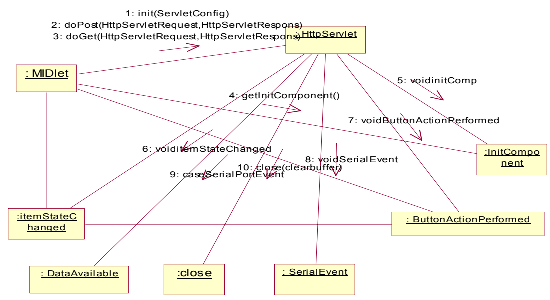


Figure 20. ***Collaboration Diagram HttpServlet***

4.3.1 State-Chart Diagram

Statechart diagrams describe the state of transition and change (from one state to another state) of an object on the system as a result of stimuli received.

4.3.2 Class Diagram

Class diagrams describe the structure and description of the class, package and objects and their relationships to each other like Containment, inheritance, associations and others.

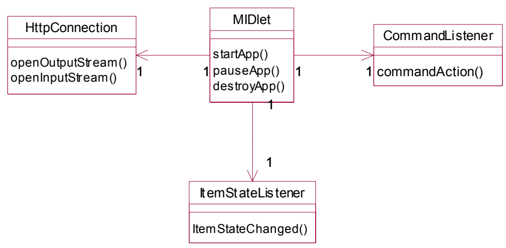


Figure 21. Diagram *Class Midlet*

Class diagram is needed in the servlet can be seen in the picture below.

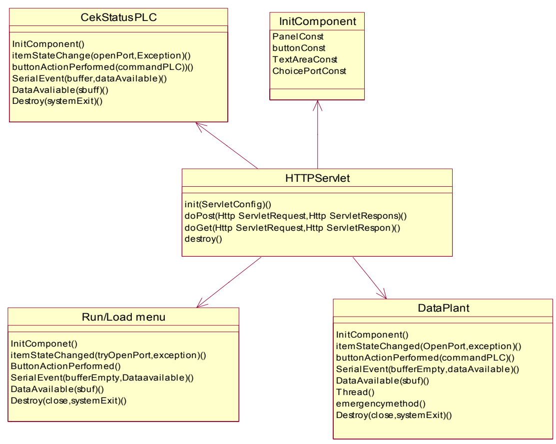


Figure 22. HttpServlet Class Diagram

4.3.3 Component Diagram

Component diagrams are physical description of the model being built.

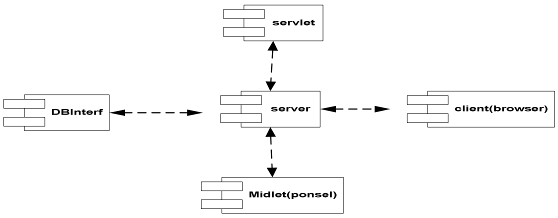


Figure 23. Component Diagram of An Application Made

Besides these applications, other applications involved in running the system are:

- o v.4.0.1 as Jakarta Tomcat web server facilities.
- o MySQL database server as a server.

4.4 Description of The System Design

Secara garis besar tahapan proses pengambilan data dibagi menjadi :

Mostly the data acquisition phase of the process is divided into:

1. Preparation of the PLC hardware requirements
   For this process required:
   Digital inputs: 4 pieces (3 pieces for on / off lights,
   Digital output: 4 pieces (3 pieces status on / off lights,
2. PLC programming Allocation of address used in PLC programming
   o Providing a digital address input and output Digital input with 0:00 until 0:03 address Digital output given 1:00 to 1:03 address
   o Set up sequencing process.
3. Building a PC with a PLC communication using C-frames Command Example C-Command command given, so no response from the device are:
   - o "@ 00SC0252 * (enter = \ r \ n)" response from the device when giving frame "@ 00SC0050 *", 00 characters worth 6.7, mean response in normal conditions (true), if the 13 FCS error, format error 14 .
   - o "@ 00KSCIO 0001003C * (enter = \ r \ n)" force command set (on) the condition of our output set to "on", the output module to address 1 (0001), the output 0 (00), FCS = 3C. response format "@ 00KS002A *", if 6.7 is worth 00 characters, the response to normal conditions (true). To force reset (off) on the transmitted light format "KRCIO @ 0001003D * (enter)". Format is almost the same response.
   - o To check the status of any input condition is active, by sending data format "@ 00RR0000000141 * (enter)".

4.5 Design Interface

Structurally, the organization can view the program described as follows:

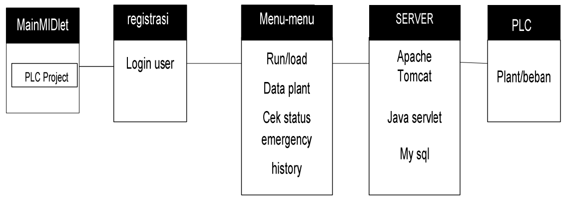


Figure 24.Design of The Main Display and The Process to The Server

5. Analysis and Results

5.1 PLC Programming

Instruction is done to perform the control and monitoring functions of the PLC are:

- o @ 00SC0252 * \ r \ n (to do the monitoring mode PLC,). Response of normal format: @ 00SC0050 * (enter).
  00KSCIO o @ 0001003C * \ r \ n (to set on for input to 1). KRCIO @ 0001003D * \ r \ n (to set off for input to 1).
- o 00KSCIO o @ 0001013D * \ r \ n (to set on for input to 2). The normal response of the PLC format is @ 00KS0040 * (enter).
  KRCIO @ 0001013C * \ r \ n (to set off for input to 2). @ 00MS5E * \ r \ n (to do the reading status of the operating conditions of the PLC CPU units). Response format (depending on the condition of the CPU units), @ 00MS000330 * \ r \ n (normal condition with the monitor mode, the number of programs 30kstep max).

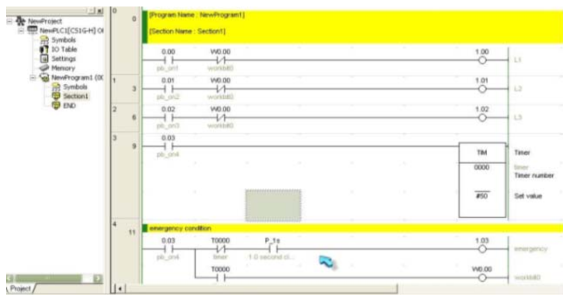


Figure 25. Ladder Diagram in The Case of This Research

5.2 Display the user interface

Display images occurred in java application (server) that direct access to the device controllers. Here are some examples of applications in the MIDlet:

Figure 26. Application User Login and View Menu to Control

By selecting the menu that will be controlled it will go to the next menu.

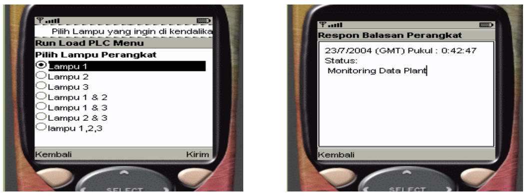


Figure 27. View Menu to Run / Load The PLC, Connect

The server and responses going from the server

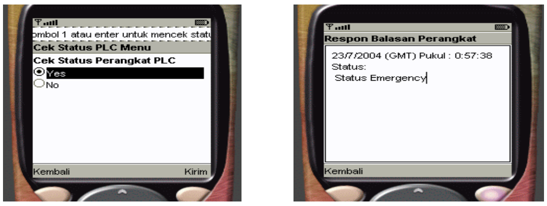


Figure 28. Commands for Monitoring Plant Data and Check The Status of The PLC CPU

From the picture above for monitoring the status of all conditions that occur at the output of the PLC device, whichever output is running. In the picture is a continuation of diataas check the status of the previous PLC devices, and reporting to the user when emergency conditions occur in prangkat / load, so that is also recorded at the time when the emergency occurred.

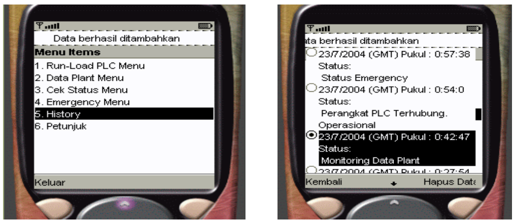


Figure 29. Status Data (History) of All Control Activities

Figure 30. Java Application Interface on The Main Menu PLC Control

Starting from port choice to be used and followed by the election of such controls performed on J2ME MIDlet application that previously. Before any control, the device must be set in the PLC monitors conditions (there are three conditions, program mode, run, and monitor), to be accessed from the application made.

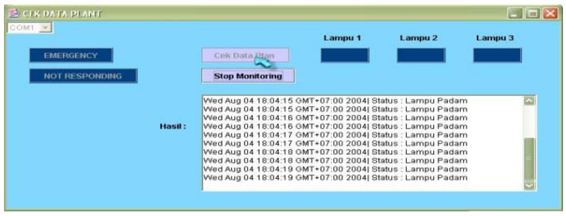


Figure 34. Monitoring The Load When The Condition is Still Out

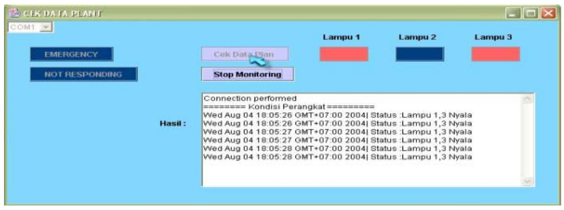


Figure 31. PLC Monitoring Devices, Emergency Conditions and Functions of Data

Visualization tool is also displayed when the output is active everywhere. Look for the light switch that changes color 1.3, so the condition is "on".

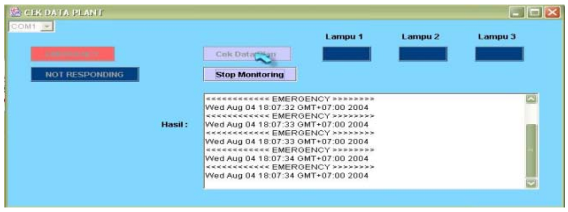


Figure 32. Emergency Status Which Occurred in The Application

Conditions are also marked emergency lights on-off turns and the timer device is set when the PLC has to do calculations for five seconds, then all will be off the output device. Menu Run / load when first switched on and connect to the serial port to access the PLC device.

Figure 33. Menu Run / Load When L1 on and L2, L3 Off

5.3. Evaluation of Latency Time

Latency is the time to analyze the data access time is calculated from the MIDlet to request data from the server and then access the PLC until the MIDlet get a response from the server returned to the MIDlet value in units of milliseconds. By calculating the Latency Time to connect to the server and the RMS, produced the following Table 3

Table 3. Data Latency Time MIDlet-Server-PLC

Time	PLC Operation	Time (ms)
16-8-2008;15.02	Run /load PLC (lamp 1 on)	4210
16-8-2008;15.07	Lamp 2 on	3866
16-8-2008;15.35	Lamp 3 on	4024
16-8-2008;15.47	Lamp 1&2 on	4137
	All lamp on	4238
30-8-2008; 9.30	dataplant Monitor	4320
30-8-2008; 9.35	Check status CPU	4023
30-8-2008; 9.38	Emergency status	3976

5.4. Analysis of Total Byte Data Transmission

Tests to determine the amount of bytes sent and received by the phone. The magnitude of this byte determines the total cost to be paid by the user.

Table 4. Table Total Byte Data Server Emulator + Local

Task Name	Upload (bytes)	Download (bytes)	Total (bytes)
Menu L1-on	7	25	32
Menu L1-L2 on	8	31	39
All the lights burning	7	31	38
Monitor data plant	7	31	38
Check the status menu	7	46	53
Emergency	7	26	33
Total (bytes)			233

5.5. Final Analysis

Mobile applications monitoring capabilities and control (functional system):

- Capable of controlling the PLC devices based on user Inputan.
- Ability to display the Run / Load the program on the device controller.
- Ability to plant data monitoring any of the current and reported when it happened.
- Ability to view history (log) on the mobile application that is known all activities that occur.
- Applications can be run on some mobile devices (cross platform).

Bibliography

[1] Allmaraju. "*Java Server Programming J2EE Edition*". WROX Press Ltd, 2000.

[2] Moss Karl, "*Java Servlet*". Mc GrawHill, 1999.

[3] Devine, Christophe, "J2EE Tutorial" devine@cr0.net. USA. 2003.

[4] Tramblett, Paul , " *Instant Wireless Java with J2ME* ". California – USA : Mc-Graw-Hill/Osborne . 2002.

[5] BoodOff Stephanie,Green Dale, "*J2EE Tutorial*". Addison Wesley .2002.

[6] Keissel E Thomass, "*Industrial Electronics*". Prentice Hall. 2000.

[7] Dunning Garry. "*Introduction to PLC*". Delmar Publisher, 1998.

[8] Omron, "*A Beginner Guide to PLC", Omron Singapore,2001.*

[9] ________, "Instruction Command Communication", Sysmac CS Series Omron, 2001.

[10] ________, "*Programmable Controllers, Operation manual*" Sysmac CS Series Omron, 2001.

[11] Samuel,Albert& Setyabudi Agus, "E-Commerce Applications with the Java Servlet and JSP", 2002.

[12] Wicaksono. Ady, "Wireless Application Programming with Java" , Jakarta : Exel Media Computindo, 2002.

[13] Ian Leonardo, "Database Programming with JAVA", Jakarta: Exel Media Computindo, 2003.

[14] Rickyanto.Isak., "*Java Server Pages*", Jakarta : Exel Media Computindo, 2002.

[15] Hutahaean JJ Junko, "System PLC based automation moves through SMS and WAP, the final project report, majoring in electrical engineeringp ITB*, 2003*

[16] Setyo Renault, "Supervision and Monitoring of distributed processes in the industry with JavaTM-based framework, Master Thesis, ITB majoring in electrical engineering 2002.

Website

[17] http://java.sun.com

[18] http://jakarta.apache.org

[19] http://www.servlets.com/jservlet2/examples/index.html

[20] http://www.dotmobile.home.pl

[21] http://www.billday.com/j2me,

Papers / Articles and Presentations

[22] www.vth.fi/tuo projectit/ohjaava/ohjaava1/atr_2001.pdf.

[23] www.intel.com/business/bss/solutions/blueprints/pdf/30064501.pdf

[24] www.manufacturing.net/ctl/article/CA188645

[25] www.lamspeople.epfl.ch/kirrmann/slides/AI_15_Architecture.ppt

[26] http//developers.sun.com/techtopics/mobility/midp/ttips/httpcon/ making HTTP connections with MIDP.

CARICARI GAME DESIGN AND IMPLEMENTATION AS A WEB BASED MINI GAME

Purba Daru Kusuma

Institut Teknologi Telkom
purbo_daru@yahoo.com

Abstract

Caricari is a web based game that is installed on www.tombongantuk.com. Here, user's mission is to find the corresponding letter in the letters box. This game is designed as a mini game. As a mini-game, Caricari is designed to be a game that is not played in a long time, short scenario, simple game, and does not make users become addicted. Game is designed to be played by users with a wide age range, that are 7 years and over. One of the factors taken into account in game design is the size of the data. As a web based game, downloadable data size becomes a crucial factor when the game is downloaded. The programming activities are divided into 2 activities, the server side programming using PHP and programming languages on the client side using Javascript. Artificial intelligence is used when downloading the words to search for letters and at scrambling the letters on the letters box.

Keyword: web based game, server side programming, client side programming

1. Introduction

Caricari is a new game on game collection web sites, www.tombongantuk.com. Caricari is designed as a web based mini game. This concept is chosen because the aim of this game is to give breaking time for users. Web based technology platform is chosen considering a web based application is multi platform. Thus, the expected range of users can be more broadly. With web based application platform, users do not need to install the application to start playing.

Caricari is developed by considering several aspects. The first aspect is fun. Caricari is designed to give fun experience for users. One of important goals of game is providing fun experience for users. The second aspect is the level of addiction. Some games can be so addictive so users can abandon their other important activities. Caricari is designed not to make users addicted. Thus, the game duration should not be too long and the game scenario is simple. The third aspect is a challenge. Caricari is designed to provide enough challenge level so that users do not get bored easily and users will repeat the game until a certain amount. The fourth aspect is the mission complexity. The mission is easy to understand. These are selected considering the user segment of Caricari is common internet user communities. Thus, users will understand the game as quick as possible. If users can't understand the game rule in a short time, users will tend to cancel playing. The fifth aspect is the interaction with other users. Caricari is designed to provide certain limit inter users interactions.

This research is expected to provide some added value in national game industry development. First, this research is expected to provide a framework for a simple fun game development. Second, this research is expected to provide alternative scenario for further game development. Third, this research is expected to provide alternative web based mini game algorithms.

Caricari is developed within several steps. The first step is searching for ideas. The second step is designing game scenario. The third step is designing algorithms which are needed for various processes in this game. The fourth step is exploring built in functions in Javascript and PHP which are used in the game. The fifth step is implementing game design to programming codes. The fifth step is launching by placing the game files in the www.tombongantuk.com server so public can access the game. The sixth step is testing and monitoring public acceptance.

2. Basic Theory

Game development processes are varies depending on the company or the project management style. However, there are common steps in game development.

The first step is the pre production. The second step is production. The third step is testing. The fourth step is completion. The fifth step is maintenance. The first step is consists of proposing game concept, designing project management, and designing game. The second step is consists of coding the game, developing model, and making sprite. The third step is consists of testing game features, scenarios, levels, and bugs. The fourth step is consists of fixing bugs and adjusting game levels, scenarios, and duration. The fifth step is consists of handling defects that are founded after the game has been launched.

Game design is process of designing the game content and rules. Generally, game design is a job that involves multiple disciplines. Game design is a job that needs coordination between several fields such as:

- game mechanics,
- visual art,
- programming,
- production process,
- audio,
- script.

3. Design

The first step in the making of Caricari is creating game concept. The basic idea is training and testing user's accuracy. Users must find a letter in a letters collection box. Letters that must be found is letters in certain words. The letters in certain words searched sequentially from the first letter to the last letter in that word. If the last letter of a word has been found then a new word will appear. In one level, there are 5 words whose letters must be founded. If all letters in all words are found in the provided timeframe, the game is finish. After the game is finished, users can upload their score into the server, upgrade level, or start new game at the same level to improve the score. Caricari is divided into 2 versions, for members and non members. The difference of the version is the user's right to upload score so other user can watch the top list. There are 4 levels in Caricari. Each level determines the number of letters in a word to be looked for matches. List of level relationships with the number of letters in a word is as follows.

Table 1. Relation between level and number of letter

Level	Number of levels
1	4
2	5
3	6
4	8

Scores is determined by the remained time. While the game is played, the score value is continuing to decline. The decrease of score is stopped if user has completed all the letters in one period of the game. Thus, the faster the user completes the game score the higher score is obtained. The decrease of score runs every 0.5 seconds. If the score has shown zero while the user has not completed the game then the game is frozen and the user is fail to finish the game. Furthermore, the user can start a new game. Initial value is determined based on the game level. Formula for determining the initial value of the score is as follows:

$$\text{Score} = \text{level} \times 120 \qquad (1)$$

Score that has emerged affects 2 things. Member user can upload the score into the database. The first effect is the score will be entered into the highest scores list. The second effect, the score will add member user point and can be used as capital to play in another game. There is possible condition that there are more than one score from same users. The reason is the score table records every game session that are uploaded and not only the highest score of each user who has played this game.

There are 2 tables in this game. The first table is a table that contains words list that will appear in the game. The second table is a table that contains score records that are uploaded by members. The word list table is consists of two fields, namely word and level. The data type of word field is varchar. The data type of level field is integer. The score table is consists of two fields, namely user and score. The data type of user is varchar. The data type of score is integer.

The game programming is done in two parts. The first part is the server side programming. The second part is the client side programming. Server side programming is focused on the processes before and after the game. Client side programming is focused on the processes during the game.

There are several processes that are coded in the server side programming. These processes are:

- deciding playing access,
- choosing words,
- creating objects,
- entering score.

There are several processes that are in the client side programming. These processes are:

- placing objects,
- setting parameters,
- setting timer,
- shuffling letters,
- catching answer,
- checking answer,
- uploading score.

Another job that is not less important is the objects visualization design. Object visual design gives attention to several aspects. The first aspect is letters box color selection. The second aspect is font size selection. The third aspect is font type selection. The fourth aspect is the collection box frame design.

Figure 1. Letters Collection Box Design

Figure 2. Word Design

Artificial intelligence in this game is performed on 2 processes. The first process is selecting 5 words in the database that will be sent to user. The second process is shuffling the letters in the letter box. The 5 words selection process is started with generating random number from 0 through 5. Next, 5 words are taken from words table based on the playing level starting from the sequence shown by the random number. Thus, the possibility of user accessing the same set of words with other user will be small. The letters shuffling processes in the letters box begins with the letters in the letters collection box. This setting is done by setting the sequence to be inserted into the 50 member array alphabetically. 50 numbers is obtained from the letters amount that will be accommodated in the box set of 50 letters. Thus there is no alphabet that is appeared more than twice and on the other hand there is no alphabet that does not appear. The next process is swapping letter position with another letter in the box randomly.

4. Implementation

The Implementation begins by changing algorithms in game design into program code. Programming language that is used in server side is PHP. Programming language that is used in the client side is Javascript. Database that is used in the server is MySQL.

Some factors are taken into account in selecting the PHP as server side programming language. The first factor is PHP is open source. The second factor is the PHP script is not need to be compiled. The third factor is the web server that is used in tombongantuk.com is Apache.

Some factors are taken into account in selecting Javascript as client side programming language. The first factor is Javascript is open source. The second factor is Javascript manual is easy to find. The third factor is now nearly all web browsers, especially Mozilla Firefox and Internet Explorer supports Javascript language.

Some factors are taken into account in selecting MySQL as a database engine. The first factor is MySQL database engine is provided in tombongantuk.com server. The second factor is PHP provides special instructions to access the MySQL database. The third factor MySQL database is managed easily by the PHPMyAdmin package.

Server side programming consists of several functions. These are the functions in the server side programming:

- logo function,
- main menu function,
- appearance function,
- score entering function,
- board showing function,
- instruction function,
- top score function,
- javascript showing function.

Client side programming consists of several functions. These are the functions in the client side programming:

- start function,
- word setting function,
- letter setting function,
- letter shuffling function,
- event handling function,
- score showing function,
- timer function,
- score uploading function.

This game was launched on August 13, 2009. Promotion activities were done by announcing the game in the mailing list such as IA-ITB, ITB, and IT-MDGT. Promotion was also done by announcing the game on Kaskus. The number of visitors on www.tombongantuk.com from August 1, 2009 until August 12, 2009, where there are no new games is released varies between 2 to 8 visitors per day. Meanwhile, with the launch of this game, the number of visitors on August 13, 2009 is 75 guests and on August 14, 2009 is 44 guests.

5. Conclusion

From the research that has been conducted, several conclusions can be obtained as follows.

- Mini game concept that can be used to train and test accuracy has been built.
- Client side and server side web based mini game functions have been built.
- This game has been launched and affected to the number of visitors of www.tombongantuk.com.

It is expected that this research may form the basis of further research. Further research is related to the development of other mini game with the following characteristics.

- The game is built using PHP and Javascript.
- The game is mini game.
- The game uses timer.

References

[1] _____, *Game Development*, http://en.wikipedia.org/wiki/Game_development, Agustus 13, 2009 13.00 WIB.

[2] ____, Game Design, http://en.wikipedia.org/wiki/Game_design, Agustus 2009 13, 13.20 WIB.

HARDWARE DESCRIPTION LANGUAGE COMPARISON SYSTEMC TO VERILOG CASE STUDY: LEAST COMMON MULTIPLE

Janita Kinsi[1], Maman Abdurohman[2]

[1,2]Informatics Faculty, Institut Teknologi Telkom
[1]jkinsi@gmail.com, [2]m_abdurohman@yahoo.com

Abstract

One phase of embedded system design is using hardware description language (HDL) for define functionality of the system. There are many HDL languages that have different characteristics between one and another. It is necessary to evaluate those characteristics between different HDL to get the convenion HDL for suitable case. In this paper we evaluate SystemC and Verilog HDL and compare they capabilites. The comparison is done in an implementation of a single-purpose processor with least common multiple (LCM) functionality, in which, compared parameters are latency, memory consumption, runtime delay, and hard disk allocation. There are some steps which should be done, LCM modeling using state-machine modeling, optimation, creating diagram block, and implementing it into SystemC and Verilog language, and finally, calculating each compared parameter. In this paper, we found several strengths and weaknesses, such as Verilog's ability in defining clock timing for controller is better than SystemC, and vice versa, ability of SystemC is better than Verilog in flexibility term.

Keywords: Systemc, LCM, Verilog, single-purpose processor

1. Introduction

In system modeling, there are several modeling that commonly used for describing hardware (especially in embedded system), those are sequential program model, communicating process model, state machine model, dataflow model, and object oriented model. One of the modeling which commonly used is state machine model. State machine model can be finite state machine (FSM) or finite state machine with datapath (FSMD). Both is used to used in defining states which may occur in system life cycle, ex. In defining single-purpose processor. For this reason, this modeling will be used for modeling single-purpose processor for less common multiple.

Once design fixed, it is simulated. For simulation, there are many languages which can be used. Those are SystemC, VHDL, Verilog, SpecC, etc. those languages may differ in some characteristics. It can be caused by their core language. Those characteristics can be memory consumption, runtime delay, file size, etc.

Thus, it is necessary needed a benchmarking among those languages so that system developer can chose which language that can be used according his needs and computer specification he has. For SystemC, there ara just few of them who campare it with other language. If this comparison exists, it is only based on user experience and not parametrically. So that, in this literature, SystemC will be compared with Verilog. To fulfill this objective, single-purpose processor for least common multiple (LCM) is needed as a study case. LCM of 2 integers is multiplication of all prime factorization of the two integers with the biggest power. Hopefully, form this research we can define strengths and weaknesses of one language to another, and what makes one language better than another. So, we need to study some literatures (syntax and modelling) then design the model using state-machine modelling, implemented in SystemC and Verilog, then measure each parameter (hard disk allocation, memory consumption, and latency & runtime delay), compared, and analyzed the causes.

2. Theory

2.1 SystemC

SystemC is a modeling language based on C++ language. The primary goal of SystemC is to enable system-level modeling. This model is modeling the systems above the register-transfer level abstraction, including systems that might be implemented in software or hardware. RTL modeling can also be performed in SystemC.

Wide range of meodels of computation and communication and methologies used in system design presents a major challenge in creating a system-level design language. To address this challenge, SystemC uses a layered approach that allows for the flexibility, higher-level construct that share an efficient simulation engine. The base layer of SystemC provides an event driven simulation kernel. This kernel works with events and processes in an abstract manner.

Modul is the basic building block for partitioning a desing. Module allows disigners to berak complex systems into smaller, more manageable, and to hide internal data representation and algorithm.

A module contains :

1. Ports, through which the module communicates with the environment.
2. Processes that describe the functionality of the module
3. Internal data and channels for maintenance of model state.
4. Hirerchically, other modules.

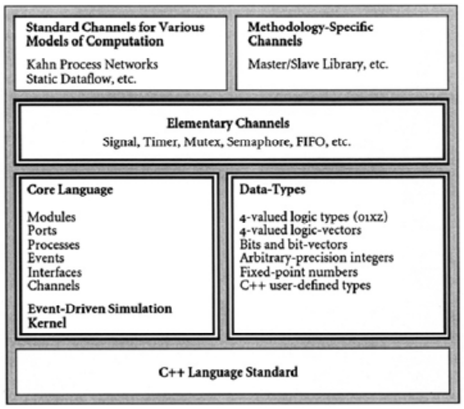


Figure 2.1. SystemC Language Architecture [2]

A module is implemented with the SC_MODULE keyword, with the body of the module enclosed in curly braces. Example :

```
SC_MODULE (Adder)
{
//ports, processes, etc
SC_CTOR (Adder){
// Process declaration,
// Sensitive list
}
}
```

An interface consists of a set of operations. It specifies only the signature of each operation, namely, the operatin's name, parameters, and return value. It neither specifies how the operations are implemented nor defines data fields, as there may be different implementations of the same interface.

SystemC provides poerts for modules to connect to and communicate with their surroundings. We represent ports by objects, not simple pointers or references. This not only makes code more readable, but also allows the SystemC kernel to manipulate them in a way that simplifies the overall syntax for the user. With port objects, we do not require that ports be bound at module instantiations time and be repeated as formal parameters to the module constructor. This benefit comes with no performance overhead (when compared against pointers or references) in any reasonable implementation.

Whereas interfaces and ports together describe what functions are available in a communications package, channels define how these functions are performed. We initiate operations through interfaces, but it is the channels that carry out these operations. Different channels may implement the same interface in different ways. A channel may implement more than one interface – provided it implements the operations specified in all of its interfaces[2].

2.2 Verilog

Verilog HDL is a hardware description language used to design and document electronic systems. An IEEE working group was established in 1993 under the Design Automation Sub-Committee to produce the IEEE Verilog standard 1364. Verilog became IEEE Standard 1364 in 1995. The IEEE working group released a revised standard in March of 2002, known as IEEE 1364-2001. Significant publication errors marred this release, and a revised version was released in 2003, known as IEEE 1364-2001 Revision C.

Verilog is a language used to describe a digital system: for example, a network switch, a microprocessor or a memory or a simple flip-flop. This just means that, by using a HDL, one can describe any (digital) hardware at any level. Here the example of D flip-flop code in Verilog [5]:

```
// D flip-flop Code
module d_ff ( d, clk, q, q_bar);
input d ,clk;
output q, q_bar;
wire d ,clk;
reg q, q_bar;

always @ (posedge clk)
begin
q <= d;
q_bar <=  ! d;
end

endmodule
```

2.3 Single-Purpose Processor

Single-purpose processor is a digital circuit which is designed to execute only one program. Basically, a processor consists of a controller and datapath. Datapath saves and manipulates system data. Datapath consists of register unit, functional unit, and interconnection unit such as wire and multiplexor. A controller defines configuration of datapath. Controller defines datapath control input like register load, and chooses input signal for multiplexor, from register unit, functional unit, and interconnection unit to make sure the configuration works properly for several interval of time. There are some steps in designing a processor, which are :

1. define sequential program
2. convert into complex state diagram
3. define functionality into controller and datapath
4. build an FSM which represents controller

3. Least Common Multiple (LCM) Design

3.1 Least Common Multiple

Least common multiple of two positive numbers *a* and *b* means a positive number with least value that can be divided by both *a* and *b*. It can be written as *lcm(a, b).*

For example : *lcm(3, 7)* = 21, *lcm(*4, 6*)* = 12, and lcm(5, 10) = 10.

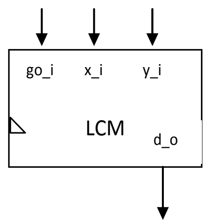


Figure 3.1. Black Box Diagram of LCM

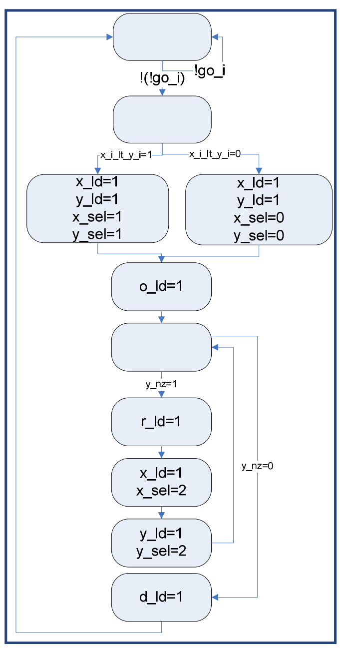


Figure 3.2. Controller State Diagram of LCM

Like greatest common divisor (GCD), *lcm* can be calculated as below.

$a = p_1 a_1 \text{ x } p_2 a_2 \text{ x } \ldots \text{ x } p_n a_n$,
$b = p_1 b_1 \text{ x } p_2 b_2 \text{ x } \ldots \text{ x } p_n b_n$

where $p_1 < p_2 < \ldots p_{n-1} < p_n$ and a_i, b_i is subset of N for $1 \leq i \leq n$,

so $lcm(a,b) = p_1 max(a_1,b_1) \text{ x } p_2 max(a_2,b_2) \text{ x} \ldots \text{x } p_n max(a_n,b_n)$

Example

$a = 60 = 2^2 \times 3^1 \times 5^1$, $b = 54 = 2^1 \times 3^3 \times 5^0$

so
$lcm(a, b) = 2^2 \times 3^3 \times 5^1 = 4 \times 27 \times 5 = 540$

while $gcd(a, b) = 2^1 \times 3^1 \times 5^0 = 2 \times 3 \times 1 = 6$

$$a \text{ x } b = gcd\ (a, b) \text{ x } lcm\ (a, b) \qquad (1)$$

3.2 Algorithm

The algorthm for LCM process show in C format, here :

```
int x, y, o, r;
while (1) {
while (!go_i);
if (x_i < y_i)        {
  x=y_i;
  y=x_i;
}
else {
  x=x_i;
  y=y_i;
}          //x must be greater than y
o=x*y;
while (y!=0) {
  r = x % y;
  x = y;
  y = r;
}
d_o=o/x;
       }
```

Above algorithm is optimized algorithm wich replace the process in while statement :

```
while (x!=y) {
 if (x<y)
 y=y-x;
 else
 x=x-y;
}
```

With this satatement :

```
while (y!=0) {
  r = x % y;
  x = y;
  y = r;
}
```

3.3 State Diagram

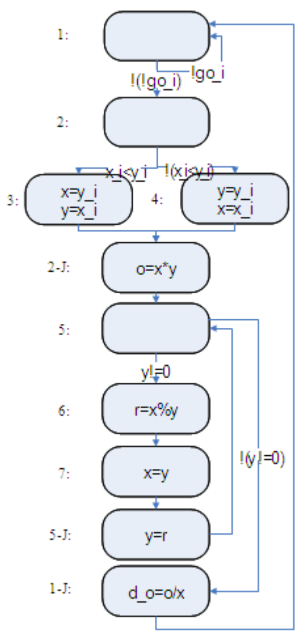
State diagram is a diagram that converted from algorithm. Figure 3.3. (a) shows the state diagram of LCM process.

3.4 Datapath

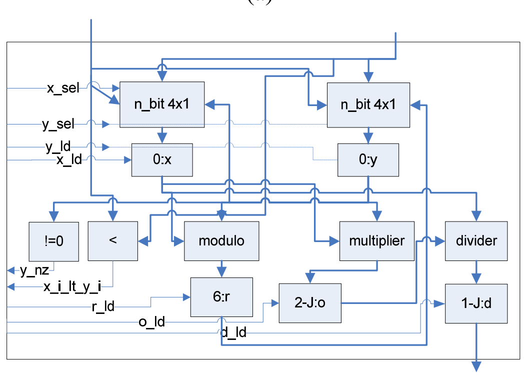
Datapath consists of 5 functional units which are slt(set less than), nz(not zero), modulo, multiplier, and divider. There are 5 registers needed, each for buffering input *x* and *y,* and modulation, multiplication and division result. Figure 3.3. (b) shows the datapath diagram of LCM.

3.5 Controller

For controller, there are 4 input signals and 7 output signals. Those four input signals are go_i as enable, clock, x_lt_y, and y_nz. Those seven output signals are 5 signals to write to each register and 2 signals to select multiplexor input. Figure 3.1.



(a)



(b)

Figure 3.3. (a) State Diagram and (b) Datapath of LCM

4. Implementation and Analysis

Tools that used in our implementation are :

- SystemC-2.2.0 : make library for SystemC
- Emacs : text editor for Verilog
- Gedit : text editor for SystemC
- GTKWave : VCD viewer
- Icarus Verilog : compiler for Verilog

Testing scenario consists of white box testing for each module, memory consumption measurement with command *top,* hard disk allocation measurement, and latency & runtime delay measurement. The results of each measurement are shown below.

4.1 Memory Consumption

Memory consumption is memory usage when two application run.

Table 4.1. Memory Consumption

	No	Verilog					SystemC				
		Virtual	Resident	Share	% Mem	Time	Virtual	Resident	Share	%Mem	Time
Ketika programjalan	1	34272	12MB	7684	2,6	0:29.99	34272	12MB	7684	2,6	0:29.57
	2	34272	12MB	7684	2,6	0:30.10	34272	12MB	7684	2,6	0:29.58
	3	34272	12MB	7684	2,6	0:30.20	34272	12MB	7684	2,6	0:29.59
	4	34272	12MB	7684	2,6	0:30.32	34272	12MB	7684	2,6	0:29.60
	5	34272	12MB	7684	2,6	0:30.42	34272	12MB	7684	2,6	0:29.61
	6	34272	12MB	7684	2,6	0:30.49	34272	12MB	7684	2,6	0:29.62
	7	34272	12MB	7684	2,6	0:30.53	34272	12MB	7684	2,6	0:29.64
	8	34272	12MB	7684	2,6	0:30.64	34272	12MB	7684	2,6	0:29.65
	9	34272	12MB	7684	2,6	0:30.74	34272	12MB	7684	2,6	0:29.66
	10	34272	12MB	7684	2,6	0:30.85	34272	12MB	7684	2,6	0:29.67
Sebelum program jalan	1	34272	12MB	7692	2,6	0:28.76	34272	12MB	7684	2,6	0:28.76
	2	34272	12MB	7692	2,6	0:28.85	34272	12MB	7684	2,6	0:28.85
	3	34272	12MB	7692	2,6	0:28.91	34272	12MB	7684	2,6	0:28.91
	4	34272	12MB	7692	2,6	0:28.96	34272	12MB	7684	2,6	0:28.96
	5	34272	12MB	7692	2,6	0:29.04	34272	12MB	7684	2,6	0:29.04
	6	34272	12MB	7692	2,6	0:29.11	34272	12MB	7684	2,6	0:29.11
	7	34272	12MB	7692	2,6	0:29.19	34272	12MB	7684	2,6	0:29.19
	8	34272	12MB	7692	2,6	0:29.26	34272	12MB	7684	2,6	0:29.26
	9	34272	12MB	7692	2,6	0:29.32	34272	12MB	7684	2,6	0:29.32
	10	34272	12MB	7692	2,6	0:29.37	34272	12MB	7684	2,6	0:29.37

Table 4.1. shows that there in increasing shared memory for Verilog when program is running. For SystemC there is no significant increasing for virutal, resident, and shared memory. This is becouse of on Linux Ubuntu the execution done on the consol.

4.2 Harddisk Allocation

Harddisk alloaction is the size of executable file of verilog and SystemC

Table 4.2. Harddisk Allocation

No	Modul	SystemC			Verilog	
		Precompiled	Compiled	Executabe	Precompiled	Executable
1	Controller	4,43 KB	168 KB	1,04 MB	271 KB	8,19 KB
2	Divider	1,96 KB	118 KB	0,98 MB	923 B	4,56 KB
3	LCM	8 KB	221 KB	1,08 MB	5,35 KB	24,9 KB
4	Modulo	1,95 KB	115 KB	0,98 MB	769 B	2,43 KB
5	Multiplier	1,95 KB	115 KB	0,97 MB	776 B	2,61 KB
6	MUX	2,61 KB	112 KB	0,98 MB	1,04 KB	3,36 KB
7	Nz	1,52 KB	138 KB	1,01 MB	444 B	1,07 KB
8	Register	1,7 KB	160 KB	1,04 MB	962 B	2,19 KB
9	Slt	1,89 KB	153 KB	1,01 MB	666 B	1,85 KB

Table 4.2. shows that SystemC has the bigger size of file than verilog. This is because of linking library C++ and SystemC when program is running and shell programming instruction.

4.3 Runtime delay

Runtime delay is amount of clock needed to execute all program.

Table 4.3. Runtime Delay

No	X_i	Y_i	LCM	SystemC(ns)		Verilog (s)	
				Runtime Delay	Latency	Runtime Delay	Latency
1	0	0	0	6	10	6	12
2	0	1	0	6	10	6	12
3	1	0	0	6	10	6	12
4	1	1	1	10	18	10	20
5	2	15	30	14	26	14	28
6	15	2	30	14	26	14	28
7	3	15	15	10	18	10	20
8	4	4	4	10	18	10	20
9	15	4	30	14	26	14	28
10	10	5	10	10	18	10	20

Table 4.3. shows that runtime delay of Verilog the same as SystemC. We conclude that run time delay is independent to HDL language.

5. Conclusion

Based on those three tables, we figure out that to define RTL-based system, Verilog show natural behavior of RTL better. Furthermore, runtime delay does not depend on language. Hard disk allocation for SystemC's executable and compiled files greater than Verilog's ones. It is because linking process which is done in compilation based on the rule which is defined in *Makefile*. Memory consumption for both of them is low, so it is easy to be implemented in more limited source.

References:

[1] Mahesh Prasad N, 2002, “SystemC: An Introduction for beginners”.

[2] Thorsten Grotker, 2002, “System Design with SystemC”, Kluwer Academic Publishers : Hingham, MA, USA.

[3] Vahid, Frank & Tony Givargis, 2002, “Embedded System Design : A Unified Hardware/Software Introduction”.

[4] Vahid, Frank, Daniel D. Gajski, Sanjiv Narayan & Jie Gong. 1994. “Specification and Design of Embedded Systems”. Upper Saddle River : Prentice Hall.

[5] Tala, Deepak Kumar. 2009. “Verification IP’s - Introduction”. http://www.asic-world.com/verilog.

[6] ____. “http://en.wikipedia.org/wiki/SystemC”. Date of download : January, 19th 2009.

[7] ____. “http://en.wikipedia.org/wiki/VHDL”. Date of download : January, 14th 2009.

[8] ____. “http://radar.ee.itb.ac.id/~suksmono /Lectures/el2009/ppt/2. Algoritma Kompl eksitas dan Teori Bilangan.pdf”. Date of download : February, 3rd 2009.

[9] ____. “http://www.asic- world.com/verilog/history.html”. Date of download : January, 19th 2009.

[10] ____. “http://www.deepchip.com/items/0373-02.html”. Date of download : February, 1rd 2009.

[11] ____. “http://www.doulos.com/knowhow/verilog_de signers_guide”. Date of download : September, 28th 2009.

[12] http://www.kdedevelopers.org/node/1445, Date of download : January, 18th 2009.

HANLE EEFFECT MODELING ON SILICON BASED SPINTRONIC SEMICONDUCTOR DEVICES

M. Ramdlan Kirom [1], Dudi Darmawan[2]

[1,2]Engineering Physics Study Program, Institut Teknologi Telkom
[1]mrk@ittelkom.ac.id, [2]ddw@ittelkom.ac.id

Abstract

Spintronic semiconductor devices are the semiconductor devices with electron's spin flows as information carrier, it is not electron or hole flows. Spintronics devices have superiority like as less power consumption and higher speed than conventional semiconductor devices. Spin flows in semiconductor from source to detector. The effects that will be happen at spin flow are spin valve and Hanle effect. In our research, only Hanle effect that has analyzed. The analysis of Hanle effect that is done by using spin transport diffusion model in semiconductor material to see dephasing phenomenon that is caused by external magnetic field. Furthermore it is done source – detector range influence, relaxation time, and drift velocity toward the voltage of source and detector along spin flow. Simulation result of diffusion model is likes measurement of V – B result at Fisika Material dan Elektronik Laboratory, Physics Department, ITB.

Keywords: Spintronic Semiconductor, electron spin, Hanle effect, diffusion model

1. Introduction

An electron that was identified as a particle in 1897 by J. J. Thomson is subatomic particle which carries a negative electronic charge. As charge carrier, electrons move freely in metal and semiconductor that was influenced by external magnetic field and Lorentz force. Zeeman effect observation in 1896 and hyperfine spectrum structure from atomic spectrum lines that first time was founded by Uhlenbeck and Goudsmith in 1925 give evidence that electron has "spin" with intrinsic angle momentum (S) and was coupled with magnetic moment [1].

In this paper, it was investigated theoritically about spin current as information carrier in spintronic devices. Spin will flow that caused electron spin alignment at semiconductor materials from one side to other side without electron movement. This is contrary with electronic devices that charge current as information carrier.

Since spin detection is a projection of the electron spin direction on a fixed measurement axis, the change in output of a spintronics device employing precession will be determined by cos θ. The angle $\theta = \omega\tau$ is spin angle, ω is spin precession frequency and τ is transit time from injector to detector over transit length L. Transit-time uncertainty $\Delta\tau$ gives rise to precession angle uncertainty $\Delta\theta = \omega\Delta\tau$ ($\omega = g\mu BB/\hbar$, where g is electron spin g-factor, μB is the Bohr magneton, $\hbar$ is the reduced Planck's constant, and B is perpendicular magnetic field) that will increase by increasing the magnetic field and causes dephasing or Hanle effect [2,3,5,7,15]. A complete model for dephasing during transport from injector to detector is therefore needed. Electron transport in a semiconductor is well-described by the drift-diffusion equation. To model spin transport accurately, finite spin-lifetime must be included within the relaxation-time approximation. Electron spin value on reference direction was depend on spin.

In quantum theory was said that spin quantum number of the electron has only two values, -1/2 and +1/2, or spin up and spin down. It is mean that Hanle effect cause change of spin direction, spin up to be spin down or spin down to be spin up, when if $\Delta\theta$ is bigger than 90^o. Hence, we can control spin direction along semiconductor materials by varying magnetic field.

2. Spin Model

The spin that was injected from ferromagnetic materials like Fe, Co, Ni, or dilute magnetic semiconductor (DMS) [2,12,16] as a source to nonmagnetic medium will cange spins polarization in medium till they align themselves along medium.

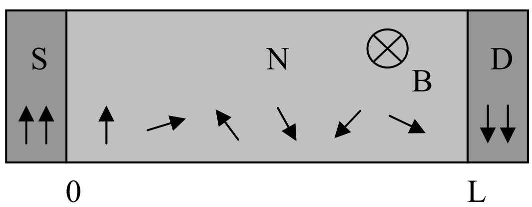


Figure 1. **Spin flows on nonmagnetik medium (N) by spin injections from source (S) to detector (D). B is external magnetic field**

In the model that would be used, electrons spin interaction, spin-orbit interaction, and exchange interaction were ignored. Spin flows was affected by diffusion process, drift velocity, and relaxation time. Therefore, we can construct the model of spin dephasing along medium with differential equation [3, 7, 13] :

$$\frac{\partial s}{\partial t} = D\frac{\partial^2 s}{\partial x^2} - v\frac{\partial s}{\partial x} - \frac{s}{\tau_{sf}} \qquad (1)$$

With s is spin density, D is diffusion constant, v is drift velocity, and τ_{sf} is relaxation time. The first term of right part in equation (1) is spin diffusion in medium, the second term is electron movement because of electric field, and third term is spin relaxation time back to original position. The equation (1) can simplify as :

$$\frac{\partial f}{\partial t} = D\frac{\partial^2 f}{\partial x^2} - v\frac{\partial f}{\partial x} \qquad (2)$$

$$s(x,t) = f(x,t)\, e^{-\frac{t}{\tau_{sf}}} \qquad (3)$$

It will be valid initial condition $f(x,0) = h(x) = \delta(x)$ which mean at $t = 0$ the spin is only in $x = 0$. The solution of above equation can be found by using Green fuction [14] and we get :

$$s(x,t) = \frac{1}{2\sqrt{\pi D t}} e^{-\frac{(x-vt)^2}{4Dt}} e^{-\frac{t}{\tau_{sf}}} \qquad (4)$$

Electron spin in detector only has two possible state, spin up ↑ and spin down ↓. It is spin up when if spin direction is in the same direction with reference and spin down when if spin direction is the opposite direction with reference. The quantities $n_\uparrow$ is density of spin up electron and dan $n_\downarrow$ is density of spin down electron menyatakan elektron spin down persatuan volume. The difference of $n_\uparrow$ and $n_\downarrow$ is :

$$n(x) = n_\uparrow - n_\downarrow = \frac{N}{V}\rho(x) \qquad (5)$$

with N is number of electrons in medium, V is medium volum, and $\rho(x)$ is density of spin probability at x for spin injection time interval T. So we can express $\rho(x)$ in :

$$\rho(x) = \int_0^T \frac{1}{2\sqrt{\pi D t}} e^{-\frac{(x-vt)^2}{4Dt}} e^{-\frac{t}{\tau_{sf}}} \cos\omega t \, dt \qquad (6)$$

Addition of cos ωt is projection of spin direction into reference. Time parameter of spin in ns, so for 1 second is too long. Hence we can approach that T close to infinity. Position at $x = 0$ is spin state in source-medium interface and $x = L$ is spin state in medium detector interface. For simplicity we can consider every constants are 1. From equation (5) and equation (6) we can draw graph of spin density as magnetic field B at detector $x = L$. For assuming $v = 0$, we get :

$$n(\omega) = \int_0^\infty \frac{1}{\sqrt{t}} e^{-\frac{L^2}{t}} e^{-t} \cos\omega t \, dt \qquad (7)$$

3. Simulation Result

Density of spin $s(x)$ for various t = 0.5; 1; 2 can be seen at fig. 2. On initial condition, spin only exist at $x = 0$. By increasing time the curve will be broaden, but density of spin will decrease. This conditions descibe spin diffusion in x-direction.

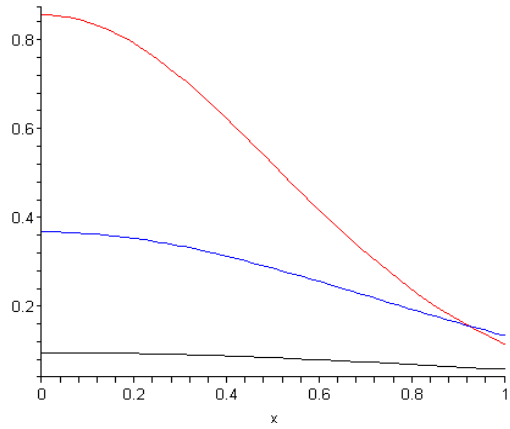


Figure 2. **Density of spin *n* along *x*. By increasing time the curve will be broaden**

The density of spin n which as function of frequncy ω for variation of L is shown in pig. 3. The value of $n = n_\uparrow - n_\downarrow$ can positive or negative that is depend on frequency or applied magnetic field. When n is positive, it is mean that spin up electrons are majority. Otherwise, When n is negative, it is mean that spin down electrons are majority. From the fig. 3 we can conclude when if medium width L was enlarged, then detector will receive less spin. It is defined B_o as external magnetic field when n equals zero. On fig. 3 we see B_o decreased when L increased.

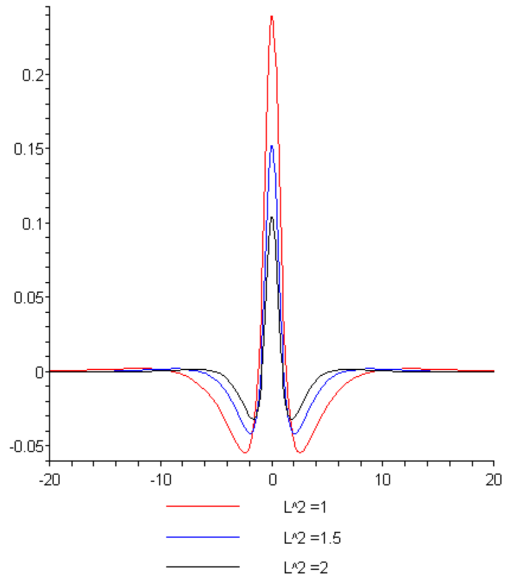


Figure 3. **Spin density $n(\omega)$ as function of frequncy ω for various *L***

The spin density n which as function of frequncy ω for variation of relaxation time τ_{sf} is shown in pig. 4. If relaxation time τ_{sf} will enlarge, then B_o will too. Otherwise, number of spin will decrease that is caused by direction of spin orientation is more random.

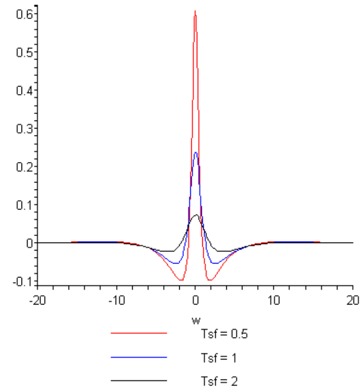


Figure 4. Spin density $n(\omega)$ as function of frequncy ω for various τ_{sf}

Influence of electron movement in the spin flow can be seen at pig. 5. Calculation result equation (7) with drift velocity v = 0,4 which was compared with zero drift velocity can be concluded that increase number of spin. It is interesting to study more detail to prove conductivity influence in spin flows.

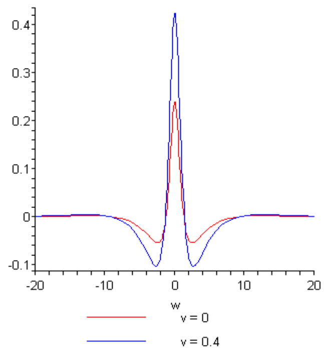


Figure 5. Spin density $n(\omega)$ as function of frequncy ω for drift velocity v

From this simulation we wish information of spin orientation that was received by detector. Output voltage at detector depend on majority of spin [9]. If spin up is majority, then output voltage is positive. However, If spin down is majority, then output voltage is negative. It is consistent with observation result at Fismatel Lab. ITB on pig. 6 below.

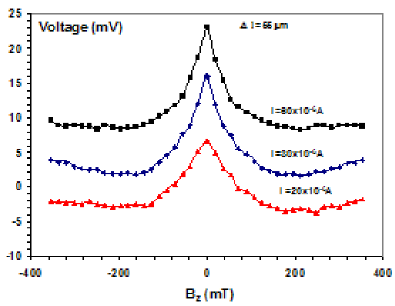


Figure 6. Observation result of output voltage versus external magnetic field for various spin up injection current

Some parameters of silicon has been obtained by Lou, et al [17] which are consist of spin diffusion constant D = 640 cm^2/s; Electron spin g-factor = 0.44; spin time relaxation $\tau_{sf} = 7.56 \times 10^{-11}$ s. If those parameters take into equation 7, then we obtain calculation result of output voltage versus external magnetic field. Fig. 7 show comparison calculation result and observation result of output voltage versus external magnetic field. For small B_z (<100 mT) these curves coincide, but for B_z > 100 mT they separate. It is probably caused by interspin interaction, spin-orbital interaction, spin-impurity interaction, or spin-lattice interaction.

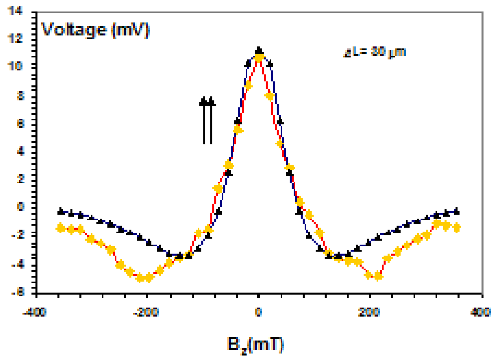


Figure 7. Output voltage versus external magnetic field. Red line is observation result, and blue line is calculation result

4. Conclusion

There is Hanle effect in semiconductor medium that was caused by perpendicular magnetic field. Spin up or spin down can be obtained at detector that was depend on magnetic field. Medium width L, relaxation time τ_{sf}, and ohmic properties will affect number of spins which was observed at detector.

Hanle effect can be expalined theoritically by using diffusion equation. However, the result isn't exactly close with observation result, especially for magnetic field $B_\perp$ > 100 mT. It is caused by ignoring spin interaction with surroundings like another spin, impurity, etc. For further research, we need concept of quantum mechanics to analize influence of spin interaction with surroundings.

Acknowledgements

Authors thank to Edi Suprianto for giving data of observation result at Fismatel Lab. ITB. This work support by IT Telkom Internal Research Grant.

References

[1] A. Beiser, (1983), *Konsep Fisika Modern*, Erlangga

[2] A. D. Giddings, (2008), *Dilute Magnetic Semiconductor Nanostructures*, thesis for PhD degree, University of Nottingham

[3] B. Huang, I. Appelbaum, (2008), *Spin Dephasing in Drift-Dominated Semiconductor Spintronics Devices*, Material Science, eprint: arXiv:0801.1790

[4] B. Huang, H. J. Jang, I. Appelbaum, (2008), *Geometric dephasing-limited Hanle effect in long-distance lateral silicon spin transport devices*, Material Science, eprint: arXiv:0809.2225v1

[5] J. Li, B. Huang, I. Appelbaum, (2008), *Oblique Hanle effect in semiconductor spin transport devices*, Applied Physics Letters, 92, 142507

[6] J. Xiao, A. Zangwill, M. D. Stiles, (2007), *A numerical method to solve the Boltzmann equation for a spin valve*, eprint: arXiv:0701387v1

[7] M Furis, D L Smith, S Kos, (2007), *Local Hanle-effect studies of spin drift and diffusion in n:GaAs epilayers and spin-transport devices*, New Journal of Physics, 9, 347

[8] M. I. Miah, E. M. Gray, (2007), *Spin transport and spin current detection in semiconductors*, Solid State Sciences, 10, pp 205 – 210

[9] F. J. Jedema, (2002), *Electrical Spin Injection in metallic Mesoscopic Spin Valves*, thesis for Doctor degree, Solid State Physics Laboratory of the University of Groningen

[10] F. Mireles, G. Kirczenow, (2002), *From classical to quantum spintronics : Theory of coherent spin injection and spin valve phenomena*, eprint: arXiv:0206009

[11] F. Assad, K. Banoo, M. Lundstrom, (1998), *The Drift-Diffusion Equation*, Solid-State Electronics, Vol. 42, No. 3, pp 283-295

[12] H. Ohno, (1999), *Ferromagnetic III-V Semiconductor, (Ga,Mn)As*, J. of Magnetism & Magnetic Mat., 200, 110

[13] Jasprit Singh, (1993), *Physics of Semiconductors and Their Heterostructures*, McGraw-Hill, International Ed., pp 324 – 356

[14] Q. H. Qin, (2007), *Green's function and boundary elements of multifield materials*, Elsevier Science

[15] S. Luryi, J. Xu, A. Zaslavsky, (2007), *Future Trends In Microelectronics*, Wiley-Interscience Publication, pp 28 – 40

[16] Y. B. Xu, S. M. Thompson, (2007), *Series in Materials Science and Engineering : Spintronic Materials and Technology*, Taylor & Francis, pp 4 - 35

[17] X. Lou, C. Adelmann, S.A. Crooker, E. S. Garlid,(2007), *Electrical Detection of Spin Transport in Lateral Ferromagnet-Semiconductor Devices*, Cond-Mat, eprint: arXiv:0701021

PON EXTENDER : CONTROLLING SYSTEM FOR GPON WITH OPTIONAL EDFA

D. Andriyanto[1], D.Tarsono[2], N. Samsuri[3], A. Ahmad[4], R. Mohamad[5]
[1,2,3,4,5]Photonic Technology Program, TM Research & Development Sdn. Bhd.
[1]drees@tmrnd.com.my, [2]dedy@tmrnd.com.my, [3]norhakimah@tmrnd.com.my, [4]anim@tmrnd.com.my, [5]romli@tmrnd.com.my

Abstract

In this paper, we proposed and developed PON Extender Controlling System for GPON (Gigabit Passive Optical Network) with optional EDFA. The amplifier for physical link performance and network traffic evaluation have been demonstrated experimentally. A laboratory setup has been made with any wavelength input signal for maintenance the automatic control output power and Gain. The OLT (Optical Line Terminal) is used to inject traffic data and signal to the system. The results will prove the ability of the developed GPON Extender system to deliver services to the subscribers.

Keywords: PON Extender

1. Introduction

Finding the right technology to cover the last few miles of any network has always been a problematic challenge for service providers. Whether wireless, copper or optical links are to be used, there is inevitably a set of competing technologies to choose from. Finding the optimum solution can be a complex process, with numerous interlinked factors to be taken into account. In this context, developments are underway in the world's market around the use of high speed Passive Optical Networks (PONs) which is going to be particularly significant - especially as far as the role of Gigabit PON (GPON) systems are concerned. A Gigabit Passive Optical Network (GPON) technology which mean there are no active elements in the field and, thus, powering and heating issues are eliminated. The system also has fewer network components than an active fiber deployment, which yields lower maintenance costs and fewer potential points of failure. The newly standardized ITU-T G.984 GPON technology, is delivering extremely high bit rates while still supporting the transmission of native formats such as IP and TDM at extremely high levels of efficiency. Due to its unprecedented offered bandwidth, GPON is the ideal technology for large-scale FTTH applications where multiple end-users are requiring an ever-growing bandwidth. Moreover, in areas populated by both business and residential customers, GPON is the most cost-effective solution.

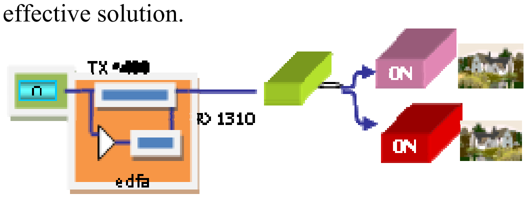


Figure 1. Proposed PON EXTENDER System

The diagram above shows the proposed system architecture and the same design could be used for commercial implementation in the access network area. Four customized transceivers with 1490/1310nm, 1550nm and 1625nm wavelength are multiplexed in one fiber transmission (bidirectional). The 1490nm wavelength is used as the downstream signal from the OLT and 1310nm is used as the upstream data from the ONU. The 1550nm is reserved for CATV application and 1625nm is useful for monitoring. In this paper, the experiment on the physical layer performance with the customized control for maintenance the signal output in the link transmission.

2. Test Bed Set up

In this experiment, the setup is similar to the proposed architecture in Figure 1 the Optical Line Termination (OLT) is used to generate the traffic data. The signal from OLT and Optical Network Unit (ONU) are placed at transmit and receive side of the access link to capture the result of the bidirectional data transmission through the single mode fiber (SMF) spool link. The measurement on output power and gain can be captured and analyzed in order to prove that the link setup is working as expected in the PON environment.

3. Description and Figure

The design uses an optical switch 1x2 for protection of the transmission line. The AWG acts as the multiplexer and demultiplexer to guide the signal from the different wavelengths into the dedicated SOAs. Each SOA of 1310nm, 1490nm, 1550nm and 1625nm is attached to an individual electronic module. All four units of the individual modules is then connected to the master card for control and monitoring of every module. Either a LCD or a GUI on a PC can be used as user interface and remote the system like NMS use TCP/IP or other will be an added feature to ease monitoring, controlling and maintenance of the PON extender.

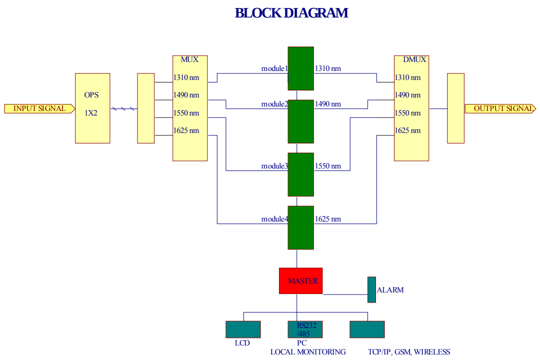


Figure 2. Block Diagram for PON Controlling System

Figure 3 show the parts of graphical user interface (GUI) for the system. The input power from signal use coupler 95/5 (%) for monitoring signal. The PD1 (Photodiode input) and PD2 (Photodiode Output) reading the data and monitoring the signal input and output. Both PDs are embedded into the electronic module of the SOA. Driver Laser can be running automatically for maintained the gain and power output .The modes of operation that being offered by the controlling system are as the following; Current Mode (CM), Automatic Power Control (APC), and Automatic Gain Control (AGC). Each mode can be selected depending on the requirement of the link loss characteristics.

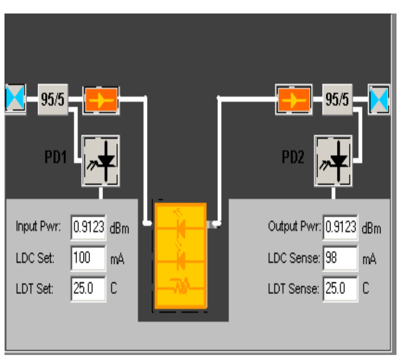


Figure 3. Block Diagram for Client Module

Figure 4 shows that the transmitted signal from the ONU which is slowly attenuated by the fiber loss can be amplified successfully by the PON extender controlling system. The optical signal of -21dBm is increased to -1.45dBm after being driven with 200mA of current. The controlling system is able to handle a maximum current increase of up to 500mA.

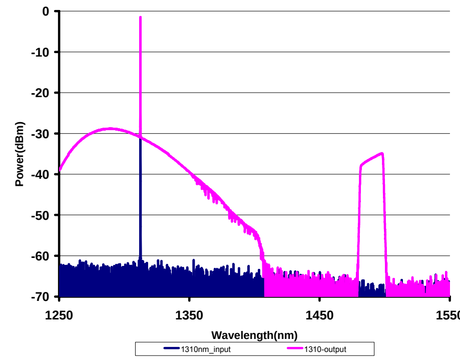


Figure 4. Input and Output of PON Extender

4. Conclusion

The PON extender controlling system is able to amplify the degraded signal from both ONU and OLT using 3 different modes. With the existence of the PON extender, the transmission distance of the OLT can reach more than 20km but is restricted to a maximum of up to 60km. More users can be introduced into the PON system as the splitting ratio can be increase to 64 and 128.

References

[1] K. Suzuki, Y. Fukada, D. Nesset, and R. Davey, "Amplified gigabit PON systems [Invited]," J. Opt. Netw. **6**, 422-433 (2007).

[2] Tien-Tsorng Shih, Pei-Hao Tseng, Tai-Wei Wu, Min-Ching Lin, Hidenori Taga and Wood-Hi Cheng, “A 40Gb/s Bidirectional CWDM-PON.

BIDIRECTIONAL OPTICAL ADD/DROP MULTIPLEXER (OADM) FOR CWDM PON SYSTEM

Sahrul Hilmi[1], Norbaizura[2], Farid[3], Saifulbahri[4], Fazli[5], M.Rizal[6], Abu Bakar[7]

[1,2,3,4,5,6]Access Network Technology Program, TM R&D Sdn Bhd.
[7]Photonics Technology Centre, Universiti Teknology Malaysia.
[1]hilmi@tmrnd.com.my, [7]bakar@utm.my

Abstract

One of the most important components in Wavelength Division Multiplexing (WDM) distribution network is the passive bidirectional optical add/drop multiplexing (OADM). Results presented so far only use OADM in ring topology but not for passive optical network (PON) system and therefore not suitable to distribution network. This paper is to propose and demonstrate a novel bidirectional OADM for CWDM-PON system, which is passive, offers low cost, low insertion loss and protection function in a distribution network with only one fiber. By adding the proposed OADM in a WDM PON architecture, it has the capability to virtually limitlessly scale the network with minimal or no alteration to the existing network. Service providers could have the flexibility to introduce new services while maintaining the current legacy services since several logical light-tree shares the same physical fiber. The use of Fiber Bragg Grating (FBG) filters in the OADM design provides flexibility in CWDM channel selection in the same casing design. Nevertheless the OADM in the network posses an extra 6.5dB power penalty but still error free transmission had achieved with a 30 km passive optical link at 2.5 Gbps.

Keywords: OADM, PON, WDM, Fiber Bragg Grating, FTTH

1. Introduction

Among all possible wire line access technologies, a PON provides an interesting option to operators to deliver very high-speed broadband access in a cost effective way. For next-generation PON systems, targets include the capturing of FTTH subscribers such as high-end users, business users and multifamily-housing users and the provision of bandwidth-guaranteed services and high-definition video services among others. To enhance the speed and capacity, focus is being placed on the research and development of next-generation PON systems such as Wavelength Division Multiplexing (WDM) PON systems. To accommodate such users and provide such services, WDM-PON systems using WDM technology that allow each user to occupy the bandwidth of a single wave using TDM (Time Division Multiplexing) technology, which increases the transmission rate from the current 1 Gbps/s to 10 Gbps/s, hold promise [1,2].

A basic FTTH-PON structure will consists of Optical Line Terminal (OLT) at Center Office (CO), an optical power splitter or a wavelength splitter at the distribution point and numbers of Optical Network Unit (ONU) as the customer premise equipment (CPE). PONs are usually recommended to operate at the range of 20 km from the OLT. For a high capacity PON (of the order of hundreds or thousands of users), it is not feasible to have one split point or remote node from which all the wavelengths are distributed to the ONUs. Therefore, multistage is a practical approach to cater to large coverage areas and to reduce the cost of fiber installation by maximization of the multiplexed signal fiber length. As shown in Figure 1, the best compromise point between power loss brought by power splitting and data rate to be served to the user are among 16, 32 or 64 split ratio.

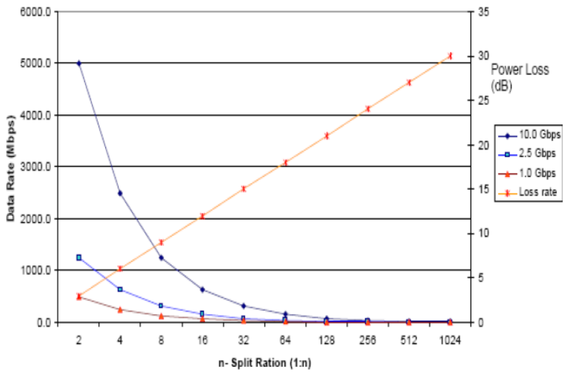


Figure 1. High Capacity and High Speed. Data Rate per User At 1.25, 2.5 And 10 Gbps as A Function of A Split Ratio at The DP

There are several techniques to upgrade capacity of a deployed passive optical network (PON) system, and the most promising options include line-rate enhancement and WDM. WDM-PON offers a variety of advantages such as the following:

(1) Permits the shared use of trunk fibers from CO to CPE (each up to 50 km long) for reducing costs and eliminating the problem of fiber shortage.
(2) Economization through the simplification of equipment installed at accommodation stations.

(3) Since each subscriber occupies one wave for downstream traffic and another wave for upstream traffic, bandwidth can be guaranteed.

Even though various topologies had been presented, most WDM-PON architectures [3,4], only allow wavelength channel separation at the AWG router, which then goes onto its individual output ports or creating multiple WDM-based light tree PONs on the same physical fiber. To the place where the population density is very high, it is difficult to re-configure the PON network layout once the fiber been installed in the field, as it will cause high operating cost. If need arises to provide another PON in between OLT and the AWG router, laying new fiber and installing new WDM-PON access network is the first thing the network operator would like to avoid.

This paper proposes a novel Dynamic WDM-PON architecture by adding a PON OADM which specifically achieve cost effective, highly scalable multiple PON configuration that allows smooth migration path and bandwidth management. It can be used to serve users in order of hundreds/thousands, and able to cover 30 km in radius that cannot be served from one split point. We also like to report a new PON OADM concept, which consist of passive optical components and a new version of current OADM normally used in Metro ring and ring based PON topologies [5].

2. Novel WDM FTTH PON

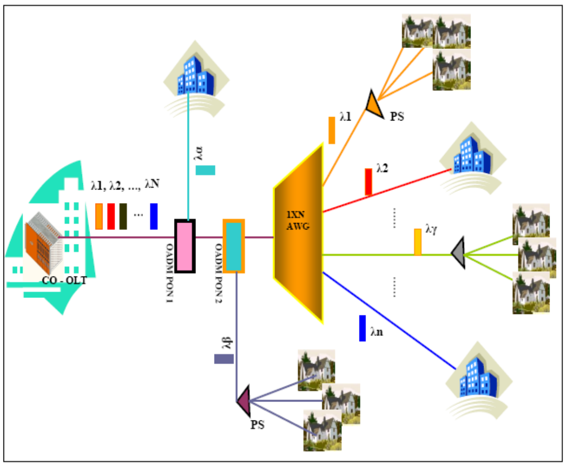


Figure 2. Novel FTTH-PON Architecture

The proposed FTTH PON is shown in Figure 2. It is a symmetric multistage architecture based upon the coarse WDM (CWDM) FTTH-PON since the technology offer a network configuration using devices with lower costs compared to dense WDM (DWDM). The novelty of the proposed architecture lies in the usage of the new design PON OADM. The Optical signals at different wavelength from separate transmitter or OLT are combined in a multiplexer and travel through the same fiber from the multiplexer to an OADM that then route one wavelength to create new PON tree or individual output port. The advantage of this new PON OADM is, a wavelength transmitted back (add) into the PON OADM will be routed to the original OLT as shown in Figure 3.

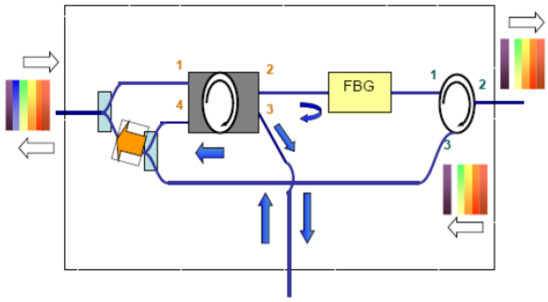


Figure 3. Configuration of New PON OADM

Depending on the type of Fiber Bragg Grating FBG) used, multiple wavelength can be routed to an OADM to have more wavelength per PON (see Figure 4). By having more than one light tree in the same physical fiber link, service providers have the flexibility to introduce new services (e.g. faster bit rate or new content) with minimal or no interruption to the legacy systems that has already been deployed.

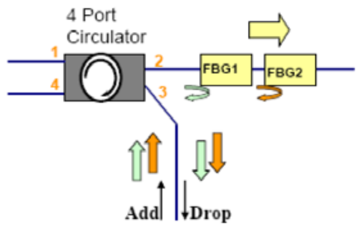


Figure 4. Multiple Wavelength Add/Drop Features

The infinite number of wavelengths in a single physical PON allows infinite scalability in the future. On the other hand, initially the upstream traffic from users will be aggregated by the ONUs and transmitted upstream towards the OLT via a TDMA technique. Initial upstream traffic will be transmitted via the time domain multiple access (TDMA) technique but as the amount of upstream traffic increases, an extra WDM TDMA link can be created using another wavelength that falls on the free spectral range. The ability to create a new light tree on the existing network allows the possibility to introduce new services while maintaining the legacy network users (i.e. low speed home users vs. high-speed business users). Use of this device makes it possible to freely add or drop signals with arbitrary wavelengths over multiplexed optical signals by assigning a wavelength to each destination.

The proposed OADM also can be used as a protection interface solution. In common network protection techniques, its implementing duplicate network resources such as fiber links, optical network units (ONUs), or array waveguide gratings (AWGs), etc., to provide network resource redundancy or the use of automatic protection switching to reroute the affected data traffic through an alternative protection path. In real optical network deployment especially for PON architecture scenario, duplication could be a good approach for mainstream fiber link. For example, links between CO to the remote node (RN) require duplication because the fiber length is more than 5 km.

However, towards user premises the same approach is not preferable. In fact, not all user premises require protection. It could be very selective among them. In the protection architecture proposed in Figure 5, additional fiber link possibly to protect intended link by using OADM proposed in Figure 3. By maintaining the original fiber connections, the original link can be considered as main / primary path, while the protection line can act as a secondary path, in order to sustain all services to the user premises. The applications not only at main fiber between the two AWG (Mux and de-Mux), but it also can be apply at the second or the next stages.

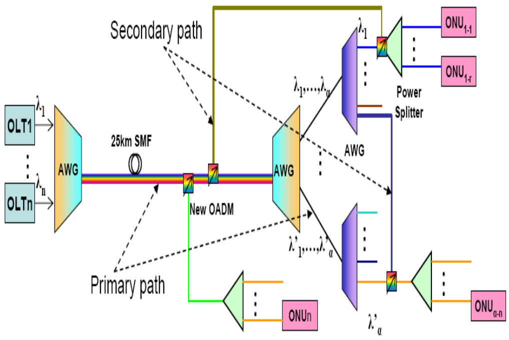


Figure 5. Proposed Protection Architecture for Multistage WDM-PON Access Topology

3. Experimental Demonstration

The experiment to demonstrate the downlink / uplink transmission incorporating the proposed OADM optical interface at the DP side is shown in Figure 6. In this experiment, two stages of 1.25Gbps and 2.5Gbps symmetric data version have been developed, using four line-up C-band transceiver (DFB lasers) in range of 1531nm ~ 1591nm, 20nm apart (approximately 2500 GHz spacing) based on CWDM standard. The output of lasers are varies from – 1 dBm ~ -5 dBm multiplexed together with and AWG Mux. To verify the performance of the newly proposed architecture, an optical transmission test was carried out (Figure 7). The Bit Error Rate Tester (BERT) is set to transmit a Non-Return-to-Zero (NRZ) 2^7-1 Pseudo Random Bit Stream (PRBS) at 1.25 Gbps and 2.5Gbps.

In order to separate transmit and receive channels at the transceiver side (considered as OLT), a 3-port circulator is used, that make an additional loss (about 1dB) compare to the original concept. The downlink channels, after traveling 10 km in single mode fiber (SMF) from OLT, are divided in the proposed OADM. Here, drop channel is directed out to a new brunch of PON, while remaining channels are routed by AWG de-Mux to it respective fiber and then split into 16 identical channels by use of a 1x16 optical power splitter before terminated at ONU.

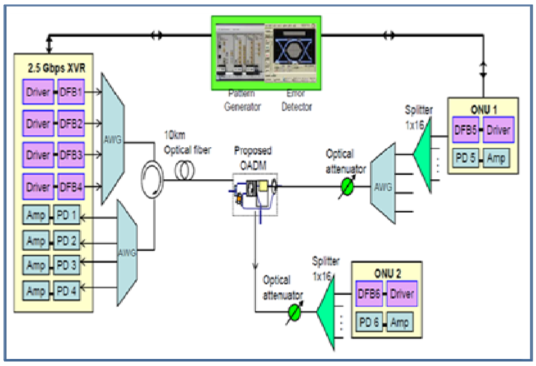
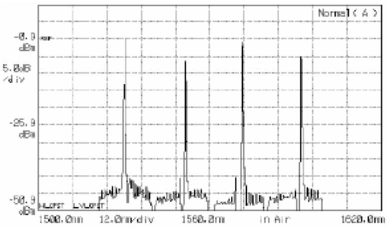


Figure 6. Experiment Setup to Demonstrate The Multistage CWDM / TDM PON Architecture.

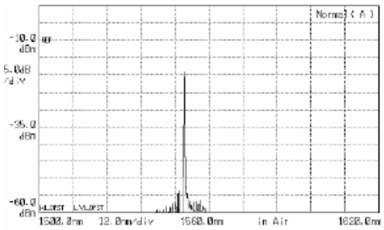
In the uplink direction, the data channels are fed back into fiber from each ONUs to its respective splitter, and then all the uplink channels are multiplexed through the AWG Mux and fed into the main link fiber. On reaching OLT, the uplink channels are again separated from downlink channels by a 3-port circulator in the OADM. An uplink channel from a brunch of PON in the intermediate node is added in the OADM, then directed to port four of 4-port circulator and combined with other channels through C-band WDM coupler. All uplink channels are again separated from downlink channels by used of a 3-port optical circulator. And then it all fed to the respective optical receiver after de-multiplexed by AWG de-Mux.

To investigate the performance of this OADM, four distributed transceiver with wavelength of 1531nm, 1551nm (dropped channel), 1571nm and 1591nm were transmitted into it. Received optical power is measured at the output port and at the drop port. For downlink, the average pass-through insertion loss for the dropped channel is about 6.0 dB and the insertion loss for the remaining channels is 6.5 dB. For uplink and added channels average pass-through, the added channel insertion loss is about 7.5 dB and for the remaining channels are about 6 dB. The relatively high insertion losses are due to the power combiner or WDM coupler and optical circulators used to separate downlink and uplink channels. The isolation and crosstalk of the optical circulator are greater than 60 dB and the FBG have a central reflective wavelength of 1550.92 nm, with a 13 nm bandwidth and a reflectivity of 99.97 %.

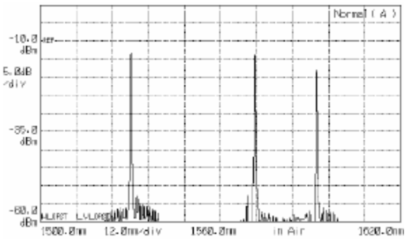
4. Results & Discussion



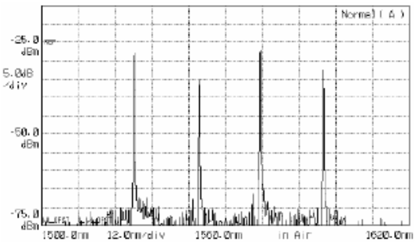
(a)



(b)



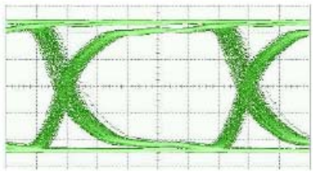
(c)



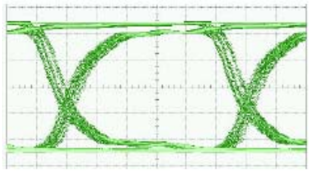
(d)

Figure 7. Optical Spectrum Result of Simulation (a) Input Channel, (b) Output of Dropped Channel, (c) Output of Remaining Channels at DP Side, (d) All Received Channels at The OLT

Figure 7 shows the optical spectrum result for the downlink and the uplink through the OADM. While Figure 8 shows the output eye diagram at 2.5 Gbps. The performance of the proposed architecture can be assessed by the BER measurement results on each channel. Figure 9 shows the BER performance against received optical power at the receiver side. For the experimental demonstration result, about 1dB performance degradation can be seen between downlink and uplink comparison for both data rates 1.25 and 2.5 Gbps. This is due to the additional power loss in the OADM cause by passive component used in the design. Comparison among downlink and uplink to all data rates shows almost a similar pattern, where there are about 1dB performance degradation between 1.25 and 2.5 Gbps. The limit of performance was set at BER 10^{-9}.

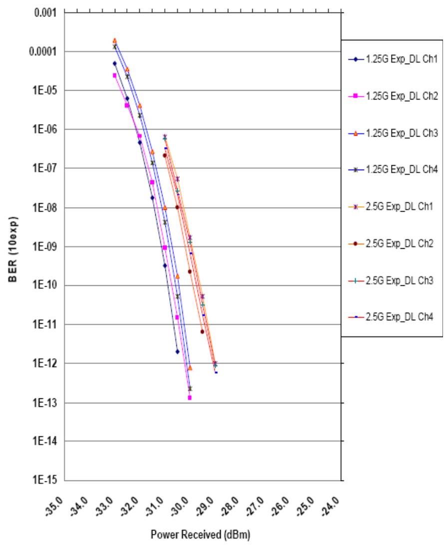


(a)

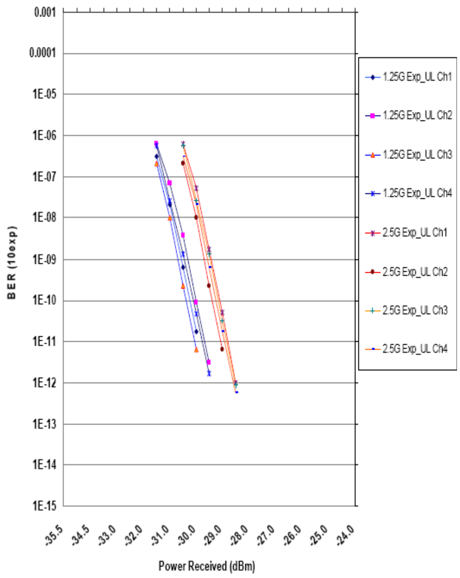


(b)

Figure 8. Typical Eye Diagram for 2.5 Gbps at -27 Dbm ONU (A) Downlink and (B) Uplink



(a)



(b)

Figure 9. BER Result At 1.25 and 2.5 Gbps Rate (A) Downlink and (B) Uplink

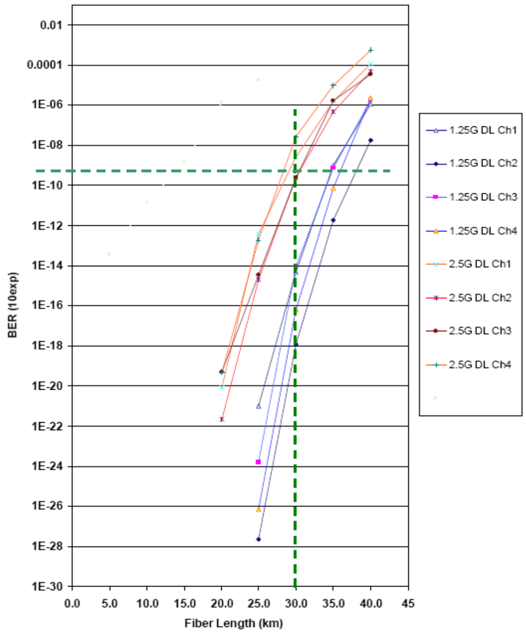


Figure 10. Result of BER Over Fiber Distance (Km)

5. Conclusions

This paper started by describing the needs and advantages of WDM-PON to the next generation access network. The trend is anticipated to change due to the high demand for large bandwidth from high-density population area such as business, commercial, education and entertainment areas. But, to have a promising access network layout is not an easy matter especially when it comes to cost effectiveness in the design architecture. Since one split point PON only can cover horizontally, to provide vertical area will require another fiber installation that will double the cost. The scalability of the proposed architecture is its main advantage over the traditionally proposed WDM-PON. Multiple logical wavelength light trees can be created on the same physical fiber link by using the new design PON OADM. Nevertheless the addition of a PON OADM into the optical link introduced an additional loss of 6.5 dB but it is still able to deliver an error-free transmission over a 30 km link with 16 optical splitting (approximately 27dB of total optical loss) as shown in Figure 10. With careful selection of launch wavelengths at the OLT, the scalability of the proposed FTTH-PON is virtually unlimited both in terms of wavelength and bit rate. The proposed architecture also has the potential of offering a wide variety of services on the same physical network.

Acknowledgement

The authors would like to express their heartiest gratitude to Telekom Malaysia Berhad for their kind support and generous funding for the projects that partially used for the work presented in this paper. The design mentioned in this paper has been submitted for the international patent filing.

References

[1] Chang Hee Lee et al., ‘Fiber-to-the-Home Using a PON Infrastructure’, Journal of Lightwave Technology, Vol. 24, No. 12, December 2006.

[2] J.-I. Kani, M. Teshima, K. Akimoto, N. Takachio, H. Suzuki, and K. Iwatsuki, "A WDM-Based Optical Access Network For Wide-Area Gigabit Access Services," IEEE Optical Communications, vol. 41, pp.S43-S48, 2003

[3] Amitabha Banerjee et al., ‘Wavelength-division-multipled passive optical network (WDM-PON) technologies for broadband access: a review’ Journal of Optical Networking, November 2005, Vol. 4, No. 11

[4] Sahrul Hilmi, Bernard H. Lee, Kaharudin Dimyati, “Scalable FTTH-PON Architecture for Unlimited User and Flexible Services”, Proceedings of SPIE, Volume 6022, Shanghai, China, December 2005

[5] David Gutierrez et al., “SUCCESS-HPON: A Next-Generation Optical Access Architecture for Smooth Migration from TDM-PON to WDM-PON” Proceedign of ECOC 2006, September 2006

BUILDING A RADAR FROM THE SCRATCH: ISRA LIPI RADAR EXPERIENCE

Mashury

Research Centre for Electronics and Telecommunications (PPET-LIPI)
[1]**mashury@ppet.lipi.go.id**

Abstract

In this paper, we present a Radar development at the Research Centre for Electronics and Telecommunications (PPET-LIPI). A coastal surveillance Radar has been developed since 2006. There are some steps for building this Radar and to realize it from the scratch. Some technical and non-technical aspects are described. Some pictures, measurement results, and Radar display are presented. The description in this paper will be used in the future for manufacturing this type of Radar in Indonesia.

Keywords: Radar, PPET-LIPI, building, aspects, manufacture

1. Introduction

Surveillance and navigation of Indonesian waters, which consist of more than 17.000 islands and 2/3 of them, are seas, will be greatly helped by the use of marine Radar. High-transmitted power Radars on coastal areas can be used to monitor the seas up to tens of nautical miles or until the border of economic exclusive zone. At the moment, a FM-CW Radar called ISRA (Indonesian Surveillance Radar) for a coastal surveillance is under development and one of the primary researches at the PPET-LIPI.

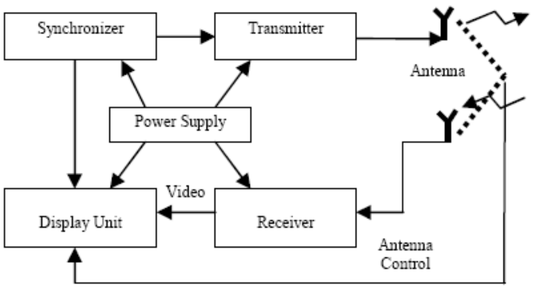


Figure 1. Block Diagram of FM-CW Radar System

A FM-CW Radar block diagram is depicted on Figure 1 [1, 2, 3, 4]. Radar system comprises two main parts: transmitter and receiver. Results of detection are shown on Radar display unit, where this unit processes the received signals into information that can be interpreted easily by the users. There are two antennas: antenna for transferring transmitted signals into electromagnetic waves and electromagnetic waves into received electrical signals. Antenna control has a function of synchronizing the antenna movement with the scanning movement on the Display unit. Synchronizer adjusts the transmitted signals with the required display of objects.

2. Basics of FM-CW Radar

A FM-CW Radar is a Radar system where a known stable frequency continuous wave radio energy is modulated by a triangular modulation signal so that it varies gradually and then mixes the signal reflected from a target object with this transmit signal to produce a beat signal. Variations of modulation are possible (sine, sawtooth, etc), but the triangle modulation is used in FM-CW Radars where both range and velocity are desired.

With the advent of modern electronics, a digital signal processing (DSP) module is commonly used for most detection processing stages. The beat signals are passed through an analog to digital converter (ADC), and, after that, a digital processing algorithm is performed on the outputs of the ADC. Most modern FM-CW Radar systems use one transmitter antenna and single antenna or multiple antennas for the receiver. Because the transmitter is on continuously at effectively the same frequency as the that of the receiver, special care must be exercised to avoid overloading the receiver stages. The form of the transmitted signal of the FM-CW is in a saw tooth shape as depicted in the following figure [1, 2, 3, 4, 5].

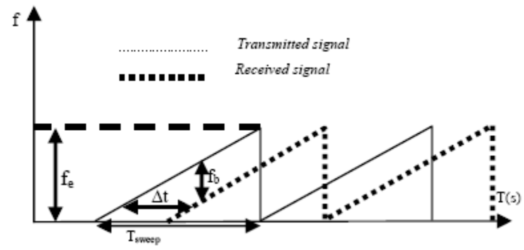


Figure 2. Shapes of the Transmitted and Received Signals Using a Frequency Modulation

Where f_e is the excursion frequency, f_b is the beat frequency (Δf) and T_{sweep} is sweep time. After reflecting from the monitored object, the received reflecting signals will be mixed with the transmitted signals to obtain beat signals as follows in Figure 1.

The beat signal can be expressed in the following equation [5, 6, 7, 8]:

$$\cos(2\pi f t) x \cos(2\pi(f + f_b)t) \qquad (1)$$

Where f is the transmitted frequency, f_b is the beat frequency, *cos()* component on the left part is the received signal and *cos()* component on the right part is the transmitted one. Using a low pass filter, the high frequency component from equation (1), i.e. $\cos(2\pi(2f + f_b)t)$ can be removed and only $\cos(2\pi f_b t)$ component remains. Thus, only signals with beat frequency f_b remain. This beat frequency f_b corresponds to the delay Δt between received and transmitted signals, and, therefore corresponds to the distance of the Radar target r.

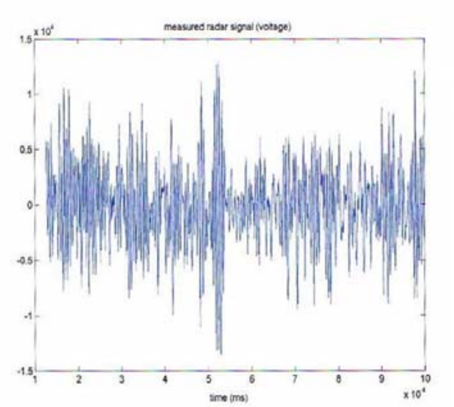


Figure 3. Beat Signal after Mixing and Subtracted by the Mean of Signal Amplitude (Sweep Time 1ms)

3. Building the Radar from the Design into a Reality

a. Radar Specifications

Before Radar is designed, the specifications must be defined. The specifications for the ISRA LIPI Radar are as follows:

- Transmitter:
  - Frequency: X band (Fc= 9.4 GHz).
  - Ranges: 64 km, 32 km, 16 km, 8 km, 4 km, 2 km. Maximum range is 64 km, which is greater than the 27 km (average distance of Radar to the horizon), to enable the Radar to detect ships beyond the horizon.
  - Output power: 2 Watts.
- Receiver:
  - IF *bandwidth*: 512 kHz.
  - Number of range cells: 512.
  - Size of Range cells: 125 meter, 62 meter, 31 meter, 16 meter, 8 meter, 4 meter.
  - PC-based signal processing.
  - PC-based display.
  - Maximum of beat frequency: 167 kHz.
  - 16bit ADC sampling for the beat signals.
- Frequency Generation
  - Main frequency generator: DRO (dielectric resonant oscillator).
  - FM – Modulation.
  - Linear saw-tooth from DDS (direct digital synthesizer).
  - Sweep Repetition Frequency: 1000Hz.
  - Fixed sweep time of 0.66mS.
  - Sweep Frequency: 2MHz, 4MHz, 8MHz, 16MHz, 32MHz, 64MHz.
- Antenna:
  - Microstrip patch arrays antenna with rectangular patch elements.
  - Antenna with flares to reduce vertical beamwidth.
  - Modular system.
  - Two antennas configuration for transmit and receive.
  - Horizontal beamwidth: 1.2°.
  - Vertical beamwidth: 10°.
  - Polarization: horizontal.
  - Rotational speed: 20 / 60 rpm.
  - Microstrip antennas have a small size, light, and never been used in FM-CW Radar.

b. Radar Design

Based on the above specifications, a design of the ISRA LIPI Radar is created:

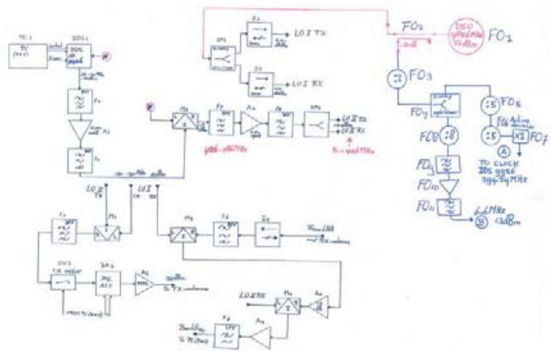


Figure 4. Block Diagram of ISRA Radar's Hardware

The above design was created based on the knowledge that has been gained in Radio Frequency (RF) and Electronics Projects, from the books on modern receiver and Radar, and also from components data sheets.

c. Components Procurement

After deciding the components that are going to be used for ISRA LIPI Radar, the components available in the market are surveyed and compared. After that, the required components are selected. Then, the local and overseas suppliers of components are contacted with the inquiries on price, delivery time, minimum quantity and guarantee.

Based on the price quote from the suppliers, a purchase order will be provided. Then, an invoice will be issued to the buyer (PPET-LIPI). It is better to buy components in bulk as the price per-item will be cheaper. After the components have been delivered, the buyer has to process the custom documents for the imported components. As a government's institution in R&D, LIPI has a privilege of free tax for imported goods.

d. Radar Assembly

Based on the block diagram in Figure 4, drawings for components assembly were created. There are three assembly drawing, i.e., transmitter, receiver and frequency generation assembly drawings. Figure 5 shows an assembly drawing for the frequency generation. Based on the assembly drawing, hardware modules, cables, connectors, and components are put together so they are all connected, see Figure 6.

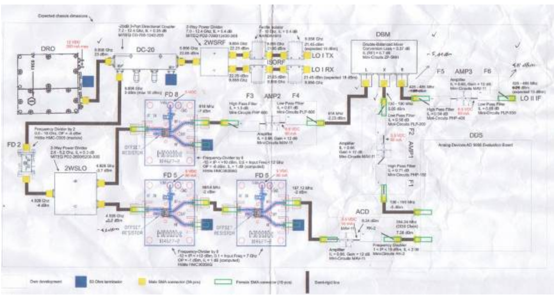


Figure 5. Assembly drawing for Frequency Generation

Figure 6. Hardware Assembly

e. Hardware Testing

To make sure that the hardware modules function as expected, some measurements are performed to check if the inputs and outputs of modules are within the acceptable levels, see Figures 7, 8 and 9. If the levels are not sufficient then improvement must be done such as adding additional amplifiers, reduce the losses introduced by the connectors and cables, check the grounding, and check the setting of modules.

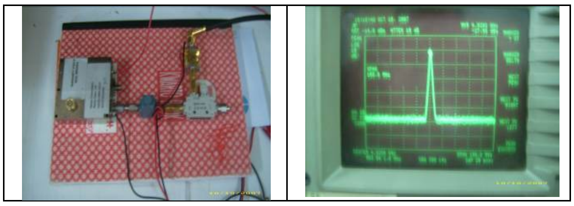
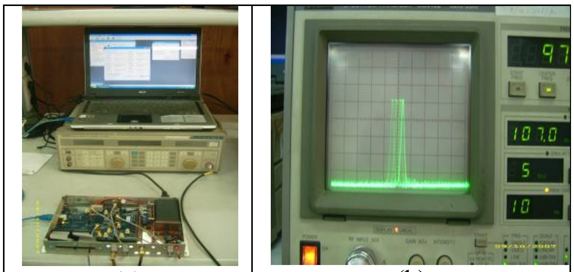


Figure 7. Meas. of DRO Output with a Divider by 2



(a) (b)

Figure 8. DDS Measurement

Figure 9. Measurement Result of LNA

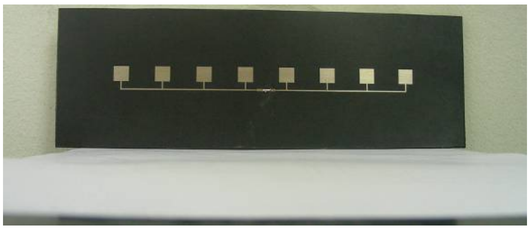


Figure 10. An Antenna Module

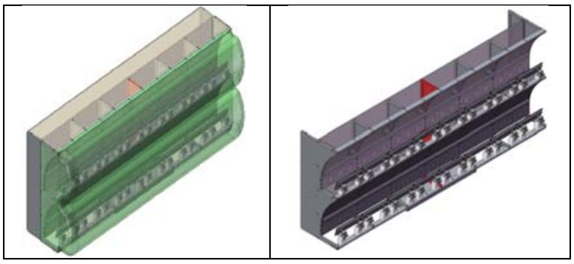


Figure 11. Design of the System Antenna

Figure 12. System Antenna with the Antenna Modules

f. Design of the System Antenna

The front end of Radar is the antennas. As shown in the Figure 1, there are two separate antennas for the transmitter (TX) and receiver (RX) of FM-CW Radar. Both of the transmitter and receiver work simultaneously. Therefore, a high isolation between the transmitting and receiving sides must be guaranteed. The minimum isolation is 60dB. Both of the TX and RX antennas are the identical. There are eight antenna modules for each TX and RX. Figure 10 shows a single antenna module for the ISRA LIPI Radar. To mount all the hardware and the antennas, a system antenna is required. The design of the system antenna is depicted on Figure 11. Figure 12 shows all antenna modules have been mounted to the system antenna.

At it can be seen on Figures 11 and 12, the holder of each antenna module can move to the left and right, up and down. This feature is important because the antenna can be adjusted to achieve an optimal transmission or reception and isolation.

g. Motor for Rotating the Radar

To rotate the antenna with a certain speed, a motor is required. This motor should be brushless in order to minimize interference to the Radar hardware. The motor is equipped with a line encoder (to feed the data on rotational speed, and azimuth angle to the PC), a gearbox (to reduce the original motor speed), a transfer gear (to reduce further the motor speed), and a motor mounting. This configuration is shown on Figure 13.

Figure 13. Motor and its Mounting

h. Software for Data Processing and Display

For the FM-CW Radar, the software plays an important role because this software performs the data and signal processing, motor control, hardware setting, and display (Radar plot).

A block diagram showing a Radar image processing system is depicted on Figure 13 [2, 4, 7, 8]. In this figure, there are five components of the system:

- Instrumentation amplifier: to amplify the filtered signals from the Radar receiver.
- Analog to digital converter (ADC): to convert analog signals into digital signals
- Interfacing and data acquisition: to adjust suitable connectors and signal levels between the Radar receiver and signal processing part. The digital data is temporarily stored prior to processing done by the image processing software.
- Image processing software: an important component for data processing, removing interference (noise and clutter) from data and extracting the wanted Radar information.
- Indicator (monitor): the retrieved information is shown on the monitor in a form of position and speed of monitored objects.

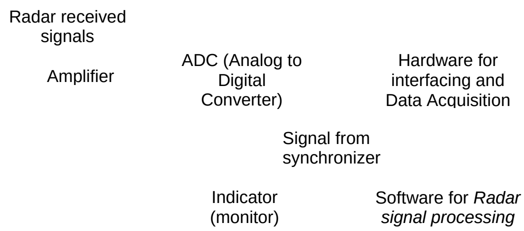


Figure 14. A Radar Image Processing Block Diagram

In the implementation, ADC, the interfacing and data acquisition components are applied by using a personal computer so that it will be more integrated and its size is reduced. In our application, we used a 16-bit ADC, then, the resolution of ADC is 0.0763 mV/step in 5 Volt as voltage reference.

The presentation of results of Radar signals processing computations are developed using a GUI algorithm. This algorithm is as follows:

1. Obtain the computed range (R), the velocity of objects V (ships) dan their radar cross sections (RCS) from the signal processing computation stage.
2. Store the obtained data in a defined format.
3. Scale the Radar display based on the maximum range it can detect.
4. Position the Radar location as the center of the Radar plot.
5. Plot the object's position based on the calculated range and its angle on the display.
6. Refresh the plot of a monitored object based on its velocity V after a certain time, which depends on the antenna rotation speed.

The latest snapshot of the Radar display is given in Figure 15. In this figure, the detected ships/objects are shown as red dots. Some hardware settings are controlled by this software. The background of this Radar display is the map of Radar station in Cilegon.

Figure 15. A Snapshot of ISRA Radar's Display

j. Radome for Protecting the Radar

A radome (the word is a contraction of Radar and dome) is a structural, weatherproof enclosure that protects a microwave or Radar antenna. The radome is constructed of material that minimally attenuates the electromagnetic signal transmitted or received by the antenna. In other words, the radome is transparent to Radar or radio waves. Radomes protect the antenna surfaces from the environment (e.g., wind, rain, ice, sand, ultraviolet rays, etc.) and/or conceal antenna electronic equipment from public view. They also protect nearby personnel from being accidentally struck by quickly rotating antennas.

There are two radomes that have been built for this ISRA LIPI Radar, see Figures 16 and 17.

Figure 16. ISRA LIPI Radar with 1st Radome

Figure 17. ISRA LIPI with the 2nd Radome

k. Human Resources

Building the Radar from the design requires a number of trained personnel. These personnel should have a knowledge on electronics, RF circuits, antennas, signal processing, programming language, PCB design, mechanical design, and measuring equipment. At the PPET-LIPI, we are lucky to have personnel with different expertise related to the fields required for Radar development as some of the personnel have experience in developing RF circuits and broadcasting systems in the past.

To improve the expertise of young personnel, the PPET-LIPI gives them opportunities to study for a higher level at local or overseas universities. Starting in 2009, the Bandung Institute of Technology (ITB) also offers a double-degree program in Telecommunications and Radar in cooperation with the International Research Centre for Telecommunications and Radar (IRCTR) of the Delft Technical University.

l. Non-Technical Aspects

There are some non-technical aspects that may hindered the progress of Radar Development, and they are:

a. Funding, the Radar development requires a lot of investment. If all funding are allocated in the start of Radar development, then building the Radar will not take a long time as compared to the case with funding is termed yearly.
b. Components delivery, the delivery time for imported components/modules may take up to six months.
c. Facilities, there should a sufficient number of measuring equipment and room for Radar assembly/testing.
d. Team work, there should be a strong team for Radar development because the personnels must focus on what they are doing and strongly motivated.

4. Conclusions

A coastal surveillance Radar is necessary for a country like Indonesia that has more than 17,000 islands. Development of ISRA LIPI Radar has been presented. There are a lot of factors involved in the Radar development. To be able to produce Radar locally is the main goal of this Radar development.

Acknowledgment

The development of ISRA LIPI Radar is in a cooperation with the International Research Centre for Telecommunications and Radar of the Technical University of Delft. The funding for the Radar development comes from the LIPI research grants (DIPA and Kompetitif programmes) and from the Ministry for Research and Technology (Insentif Programmes).

References

[1] Leo P. Ligthart, September 2005, ’Short Course on Radar Technologies’, International Research Centre for Telecommunications and Radar, TU Delft.
[2] Bassem R. Mahafza, 2005, ‘Radar Systems Analysis and Design Using MATLAB’, Chapman & Hall.
[3] M.I. Skolnik, 1990, “Radar Handbook”, McGraw-Hill.
[4] Mark Richards, 2005, ’Radar Signal Processing’, McGraw-Hill.
[5] S.H. Heijnen, October 2003, ‘TARA Data Processing’, Report for Cloudnet.
[6] S.H. Heijnen, J.S. Van Sinttruijen, W.F. Van der Zwan, L.P. Ligthart, 2000, ’A Dedicated Computer System for FM-CW Radar Applications’, Proceeding of 13th International Conference on Microwaves, Radar and Wireless Communications.
[7] M. Wahab, P. Daud, 2006, ‘Image Processing Algorithm for FM-CW Radar’, TSSA/WSSA Int. Conference 2006, ITB Bandung.
[8] Mashury, 2006, ’ Development of Radar Image Processing Algorithm’, Information and Communication Technology Seminar 2006, ITS Surabaya.

DESIGN AND SIMULATION OF HIGH GAIN HIGH STABILITY POWER AMPLIFIER FOR ISRA LIPI RADAR

Mashury

Research Centre for Electronics and Telecommunications (PPET-LIPI)
mashury@ppet.lipi.go.id

Abstract

Power amplifier (PA) is an important part of transmitter's front end. The cost for a PA module for applications in Radar and Microwave frequencies is very expensive. In this paper, we present a research and development on a PA module for Radar developed by PPET-LIPI, which is called ISRA (Indonesia Sea Radar). The in-house development of the PA module can significantly reduce its price. The development of PA is started by defining the specifications, design and simulation, and optimization. The component, i.e., the SMD transistor, has been purchased and its data are used in the design and simulation. The designed result will be verified if it matches or exceeds the specifications. The expected gain of this PA is about 40 dB with a high stability. The simulated results are presented to show that this PA achieves the expected performance.

Keywords: Power Amplifier, design and simulation, ISRA Radar, gain, stability

1. Introduction

Surveillance and navigation of Indonesian waters, which consist of more than 17.000 islands and 2/3 of them, are seas, will be greatly helped by the use of marine Radar. High-transmitted power Radars on coastal areas can be used to monitor the seas up to tens of nautical miles or until the border of economic exclusive zone. At the moment, a FM-CW Radar called ISRA (Indonesian Surveillance Radar) for a coastal surveillance is under development and one of the primary researches at the PPET-LIPI.

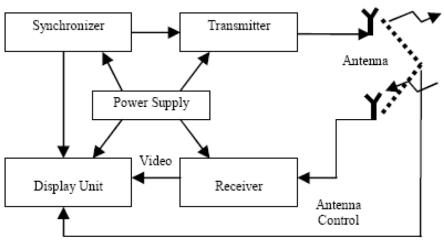


Figure 1. Block Diagram of FM-CW Radar System

A FM-CW Radar block diagram is depicted on Figure 1 [6, 7, 8, 9]. Radar system comprises two main parts: transmitter and receiver. Results of detection are shown on Radar display unit, where this unit processes the received signals into information that can be interpreted easily by the users. There are two antennas: antenna for transferring transmitted signals into electromagnetic waves and electromagnetic waves into received electrical signals. Antenna control has a function of synchronizing the antenna movement with the scanning movement on the Display unit. Synchronizer adjusts the transmitted signals with the required display of objects.

As it can be shown in Fig. 1, PA is an important part of the front end of the Radar transmitter (TX). This PA amplifies the transmitted signals before being converted into electromagnetic waves by the TX antenna. The PA should have a high gain and high stability so that it will amplify the signals without turning into an oscillator [1, 2, 3]. A high performance PA contributes to the high accuracy of the Radar detection because the received signals will be stronger.

Currently, an imported PA Module for Radar applications costs expensive because it has high requirements. High labor costs in Europe and US also contribute to the high price of PA. Therefore, an own-developed PA module at the PPET-LIPI can significantly reduce the Radar development cost. PPET-LIPI has sufficient facilities to develop the PA module. The active components, i.e., transistors or amplifier ICs, are commercially available and imported from overseas manufacturers.

2. Basics of FM-CW Radar

A FM-CW Radar is a Radar system where a known stable frequency continuous wave radio energy is modulated by a triangular modulation signal so that it varies gradually and then mixes the signal reflected from a target object with this transmit signal to produce a beat signal [6, 7]. Variations of modulation are possible (sine, sawtooth, etc), but the triangle modulation is used in FM-CW Radars where both range and velocity are desired [5, 6, 7, 8].

With the advent of modern electronics, a digital signal processing (DSP) module is commonly used for most detection processing stages. The beat signals are passed through an analog to digital converter (ADC), and, after that, a digital processing algorithm is performed on the outputs of the ADC.

Most modern FM-CW Radar systems use one transmitter antenna and single antenna or multiple antennas for the receiver. Because the transmitter is on continuously at effectively the same frequency as that of the receiver, special care must be exercised to avoid overloading the receiver stages. The form of the transmitted signal of the FM-CW is in a saw tooth shape as depicted in the following figure [6, 7, 8, 9].

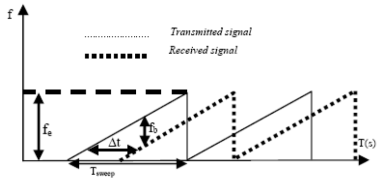


Figure 2. Shapes of the Transmitted and Received Signals Using a Frequency Modulation

Where f_e is the excursion frequency, f_b is the beat frequency (Δf) and T_{sweep} is sweep time. After reflecting from the monitored object, the received reflecting signals will be mixed with the transmitted signals to obtain beat signals as follows:

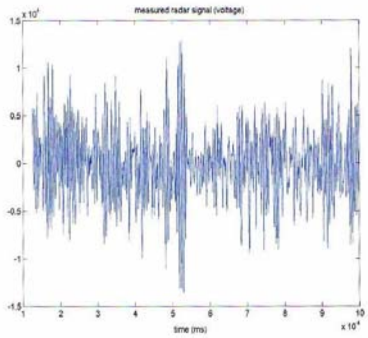


Figure 3. Beat Signal after Mixing and Subtracted by the Mean of Signal Amplitude (Sweep Time 1ms).

The beat signal can be expressed in the following equation [5, 6, 7, 8]:

$$\cos(2\pi f t) x \cos(2\pi (f + f_b) t) \qquad (1)$$

Where f is the transmitted frequency, f_b is the beat frequency, *cos()* component on the left part is the received signal and *cos()* component on the right part is the transmitted one. Using a low pass filter, the high frequency component from equation (1), i.e., $\cos(2\pi(2f + f_b)t)$, can be removed and only $\cos(2\pi f_b t)$ component remains. Thus, only signals with beat frequency f_b remain. This beat frequency f_b corresponds to the delay Δt between received and transmitted signals, and, therefore corresponds to the distance of the Radar target r.

3. Design Considerations

Some considerations must be taken into account in designing the PA. These important considerations are:

- The stability factor K >> 1.
- Input and the output match to the source and the load, where the impedance is 50 Ω + j0.
- VSWR < 2.00:1.00.
- Optimum gain for the required bandwidth.
- The PA can be single stage and multi stage, which requires interstage matching.
- An isolator should be inserted at the input, and the interstage.

4. PA Design

Before designing the PA, specifications must be defined. The specifications are as follows:

- Gain: 30 dB, min
- Daya output (@1dB comp): +33 dBm, max
- VSWR: max 1.5:1
- Gain flatness vs freq: +/- 0.5 dB, max
- Noise Figure: 5 dB, max
- Maximum input level, Pin Max: 5 dBm
- Input and output impedance: 50 ohm
- Desired working frequency: 9.370-9.430 GHz

PA design was performed using RF design software called LINC from Applied Computational Sciences (ACS). This software has a capability of analyzing and synthesizing RF circuits based on the input and output parameters. The RF circuits can be implemented using lumped elements, active components and microstrip. Performance improvement can be done via optimization feature of this software.

5. Simulated Results

In this section we present some simulated results on the PA design. The schematic of this PA design was generated using synthesis facility of the LINC software.

Figure 4 shows a schematic diagram of the PA. The PA has an input and output matching to the load and source of 50 ohm. The matching stages have been converted into microstrip lines with a substrate of duroid 5880. This substrate has a permitivity factor of 2.2. This substrate is suitable for RF applications in a high humidity environment.

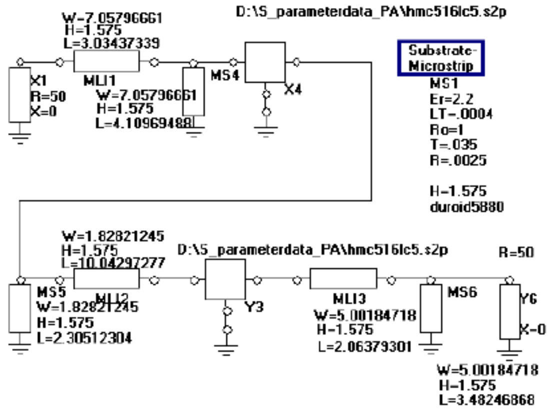


Figure 4. Schematic Diagram of PA

The performance of S21 (forward gain) and S11 (return loss) are depicted on Figure 5. In this figure, the gain of PA for the frequency range of 9,370-9,430 MHz is achieved at about 37 dB. Although the maximum gain is not achieved at this frequency range, where the maximum gain is about 39 dB at 10 GHz, the lowest level of S11 is achieved within this frequency range. This shows that the PA design results in the expected performance.

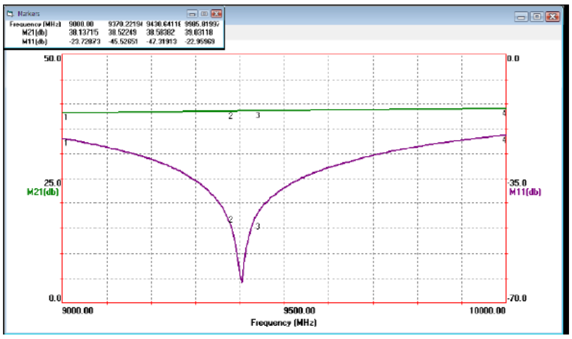


Figure 5. The Gain S21 and Return Loss S11 of the PA

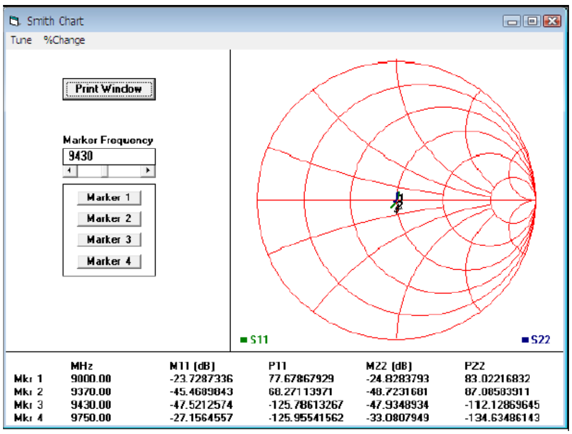


Figure 6. Input and Output Return Loss on Smithchart

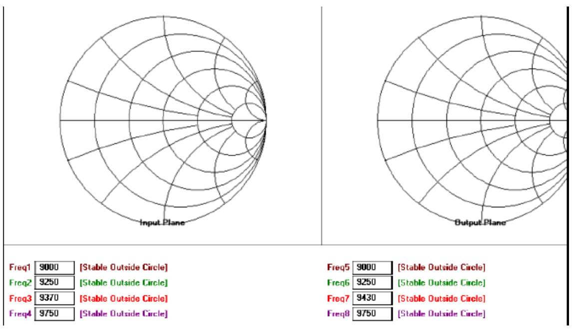


Figure 7. Stability Circles of the PA

Figure 6 shows input and output return loss S11 and S22 of the designed PA on the Smith chart. For the desired frequency range of 9370-9430 MHz, the return loss S11 achieves a very low level, which is less than -45 dB. This means that the input source is really matched to the device input. While, the output return loss S22 is also less than -47 dB. This shows that the input and output matching stages are really matched to the devices.

Stability circles are shown in Figure 7 for the input and output planes. Both of the stability circles are stable outside, which are far from the smith chart so that it can not be seen in the above figure. However, for the requested frequency range 9,370-9430 GHz, they are stated as stable outside circle. This means that the designed PA has a very good stability with K factor >>1.

6. Evaluation

Based on the simulated results, the gain for the designed PA with a working frequency of 9.370-9,430 MHz is sufficient with a value of 37 dB. The VSWR is close to 1:1 and this is due to very small values of the return loss S11 and S22, which is less than -45 dB. The stability factor K is shown to be much greater than 1.

According to the simulated results, the design of PA meets the required specifications. An optimization can be performed on the PA implementation to the microstrip with Duroid substrate by trimming the size of the line's width and length.

Thus, the designed PA achieves a high performance for a high stability factor and high gain. This high gain was achieved by implementing multiple stages PA.

7. Conclusions

PA module plays a significant play role at the Radar Transmitter. Design and simulation of a PA module for the FM-CW Radar have been presented. The in-house development of a PA can significantly reduce the cost for Radar development. Improvement and optimization will be performed to the PA during implementation.

Acknowledgment

The development of ISRA LIPI Radar is in a cooperation with the International Research Centre for Telecommunications and Radar of the Technical University of Delft. The funding for the Radar development comes from the LIPI research grants (DIPA and Kompetitif programmes) and from the Ministry for Research and Technology (Insentif Programmes).

References

[1] Michael Tse, "Power Amplifier Design", http://cktse.eie.polyu.edu.hk/eie403/PowerAmplifier.pdf.

[2] Philips Semiconductor, "Power Amplifier Design ", http://www.nxp.com/acrobat_download/various/SC19_POWER_AMPL_DESIGN_1.pdf.

[3] Applied Computational Sciences,"LINC2," version 2.72H, 2009.

[4] Mark Richards, 'Radar Signal Processing', McGraw-Hill, 2005.

[5] Bassem R. Mahafza, 'Radar Systems Analysis and Design Using MATLAB', Chapman & Hall, 2005.

[6] M.I. Skolnik, 'Radar Handbook', McGraw-Hill, 1990.

[7] Leo P. Ligthart, 'Short Course on Radar Technologies', International Research Centre for Telecommunications-transmission and Radar-TU Delft, September 2005.

[8] S.H. Heijnen, October 2003, 'TARA Data Processing', Report for Cloudnet.

[9] S.H. Heijnen, J.S. Van Sinttruijen, W.F. Van der Zwan, L.P. Ligthart, 2000, 'A Dedicated Computer System for FM-CW Radar Applications', Proceeding of 13th International Conference on Microwaves, Radar and Wireless Communications.

SIMULATION OF SEQUENTILA FAST ADC FOR GROUND PENETRATING RADAR (GPR) APPLICATIONS

Mery Kombo[1], Heroe Wijanto[2], Iswahyudi Hidayat[3]

[1,2,3]Electro and Communication Faculty, Institut Teknologi Telkom
[2]**hrw@ittelkom.ac.id**, [3]**isw@ittelkom.ac.id**

Abstract

Ground Penetrating Radar (GPR) is an electromagnetic method with high frequency and resolution. GPR use electromagnetic wave radiation at band spectrum frequency range of 40 MHz up to several GHz to identification the reflection shape from underground with penetration deeps. The transmit pulse will be differ because characteristic each underground layer is different (depend on dielectric constant each layer). The received reflection pulse converted to digital by Analog to Digital Converter (ADC). Output of this ADC can give profit in Digital Signal Processing (DSP) to identification object more detail. ADC that uses in this GPR systems is *Sequential Fast ADC*. It is because this GPR use frequency 1 GHz (frequency selection base on operational frequency of GPR and penetration deeps) so needs high sampling rate and resolution. Unit of *Sequential Fast ADC* consist of sampling, quantization, and encoding. To compare resemblance between digital signals and analog signals so the digital signals must convert to analog by interpolation polynomial. In this paper was compared the effect of dynamic range GPR signals to Signal to Quantization Noise Ratio (SQNR). The effect of uniform and non-uniform quantization to quantization error. The effect of sampling rate (oversampling) and quantization levels to interpolated signals. From the simulation dynamic range of GPR signal must appropriate with dynamic range of ADC. The quantization that compatible with GPR is uniform quantization, the resolution (number of bit) of ADC that effective for GPR is 8 bit and the Oversampling (space between samples was closed) make the error reconstruction decrease. Linear interpolation is the sufficient method to reconstruction the GPR signal.

Keywords: Ground Penetrating Radar (GPR), Analog to Digital Converter (ADC), SQNR, Dynamic Range, Interpolation, and Reconstruction Signal.

1. Background

Ground Penetrating Radar (GPR) is a new explorating technique that use electromagnetic wave with high frequency (40 MHz until 3 GHz). High frequency in this technique can give good resolution. GPR used in many sites like geology, construction, archeology, forensic, etc The use oh high frequency make conversion of GPR signal (analog signal) to digital signal is faster.

Sequential fast ADC (Analog to Digital) in GPR system is essential. It can change the analog signal (GPR signal from the ground) to digital with high sampling rate.

2. Theorem

2.1. Ground Penetrating Radar (GPR)

Ground Penetrating Radar (GPR) is known as georadar or surface penetrating radar. GPR is a geophysics that used electromagnetic wave that can give information of continue cross-sectional profile without th needed to mining the ground. This method is used to know the condition and characteristic surface of underground.

The radar method using electromagnetic wave with higher frequency (40MHz until 3GHz). Its sensitive with the contras conductivity.

Block system of GPR system can be look at figure 1. Like the other radar system, GPR system have transmitter, receiver, and digital signal processing and image processing. The transmit signal is Gaussian signal pulse with center frequency 1 GHz. This signal will translate to digital format by Sequential fast ADC and then will process to DSP processing

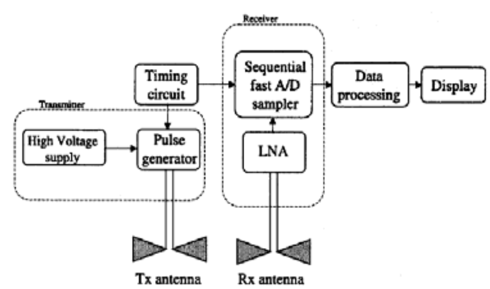


Figure1. Blok System GPR

2.2. Analaog to Digital Converter

Analog to Digital Converter (ADC) is used to convert the analog signal to digital. ADC system have sampling process, quantization process and encoding process. For radar application ADC use high resolution and high sampling rate

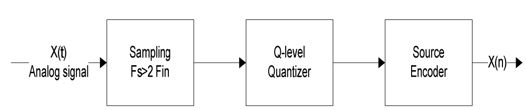


Figure 2. Blok Sistem ADC

One of ADC method is Sequential Fast ADC or is known as Flash ADC.

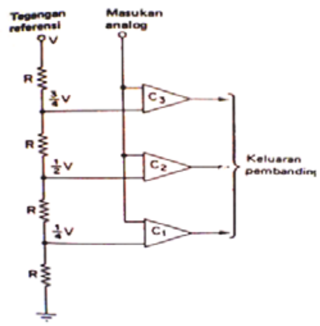


Figure 3. Sequential ADC

The circuit at figure 3. use some circuit to compare the signal. The number of comparator circuit is suitable to ADC resolution. Every comparing circuit have a reference. The number of comparator is equal to 2^n-1, where n is number bit binary.

The advantages of Flash ADC is a simple circuit and the operating rate that suitable to convert GPR signal. But it needs many comparator so the system be larger and expensive.

3. Simulation Model

Design model for the simulation Sequential ADC for GPR application can be see in figure

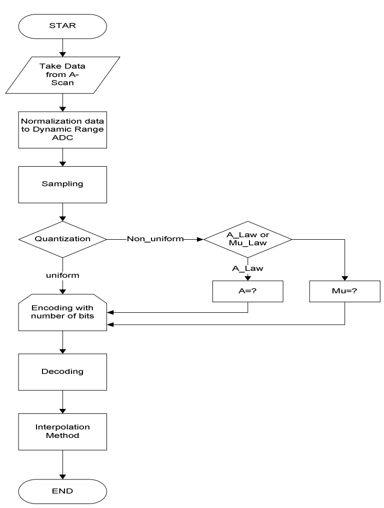


Figure 4. Model of Simulation Sequential Fast ADC for GPR Application.

4. Analysis of Simulation

4.1. Effect of Dynamic Range Signal GPR to Signal to Quantization Noise Ratio (SQNR)

Assume that Dynamic Range of ADC are -5V until 5V and using uniform Quantization, Showing that for lower bit level, if GPR signal is not suitable to dynamic range ADC its effect to smaller SQNR, so it can be effected to performance of the signal, where SQNR=10 log $\frac{Px}{Pe}$, Px is power signal and Pe is Power of error quantization.

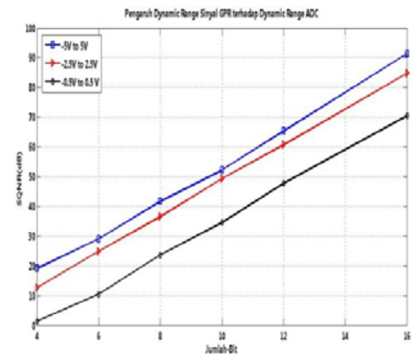


Figure 5. Effect Dynamic Range to SQNR

Table 1. Dynamic Range to SQNR

Dynamic Range GPR Signal	Signal to Quantization Noise Ratio(SQNR) (dB)(number of bit)					
	4	6	8	10	12	16
-5 to 5	19.1	**29.02**	41.58	52.5	65.3	9140
-2.5 to 2.5	12.57	24.85	**36.52**	49.37	60.66	8440
-0.5 to 0.5	1.389	10.41	23.64	**34.52**	47.73	7050

Table show, that if GPR signal is smaller than dynamic range, ADC will need larger number of bit to get the same SQNR. So it need to be suitable to dynamic range ADC. Because of ADC that used in GPR application is Flash ADC where a number of comparator is used to form sampling signal is equal to $2^{\text{number of bit} - 1}$. So the higher resolution is need many comparator, than it make the system be expensive and large.

4.2. Effect of Uniform Quantization and Non Uniform to Quantization Error

Figure 6 show when A equal to 1 its present to uniform quantization. Figure 6 showing that uniform quantization is effective to lower number of quantization bit (<8).

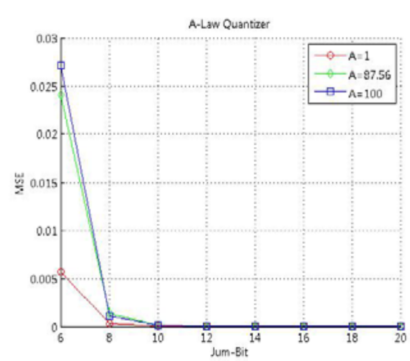


Figure 6. Performance of A law Error Quantization

Companding signal make error, using a higher level quantization make an error will be larger than before, so it need to expand. Because of this Non uniform quantization need to use narrow level quantization or using larger number of bit (>100).

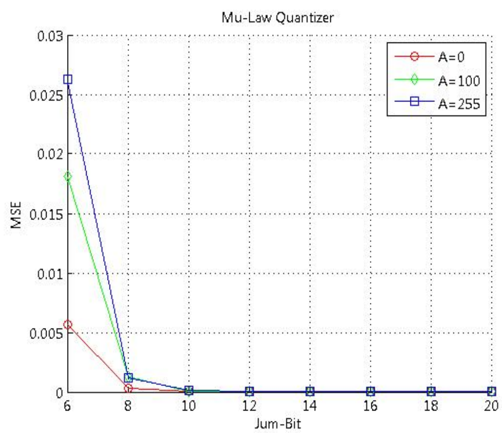


Figure 7. Performance of µ-Law Error Quantization

Figure 7 showing that the rising number of bit make smaller error quantizied . For A equal to 0, for the same level quantizied, error in linear quantization is larger than non uniform quantization. Figure 6 and 7, showing that expand method in µ law for lower level is more reliable than A law.

4.3. Effect Number bit binary to Reconstruction GPR signal using Interpolation Methode

Figure 8 showing that larger number of bit that used interpolation method make better reconstruction signal. Larger number of bit make quantization level be closer than before, and the forming of signal can be accurate.

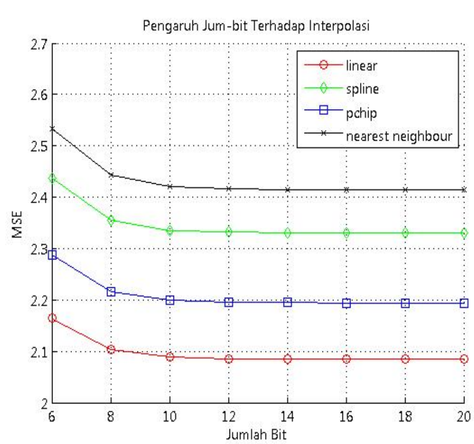


Figure 8. Performance of Interpolation Method Effected by Number of Bit.

Table 2. MSE Interpolation Effected by Number of Bit.

Interpolation polynomial	Number of bit	MSE
Nearest neighboar	6	2.711
	8	2.607
	10	2.582
	12	2.578
	14	2.577
	16	2.576
	18	2.576
	20	2.576
spline	6	2.534
	8	2.447
	10	2.424
	12	2.42
	14	2.419
	16	2.418
	18	2.418
	20	2.418
Cubic spline (pchip)	6	2.417
	8	2.339
	10	2.318
	12	2.315
	14	2.314
	16	2.314
	18	2.314
	20	2.314
Linear	6	2.307
	8	2.24
	10	2.222
	12	2.219
	14	2.218
	16	2.218
	18	2.218
	20	2.218

For a number of bit 12 until 20, MSE signal is relative same. Number of bit 8 is relative efficience to convert GPR signal to digital. Using 8 bit make quantization level equal to =(5-(-5))/2^8=0.0390, and spacing of closer GPR signal is 0.1265 so delta equal to 0,06325. Because of delta is larger than quantization level, the signal will not to rounding to lower or above quantization level. MSE in figure 8 showing that linear interpolation is a good reconstruction method to GPR signal, and Neighbor interpolation is not suitable for reconstruct GPR signal.

4.4. Effect Oversampling (Space at sampling signal) to Reconstruction GPR signal

In this case oversampling will be analogy to space between sampling signal. Higher oversampling effect to narrow space between sampling signal. It can be done because of sampling signal is an A-Scan digital (computation in matlab is digital). In this simulation A-Scan signal is known as analog signal.

Figure 9 show that lower space between sampling signal, make the reconstruction signal for every interpolation method be better because the order of polynomial coefficient be larger so the performance of the signal can be better.

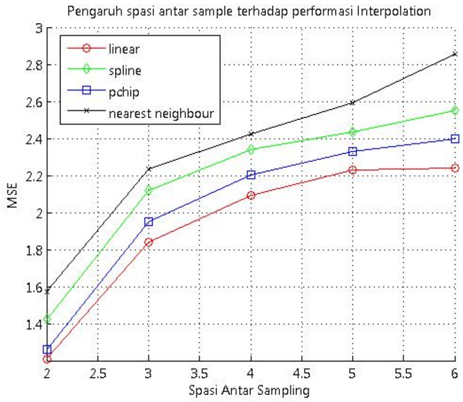


Figure 9. Performance of Interpolating Effected by Oversampling

Figure 9 shown that linear interpolation is a good method interpolation because of narrow space between the signal. Narrow space between signal make Lower MSE. Neighbor interpolation is not suitable for GPR signal because this method use number between closer data to interpolate signal.

Table 3. Effected MSE to Space between Sample

Interpolasi	Spasi Antar Sample	MSE
Linear	2	1.028
	3	1.839
	4	2.091
	5	2.231
	6	2.238
Cubic spline	2	1.422
	3	2.118
	4	2.34
	5	2.434
	6	2.55
Cubic Hermite	2	1.259
	3	1.948
	4	2.202
	5	2.328
	6	2.397
Nearest neighbor	2	1.571
	3	2.234
	4	2.423
	5	2.594
	6	2.858

From figure 9, larger sampling faquency ADC make number sample for reconstruction be larger too, and the error of reconstruction can be lower. GPR signal have closer space so it need faster sampling rate. Right now ADC high speed have conversion rate 4GHz. For implement sampling rate must be suitable to MSE performance and sampling rate ADC.

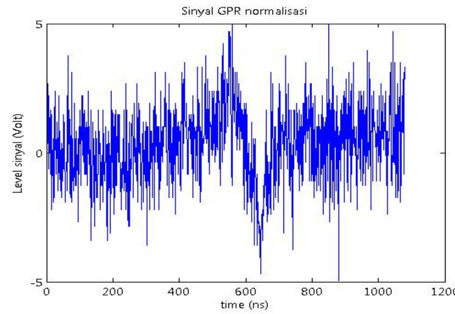


Figure 10. Normalised GPR Signal

Figure 11. shown reconstruction of GPR signal using some method of interpolation polynomial. This figure used space between signal is 5 and used 8 bit for resolution ADC. From four interpolation metode, reconstruction signal by linear interpolation is like the original signal because its have smaller MSE.

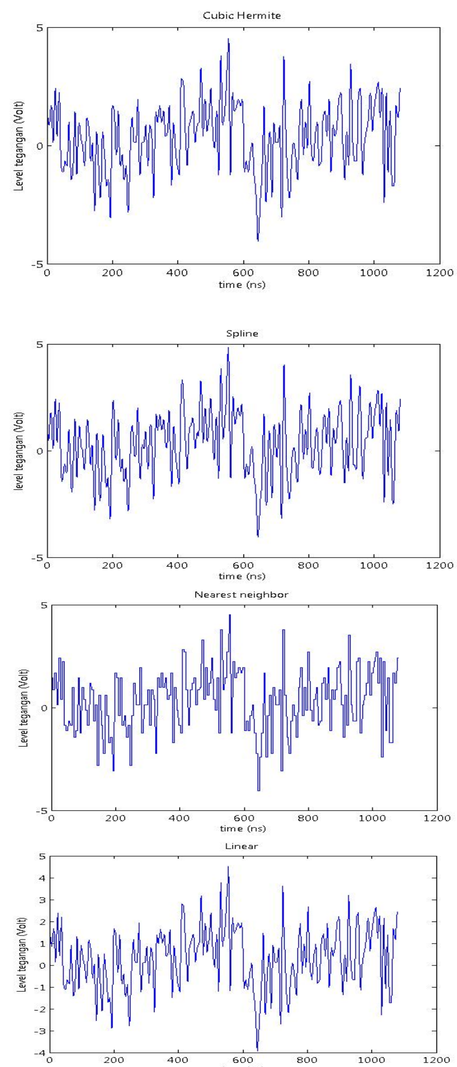


Figure 11. Interpolating Signal

5. Conclusion and Suggestion

5.1. Conclusion

1. GPR signal have narrow bandwidth pulse, and the fluctuation signal (dynamic range of the signal is effected by permitivity and permeability of surface grounding layer.
2. GPR signal is better when it follow the dynamic range of ADC.
3. When the number of bits that used is larger then the Signal to Quantization Noise Ratio (SQNR) is larger too.
4. GPR need higher sampling rates.
5. From the pdf GPR signal, non uniform quantization is suitable to quantizied GPR signal, but this method is not effective because companding process before and after quantization process.
6. Uniform quantization have a good performance than non uniform quantization in lower bit level (<8).
7. The distorsion of quantization is larger when ADC have small level quantization.
8. Resolution of ADC using 8 bit is effective to quantizied GPR signal.
9. The reconstructing signal when using interpolating method will like the original signal when space in the signal is smaller.
10. From the simulation, Linear interpolation is suitable for GPR signal because it have lower MSE than the other polynomial interpolation.

Reference

[1] Demler , Michael J,"High Speed Analog to Digital Converter", Academic Press, San Diego, California.

[2] Proakis, John.G., and Manolakis G.D. 1996. "Digital Signal Processing,"Prentice Hall, New Jersy

[3] Iskandar, Magdy.F.1992."Electromagnetic Field and Waves,"Waveland Press, United State of America.

[4] Nils Bjorklun,Tore Johson,"Real Time Sampling Of Ground Penetrating Radar and Related Processing",University of Technology Departemen Electrical Enginering Division Signal Processing.

[5] Athanassion Manikas,"Digital Comunication An Overview of Fundamental",Deprtement of Electrical and Electronic Engineering.

[6] A.G.Yarovoy, P.van Genderen, L.P.Ligthart, 'Ground Penetrating Radar For Landmine Detetion", Faculty of International Technology and System, Nederland.

[7] Shalabh, Goyal and Abhijit, chatterje, " A low-Cost Test Methodology for Dynamic Specification Testing Of High Speed Data Converter", Journal Of Electronic Testing, United State.

[8] Jun He," A Prototype of A New Class of Oversampling ADC",Sanghai Jiao Tong University,Shanghai,Cina.

[9] M.Frey and H.-A.Loiger,"on Flash A/D Converter with low-precision Comparator,"in Proc.IEEE Int, Kos Island.

[10] Cheng-Liang,Cheng,"Probability, Statict and Interpolation",PSE laboratory, National Taiwan University.

[11] C.De Boor, A Practical Guide to Splines, Springer-Verlag,New York.

[12] Jeffery, R.Price,"Resampling and Reconstruction With Fractal Interpolation Funtion", student member,IEEE.

[13] Alfredeson, Jon," Design of Paralel A/D Converter system on PCB for High Speed Sampling and Timming Error Estimation",Linkoping University, Swedwn.

PERFORMANCE ANALYSIS OF MIMO-STBC SYSTEM IN HSDPA OVER FADING RAYLEIGH CHANNEL

Bayu Jati Laksono [1], Budi Prasetya [2], Ida Wahidah [3]

[1,2,3] Faculty of Electrical and Communication Engineering, Institut Teknologi Telkom
[1]**b_laxono@yahoo.com**, [2]**bpy@ittelkom.ac.id**, [3]**idw@ittelkom.ac.id**

Abstract

Release 6 of 3GPP (Third Generation Partnership Project) adds HSDPA (High Speed Downlink Packet Access) in WCDMA (Wideband Code Division Multiple Access) as an effort to make system more efficient for packet data applications by increasing peak data rates and reducing packet latency. Theoretical data rate for HSDPA is approximately 14 Mbps. The actual HSDPA rate achieved is still much lower than that. Therefore the use of Multiple-antenna on transmitter and receiver, known as a Multiple-Input Multiple-Output (MIMO) technique, considered to be able to improve the system performance of physical layer by increasing the capacity and getting gain diversity. This research simulates the physical layer model of HSDPA depicted in 3GPP standards. These include generation of transport blocks, turbo coding, rate matching, scrambling and mapping as the modulation. The MIMO Space Time Block Code (STBC) scheme with two transmitters and two receivers is integrated in physical layer of HSDPA. Some important features of HSDPA are also simulated as the test performance of STBC 2x2 which is compared by Single-Input Single-Output (SISO) HSDPA. These features are Fast Retransmission, Adaptive Modulation and Coding (AMC), the use of Hybrid Automatic-Repeat-Request (HARQ), and 2 ms Transmission Time Interval (TTI). These systems are tested in rayleigh fading channel with Gaussian noise (AWGN). The simulation result shows that STBC 2x2 HSDPA have better performance with improvement average 2,63 dB for each transmission while SISO HSDPA only reach 0,9 dB in doppler frequency of 19,4 Hz. The STBC 2x2 also reduces the power about 4,9 dB for achieving 384 kbps HSDPA throughput compared with SISO HSDPA in 0 Hz doppler frequency.

Keywords: HSDPA, STBC, AMC, Fast Retransmission, Throughput

1. Introduction

HSDPA (*High Speed Downlink Packet Access*) is the WCDMA (Wideband Code Division Multiple Access) evolution in downlink which often named 3.5G technology. HSDPA has more advantages than the WCDMA, such as the higher throughput and the shorter delay.

Throughput indicates the performance system in HSDPA. The HSDPA possibly achieves the throughput up to 14 Mbps with 5 MHz bandwidth. These improvement can be reached because of the new physical channel, the Adaptive Modulation and Coding (AMC) implementation, the Hybrid Automatic Repeat Request (HARQ) retransmission, the shorter (2 ms) Transmission Time Interval (TTI) than the WCDMA (10 ms), also the Fast Scheduling, and Fast Cell Selection (FCS) in WCDMA platform.

The use of MIMO (Multiple Input Multiple Output) potentially increases the received data rate with high power efficiency in HSDPA. This final thesis uses the Space Time Block Codes (STBC) method with 2 transmitter antennas and 2 receiver antennas. The purpose of this diversity method is increasing the diversity gain without channel knowledge in transmit antennas.

2. System Modeling

2.1 STBC Model in HSDPA

The design of HSDPA system is divided in two main parts: the HSDPA transmitter and the HSDPA receiver. The STBC-HSDPA modeling is simulated in MATLAB 7.1 as depicted in figure 1.

2.2 HSDPA Transmitter

a. Random Data Generator

The transmitted data of HSDPA system is in binary (0 or 1) that uniformly distributed. The data length is determined by Transport Block Size (TBS) from Channel Quality Indicator (CQI) calculation as following [15].

$$CQI = \begin{cases} 0 & SNR \leq -16 \\ \left\lfloor \frac{SNR}{1,02} + 16.62 \right. & -16 < SNR < 14 \\ 30 & SNR \geq 14 \end{cases} \quad (1)$$

The value of TBS is used in HSDPA throughput simulation. While the Bit Error Rate (BER) and Frame Error Rate (FER) simulation use fixed data length 3202 bits for each frame according to Fixed Reference Channel-5 (FRC-5) of 3GPP [1] plus 24 CRCbits.

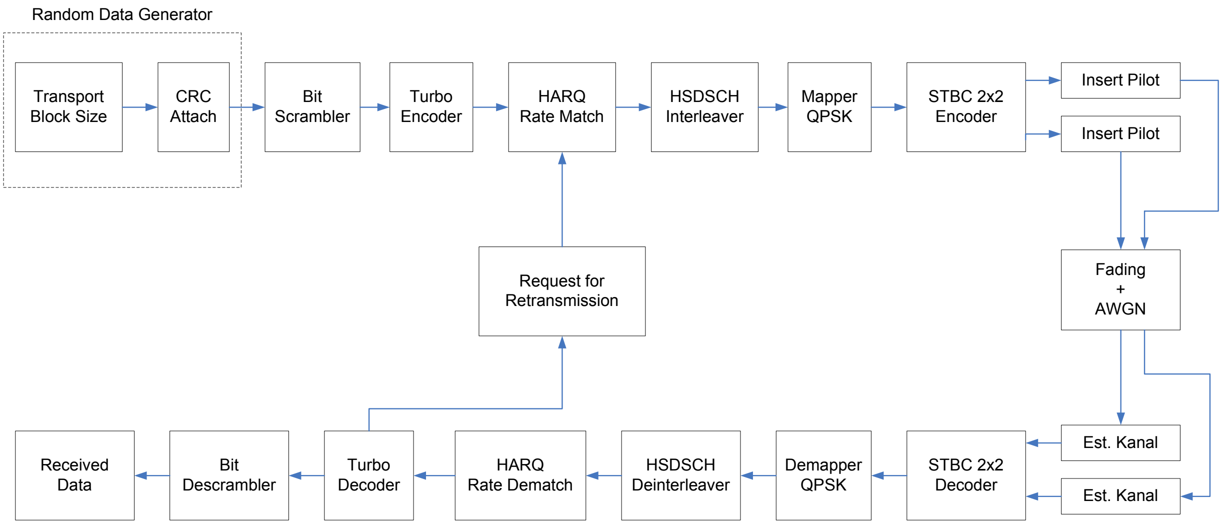


Figure 1 STBC-HSDPA Model

b. Bit Scrambler

The output bits from data generator are scrambled using bit scrambler. The output bits of scrambling process are denoted as $d_{im,k}$ which is defined by the following relation [3].

$$d_{im,k} = \left(b_{im,k} + y_k\right) \bmod 2 \qquad k = 1,2,\ldots,B \tag{2}$$

$$y_\gamma = \left(\sum_{x=1}^{16} g_x \cdot y'_{\gamma-x}\right) \bmod 2 \quad 1<\gamma\le B \tag{3}$$

Where $b_{im,k}$ is the input bits from scrambler with y_k as the scrambling sequence in k number bits and $g = \{g_1, g_2, \ldots, g_x\} = \{0,0,0,0,0,0,0,0,0,0,1,0,1,1,0,1\}$.

c. Turbo Encoder

The Structure of turbo encoder is a Parallel Concatenated Convolutional Code (PCCC) with two 8-state constituent encoder and one internal interleaver. The coding rate of turbo encoder is 1/3 with turbo polynomial generator $g0(D) = 1 + D2 + D3$ dan $g1(D) = 1 + D + D3$ as depicted in figure 2.

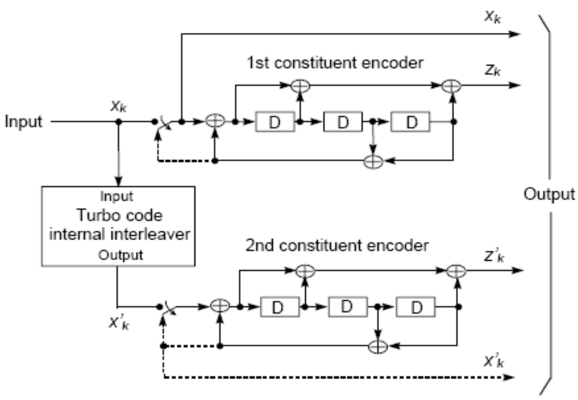


Figure 2 Structure of turbo encoder [3]

d. HARQ Rate Match

The HARQ rate match is used to matches the output bits from turbo encoder to the number of bits in physical channel (HS-PDSCH). HARQ functionality is controlled by Redundancy Version (RV) parameters within two rate match stages as shown in figure below [3].

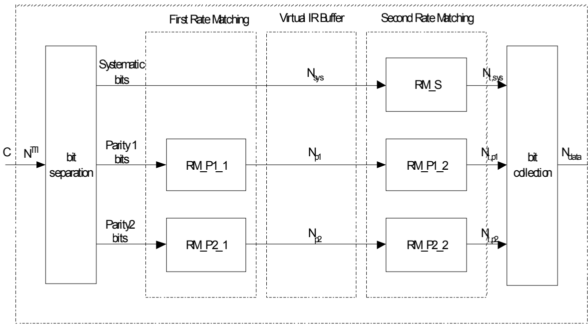


Figure 3 HARQ functionality in HSDPA

e. HSDSCH Interleaver

The interleaving process is done by divides the HARQ rate match outputs into 32x30 (R2=32 rows, C2=30 columns) matrix with performing inter-column permutation pattern as defined in table below.

Table 1. Inter-Column-Permutation pattern of HSDSCH Interleaver [3]

Number of columns C2	Inter-column permutation pattern < P2(0), P2(1), …, P2(C2-1) >
30	<0, 20, 10, 5, 15, 25, 3, 13, 23, 8, 18, 28, 1, 11, 21, 6, 16, 26, 4, 14, 24, 19, 9, 29, 12, 2, 7, 22, 27, 17>

f QPSK Mapper

The QPSK mapper is used to maps each two bits data into one modulated symbol. There are four symbol levels to represent the combination of two bits as stated below.

s_{QPSK} = 0.7071+0.7071i = 1 ∠ 45°; for bits of '00'
s_{QPSK} = 0.7071-0.7071i = 1 ∠ 135°; for bits of '01'
s_{QPSK} = -0.7071+0.7071i = 1 ∠ 225°; for bits of'10'
s_{QPSK} = -0.7071-0.7071i = 1 ∠ 315°; for bits of '11'

g STBC Encoder 2x2

The output symbols from QPSK mapper is then transmitted with STBC method for two transmitter and two receiver antennas as described in [13]. At a given symbol period, two signals are simultaneously transmitted from the two antennas. The signal transmitted from antenna Tx-1 is denoted by S_0 and from antenna Tx-2 by S_1. During the next symbol period signal $-(S_1)^*$ is transmitted from antenna Tx-1, and $(S_0)^*$ signal is transmitted from antenna Tx-2 where ( )* is the complex conjugate operation.

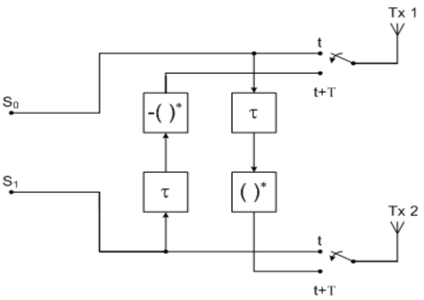


Figure 4. STBC Transmission Method

3.3 HSDPA Receiver

a. Channel Estimation

By comparing the transmitted pilot wih the receive pilot, the receiver can estimate the channel effect by utilizing the STBC orthogonality as following matrix relation.

$$\begin{bmatrix} \tilde{h}_0 \\ \tilde{h}_1 \end{bmatrix} = \frac{1}{|p|^2 + |p|^2} \begin{bmatrix} p & p \\ -p & p \end{bmatrix}^H \begin{bmatrix} r_0 \\ r_1 \end{bmatrix} = \frac{1}{2p^2} \begin{bmatrix} p & -p \\ p & p \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \end{bmatrix} \qquad (11)$$

b. STBC Decoder

Figure 7 shows the signal receiving between two receiver antennas.

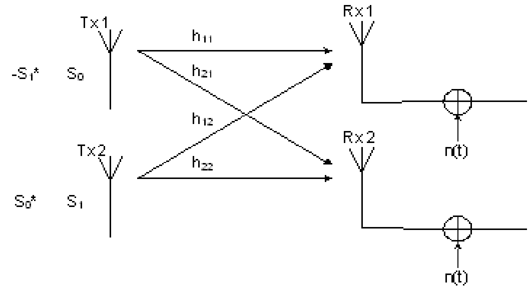


Figure 5 The STBC Receiver

At the t time, Rx 1 and Rx 2 will receive the incoming signal from Tx 1 and Tx 2 through the different paths. The received signal for Rx 1 at the t time is stated as following equation:

$$y_{11} = h_{11}.s_0 + h_{12}.s_1 + n_{11} \qquad (12)$$

While the received signal for Rx 2 is stated as:

$$y_{21} = h_{21}.s_0 + h_{22}.s_1 + n_{21} \qquad (13)$$

At the $t+T$, the received signal for Rx 1 is stated as the following equation:

$$y_{12} = -h_{11}.s_1^* + h_{12}.s_0^* + n_{12} \qquad (14)$$

While the received signal for Rx 2 At the $t+T$ is stated as:

$$y_{22} = -h_{21}.s_1^* + h_{22}.s_0^* + n_{22} \qquad (15)$$

The channel response h_{11}, h_{12}, h_{21}, h_{22} is got after the channel eatimation process. Then it will be combined with the received signal at time of t and $t+T$ according to the following equation:

$$\tilde{s}_0 = h_{11}^*.y_{11} + h_{12}.y_{12}^* + h_{21}^*.y_{21} + h_{22}.y_{22}^* \qquad (16)$$

$$\tilde{s}_1 = h_{12}^*.y_{11} - h_{11}.y_{12}^* + h_{22}^*.y_{21} - h_{21}.y_{22}^* \qquad (17)$$

c. QPSK Demapper

The received signal from STBC Decoder is transformed into log-likelihood QPSK symbols by this soft demapper.

d. HSDSCH Deinterleaver

This block will rearrange the data output from QPSK demapper into the same data sequence in interleaver input at the transmitter side.

e. Rate Dematch

The function of Rate Dematch is to conduct the rearrangement of codewords from turbo encoder which have been matched with the bit amount in HS-PDSCH (expressed in U). The valuable U is 960 bits for QPSK modulation. Parameter used in Rate

Dematch is identical with the parameter used in the Rate Match.

f. Turbo Decoder

The turbo decoder is used to decodee the sequence of received Parallel Concatenated Convolutional Code (PCCC). The decoder used in this final thesis is constant log-MAP with the input is in form of log-likelihood ratio (LLR).

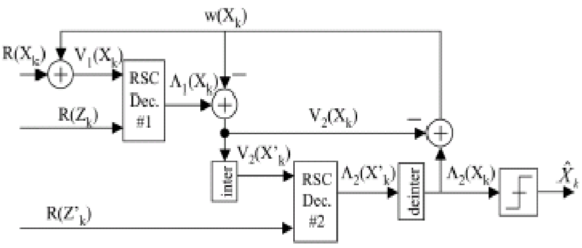


Figure 6. Structure of HSDPA Turbo Decoder [18]

g. Bit Descrambler

To get the unscrambled data hence the bit output from turbo decoder is passed to the Bit Descrambler. The descrambling process is conducted with the same method at Bit Scrambler as following relation:

$$\text{Unscramble data} = (\hat{X}_k + y_k) \bmod 2 \qquad (18)$$

k=1,2,..,B

Where $\hat{X}_k$ is the output from turbo decoder and y_k is the scrambling sequence generated in Bit Scrambler.

3.4 Simulation Parameter

Table 2. Simulation Parameter

Parameter	Nilai
User equipment class	UE category 12 (QPSK Only)
Number bit per frame	3202 bits
Number transmitted frame	1000 frames
Transmission Time Interval (TTI)	2 ms
Inter-TTI Distance	1 TTI
BER Target	10^{-5}
Sperading factor	16
Carrier Frequency	2.1 GHz
Chip rate	3,84 Mcps
Channel coding	Turbo Code; Rate = 1/3
Packet Retransmission	HARQ with *Increment Redundancy*
Relative Mean Power	[0 -10] dB
Relative Path Delay	[0 976] ns
User Velocity	0 kmph, 3 kmph, 10 kmph
Doppler Frequency	0 Hz 5.8 Hz 19.4 Hz
UE delay	25 ms
Node B delay	15 ms
Iub delay	10 ms
RNC delay	10 ms
Iu-Core delay	3 ms

4. Simulation Result Analysis

4.1 SISO performance in HSDPA retransmission

Figure 7 shows the SISO HSDPA performance using 4 times transmission. Retransmission in HSDPA is possible if there is wrong received data packet at the User Equpment (UE) side. This simulation uses only maximum 4 times transmission for each data frame.

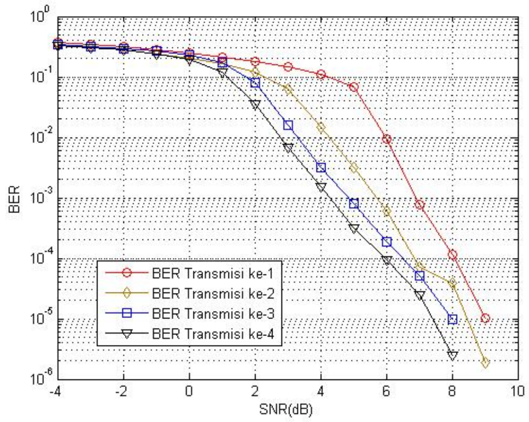


Figure 7 (a) SISO BER retransmission in fd=0 Hz

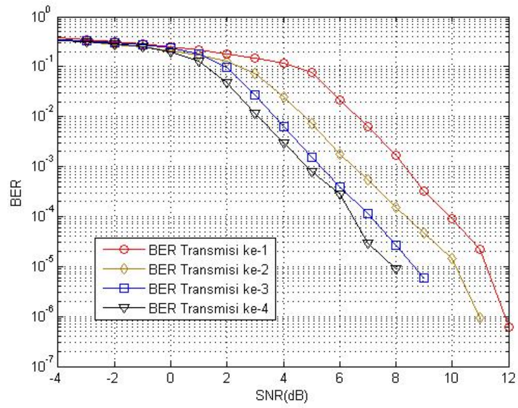


Figure 7 (b) SISO BER retransmission in fd=5.8 Hz

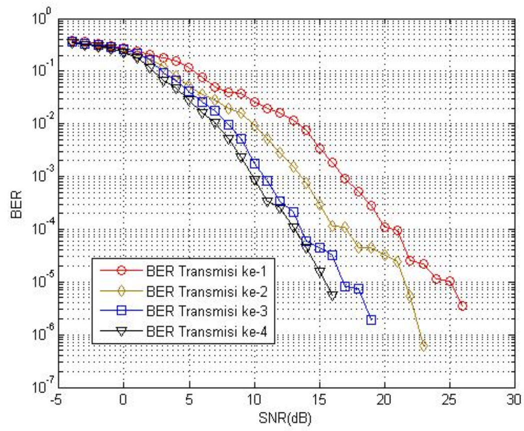


Figure 7 (c) SISO BER retransmission in fd = 19.4 Hz

From the BER to SNR graphs above can be seen that the use of retransmission in HSDPA system gives the SNR improvement in each channel condition. It is possible because the HARQ retransmission controlled by RV (Redundancy Version) may increase the probability of decoding data packet correctly in turbo decoder. The delivered data packet after the first transmission will get the redundant bits, so for the next transmission. The more retransmission take palace, the more delay between UE and Node B. Table 3 shows the SNR comparison for each transmission in SISO HSDPA.

Table 3. SNR for BER 10^{-5} in SISO HSDPA

Doppler Frequency	SISO HSDPA SNR for BER 10^{-5} (dB)			
	1st Trans	2nd Trans	3rd Trans	4th Trans
fd = 0 Hz	9,2	8,5	8	7,4
fd = 5,8 Hz	11,4	10,2	8,8	8
fd = 19,4 Hz	24	21,5	16,9	15,5

4.2 STBC 2x2 Performance in HSDPA Retransmission

Figure 8 shows the graphs of STBC 2x2 performance using 4 times transmission. The STBC 2x2 HSDPA is conducted in selective fading channel with independent and identically Rayleigh distributed (i.i.d).

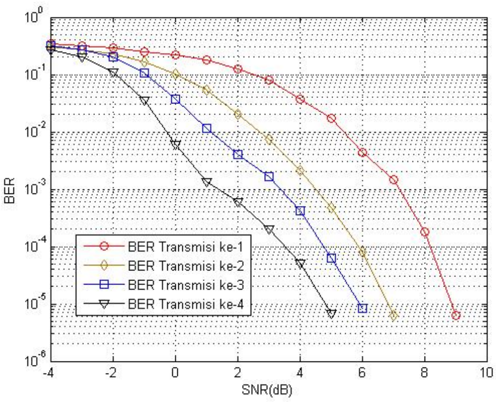


Fig 8 (a) STBC 2x2 BER retransmission in fd = 0 Hz

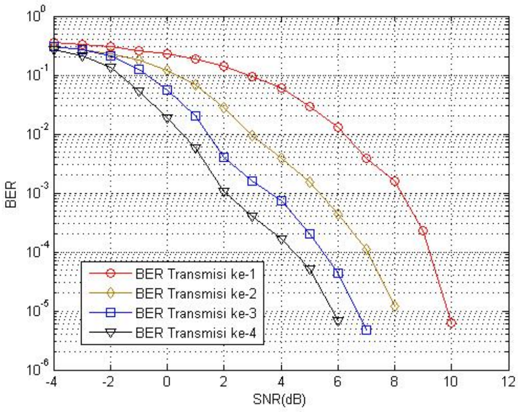


Fig 8 (b) STBC 2x2 BER retransmission in fd = 5.8 Hz

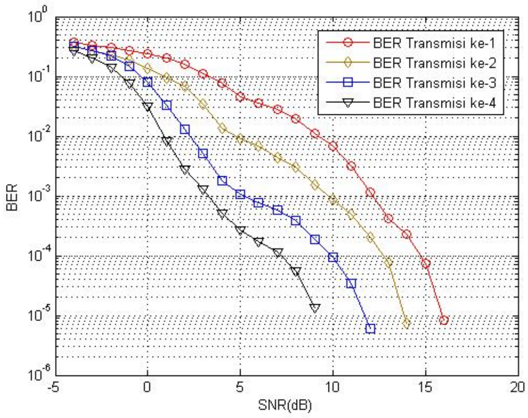


Fig 8 (c) STBC 2x2 BER retransmission in fd = 19.4 Hz

From the graphs BER to SNR in figure 8 can be seen that the use of MIMO STBC 2x2 in HSDPA retransmission gives the SNR improvement to the SISO (figure 7) for each transmission. It is possible because MIMO STBC 2x2 offers gain diversity as the improvement in transmitter side depend on the number of used antennas at transmitter and receiver. This simulation uses only 2 transmitter antennas and 2 receiver antennas which is integrated with HSDPA system. Table 4 shows the SNR comparison for each transmission in STBC 2x2 HSDPA and table 5 shows the SNR improvement for STBC 2x2 HSDPA to the SISO HSDPA.

Table 4. SNR for BER 10^{-5} in STBC 2x2 HSDPA

Doppler Frequency	STBC 2x2 HSDPA SNR for BER 10^{-5} (dB)			
	1st Trans	2nd Trans	3rd Trans	4th Trans
fd = 0 Hz	8,9	6.9	5,9	4.9
fd = 5,8 Hz	9,9	8	6,8	5,8
fd = 19,4 Hz	16	13,9	11,7	9

Table 5. SNR improvement for STBC 2x2 to SISO

Doppler Frequency	SNR improvement for each Transmission (dB)				Average Improvement (dB)
	1st Trans	2nd Trans	3rd Trans	4th Trans	
fd = 0 Hz	0,3	1,6	2,1	2,5	1,625
fd = 5,8 Hz	1,5	2,2	2	2,2	1,975
fd = 19,4 Hz	8	7,6	5,2	6,5	6,825

4.3 AMC performance in Throughput SISO HSDPA and STBC 2x2 HSDPA

Throughput in this simulation is defined as the number of correct received bits from the transmitted frames divided by the number of Transmission Time Interval (TTI) from the transmitted frames including TTI from retransmitted frames. The Adaptive Modulation and Coding (AMC) is expressed by changing the length of Tranport Block Size (TBS) in one frame based on the SNR and RV parameter in HARQ rate match for retransmitting packet. The

TBS and RV parameter will determine the effective code rate of HSDPA. This simulation is conducted in selective Rayleigh fading channel with the Doppler frequency 0 Hz.

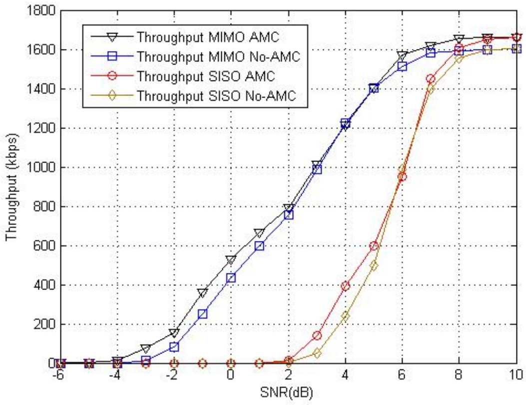


Figure 9 Throughput comparisons of SISO HSDPA and STBC 2x2 HSDPA

From figure 9 can be seen that the use of AMC in HSDPA may increase the efficiency for transmitting data packet. It means that the delivery packet in less ideal channel condition (low SNR) is preffered to send small frame length with low code rate (high addition redundancy). Figure 10 shows that 384 kbps throughput in SISO-AMC is achieved in 3.7 dB SNR but the SISO-Non AMC achieves the 384 throughput at SNR of 4.7 dB. The use of MIMO STBC 2x2 in HSDPA is reducing the tranmit power compared with SISO in achieving the same throughput. In this case there is 3 dB differences between the SISO and STBC 2x2 HSDPA when achieving throughput of 1000 kbps.

4.4 AMC performance in RTT SISO and STBC 2x2 HSDPA

Round Trip Time (RTT) in this simulation is defined as the delay of packet traveling from the receiver terminal (UE) through entire network elements to the application server and back to the receiver. The delay assumptions from network elements are the UE delay (25 ms), Node B delay (15 ms), air interface delay (10 ms), Iu + Core Network delay (3 ms). This simulation is conducted in selective Rayleigh fading channel with 0 Hz Doppler frequency.

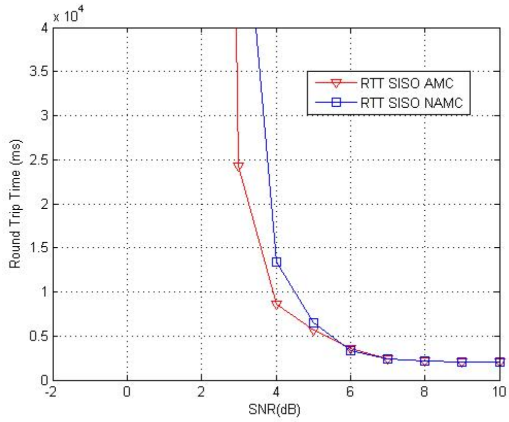


Figure 10 (a) RTT Comparison in SISO HSDPA

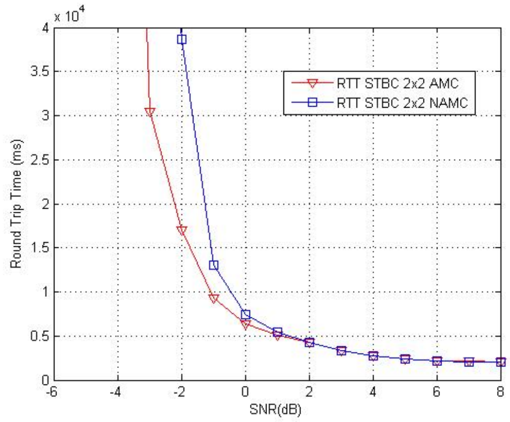


Figure 10 (b) RTT Comparison in STBC 2x2 HSDPA

From figure 10 can be seen that the RTT a packet is determined by the throughput at the certain SNR. The SISO-Non AMC RTT with SNR less than -2 dB is towards infinity. The STBC 2x2 Non AMC also achieves the RTT towards infinity at SNR below -4 dB. It possibly happened because the reached throughput at that SNR is 0 kbps. The increasing throughput because of the AMC gives the better performance than the Non-AMC system. The use of diversity which is represented by STBC 2x2 also gives better performance than the SISO. It means that the order of RTT for the HSDPA system using MIMO and AMC will be more lower compared to which only use one of the AMC and MIMO or not at all.

4.5 STBC 2x2 and SISO Performance in Multi User HSDPA

Multi user in this simulation mean that there are ‘n’ numbers of user in a HSDPA cell but only a user which is be investigated for the test performance. The (n-1) users will act as the interferer for the investigated user. This simulation uses maximum 4 times transmissions. Figure 11 shows the BER comparison for each transmission with n = 4 users and n = 16 users. This simulation is conducted in selective Rayleigh fading channel with the Doppler frequency 0 Hz.

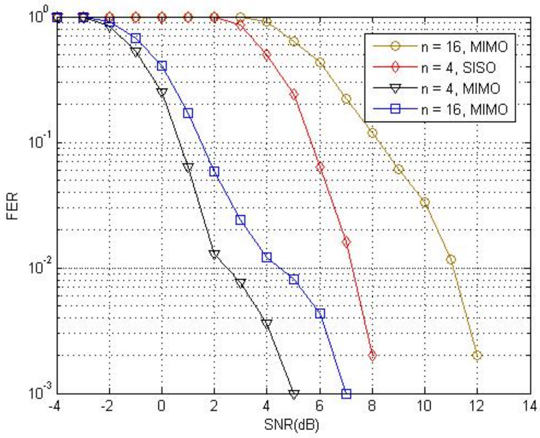


Figure 11 (a) SISO BER in Multi user HSDPA

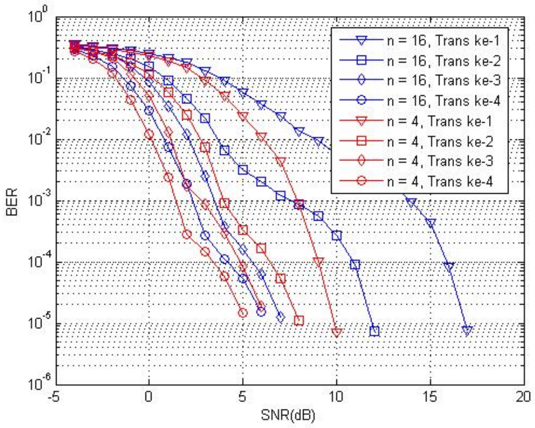


Figure 11 (b) STBC 2x2 BER in Multi user HSDPA

From figure 11 can be seen that the addition of number users will result the dedradation of system performance. It is possible because the increasing number of users will cause the SNR shifting to achieve the same target BER as in the single user system. It means that the interferences are becoming tha additive noise in received SNR. The SNR comparison in SISO multi user and STBC 2x2 multi user. The SNR improvement of multi user STBC 2x2 HSDPA to multi user SISO HSDPA is shown in table 6.

Table 6. SNR IMPROVEMNET OF STBC 2X2 TO SISO IN MULTI USER HSDPA

Number of User (n)	Perbaikan SNR per Transmisi (dB)				Improvement Average (dB)
	1st Trans	2nd Trans	3rd Trans	4th Trans	
n = 4	0,1	0,5	1,8	2,3	1,175
N = 16	1,1	3,5	6,4	5,4	4,1

4.6 Throughput comparison for STBC 2x2 and SISO in multi user HSDPA

The throughput calculation for multi user HSDPA system is done with assumption that the existence of entire users is ruled is Round Robin (RR) scheduling. This scheduling distributes the same physical resources to entire users in a cell. Figure 12 shows the throughput comparison for n = 4 users and n = 16 users. This simulation is conducted in selective Rayleigh fading channel with the Doppler frequency 0 Hz.

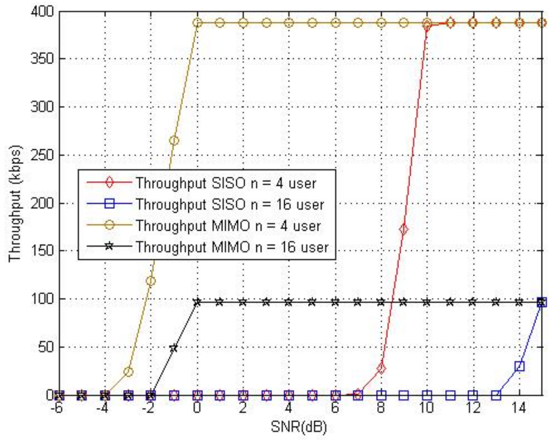


Figure 12 Throughput comparisons in multi user HSDPA

From figure 12 can be seen that by dividing physical resources to the entire user causes the throughput degradation significantly compared to the single user system. The achieved maximum throughput for n = 4 users is less than 1660/4 kbps by assuming that the maximum cell throughput is the maximum throughput for single user system (1660 kbps). For n =16, the maximum achieved throughput is less than 1660/ 16 kbps. However the use of MIMO STBC 2x2 in HSDPA reducing the power needs to get the same throughput as in SISO system.

5. Conclution

1. STBC 2x2 HSDPA have better performance with improvement average 2,63 dB for each transmission while SISO HSDPA only reach 0,9 dB in doppler frequency of 19,4 Hz.
2. The STBC 2x2 also reduces the power about 4,9 dB for achieving 384 kbps HSDPA throughput compared with SISO HSDPA in 0 Hz doppler frequency

References

[1] 3GPP TS 25.101 v6.6.0 Release 6. "Technical Specification Group Radio Access Network; User Equipment (UE) radio transmission and reception (FDD)". 2006

[2] 3GPP TS 25.211 v6.7.0 Release 6. "Technical Specification Group Radio Access Network; Physical channels and mapping of transport channels onto physical channels (FDD)". 2005

[3] 3GPP TS 25.212 v6.8.0 Release 6. "Technical Specification Group Radio Access Network;

Multiplexing and channel coding (FDD)". 2006

[4] 3GPP TS 25.214 v6.9.0 Release 6. "Technical Specification Group Radio Access Network; Physical layer procedures (FDD)". 2006

[5] Berger, Lars Torsen. "Performance of Multi-Antenna Enhanced HSDPA". 2005

[6] Haichuan, Zhao and Wu Jianqiu. "Implementation and Simulation of HSDPA functionality with NS-2". 2005

[7] Holma, Harri and Anti Toskala, "WCDMA for UMTS", John Wiley & Sons. 2002

[8] Bannister, Jeffrey, Paul Mather and Sebastian Coope. "Convergence Technologies for 3G Networks". John Wiley & Sons. 2004

[9] Agilent. "Concepts of High Speed Downlink Packet Access: Bringing Increased Throughput and Efficiency to W-CDMA". 2005

[10] Wahono, Tri dan Yusep Rosmansyah. "Teknologi HSDPA dalam Infrastruktur Telekomunikasi". 2005

[11] D. McCarthy. "UMTS Evolution to High Speed Downlink Packet Access". 2003

[12] Agere Systems. "HSDPA Mobile Broadband Data A Smarter Approach to UMTS Downlink Data". 2005

[13] Alamouti, S.M, "*A Simple Transmit Diversity Technique for Wireless Communication*", IEEE Journal on Selected Areas in Communication, vol 16 No.8, October 1998

[14] Laiho, Janaa and Achim Wacker. "Radio Network Planning and Optimisation for UMTS Second Edition". 2006

[15] F. Brouwer, et al, "Usage of link-level performance indicator for HSDPA network-level simulations in UMTS", Proceedings of IEEE ISSSTA, 2004.

[16] Holma, Harri and Anti Toskala, "HSDPA/HSUPA for UMTS", John Wiley & Sons. 2002

[17] R. Mishra, Ajay. "Advanced Cellular Network Planning and Optimisation". John Wiley & Sons. 2007

[18] Valenti, Mathew C. And J. Sun. "The UMTS Turbo Code and an Efficient Decoder Implementation Suitable for Software-Defined Radios". International Journal of Wireless Information Networks, Vol. 8, No. 4, October 2001.

THE PERFORMANCE ANALYSIS OF COMBINED MUD DECORRELATOR AND PIC IN DS-CDMA

Rahmawati Amalia Kartini[1], A. Ali Muayyadi[2], Uke Kurniawan Usman[3]

[1,2,3] Electrical and Communication Engineering Faculty of IT Telkom
[1]amaliakartini@gmail.com, [2]aly@ittelkom.ac.id, [3]uku@ittelkom.ac.id

Abstract

One major problem which can seriously affect the capacity of CDMA system, is the presence of *multiple access interference (MAI)*. MAI appears when the *cross correlation* value between two *spreading codes* is high meaning not *orthogonal* each other. This problem occurs when several transmitted signal experience the propagation phenomenon such as reflection and scattering caused by many objects within the mobile radio environment which break the orthogonality among codes. Multiuser detection is a method which is applied to improve the performance of the receiver in cellular network. This method reduces the MAI effect. There are many types of multiuser detection which can be classified into *Successive Interference Cancelation (SIC), Parallel Interference Cancelation (PIC), Decorrelator, MMSE*, etc. PIC is the simplest suboptimum multiuser detection in many applications. In this paper, the combination between decorrelator and PIC multiuser detection scheme is shown. PIC consists of several stages where the previous stage's result affects significantly to the decision in the next stage. The combination of decorrelator and PIC in the first stage can improve the performance of PIC receiver. This research also simulates a DS-CDMA system having the number of user : 1, 3, 5 and 8 users. Particularly this research uses two different conditions, i.e. synchronous and asynchronous, and with two different spreading codes, i.e. walsh-hadamard, and m-sequence. The simulation result shows that the performance of combination between decorrelator and PIC is better. The BER performance of this system is 10^{-5} at 23 dB of SNR, whereas for PIC and decorrelator is 10^{-4} and 10^{-1} respectively at the same SNR. The improvement given by the combination of decorrelator and PIC can only take effect for asynchronous condition. On the other side, by varying the number of user, we can conclude that the system performance is poorer when the number of users is increased.

1. Introduction

Code Division Multiple Access (CDMA) technology enables multiuser transmission using the same frequency at the same time. However, the CDMA capacity is reduced by *multiple access interference* (MAI).

The multiuser detection (MUD) is proposed to combat this MAI in CDMA. Generally, these MUD techniques are divided into optimum and suboptimum MUD. Subobtimum MUD is less complex and provide near optimum BER performance. Suboptimum MUD consist of linear (MMSE and decorrelator) and nonlinear (SIC and PIC) suboptimum. This research tried to analyse the performance of combined decorrelator and PIC MUD.

The principle of CDMA allows continues signals transmission occupying the whole frequency at the same time. Each user is assigned with a unique code sequence so the receiver is able to distinguish among users. The unique code assignment allow the set of users transmit by occupying the same frequency.

In radio channel, the multipath signals reflected and scattered due to some obstacles between transmitter and receiver. This phenomenon so called *dispersive multipath propagation* lower the cross correlation among the codes such that it create more intersymbols interference (ISI) which result in poor BER performance.

2. System Model

The spreading sequence used in the research here are PN sequence and walsh-hadamard sequence.

2.1 PN Code

The CDMA transmitter applies a unique code sequence which could only be recognized by intended receiver. One of these codes is so called pseudo-random or pseudo-noise (PN) code.

CDMA alocate a unique PN sequence for each users. The sequences should have meet some requirements, such as :

- The codes allocation for each user should have same length.
- *Orthogonal/nearly-orthogonal* characteristics should have been achieved.

PN sequences are mostly applied in the CDMA transmission. This is due to their simple generation just using LSFR (linear feedback shift register).

2.2 Walsh Code

The advantage of using *Walsh* code is to achieve better orthogonality. Besides, the *walsh* code is relatively simple to be created by invoking *hadammard* matrix H_n. Hadammard matrix is *n x n* orthogonal matrix, where $n=2^m$,m is a positive integer number. It consists of +1 and -1, where +1 represents 1s and -1 represent 0s. The general form of *Hadammard* matrix is:

$$H_{2n} = \begin{bmatrix} H_n & H_n \\ H_n & \overline{H_n} \end{bmatrix} \quad (1)$$

2.3 Massage Recovery

In transmitter side, the data bits *k* is multiplied by code sequence c_k (t). Then the multiplied datas modulate the carrier sinusoida to give the transmitted signal as follows:

$$s_k(t) = Ad_k(t)c_k(t)\cos(\omega_c t) \quad (2)$$

ω_c is frequency in rad/sec and A is the carrier amplitude. At receiver side, the set of signals envelope are combined. By ignoring the AWGN notation first, the receive signal equation is as follows:

$$r(t) = \sum_{k=1}^{K} s_k(t-\tau_k) \quad (3)$$

τ_k noted the propagation delay from k^{th} user transmitter towards DS-CDMA receiver.

At the receiver, the received signals *r(t)* of k^{th} user are multiplied again by similar PN code $c_1(t)$ in order to despread the signal spectrum. However, other users bandwidth spectrums are not despread. This is due to nearly-orthogonal characteristics among codes:

$$\hat{s}_1(t) = \sum_{k=1}^{K} s_k(t-\tau_k)c_1(t-\tau_1') \quad (4)$$

The despread signal demodulate carrier sinusoida and proceed to detector and thresholding device. The detector output is:

$$z_1(t) = \int_{t_1}^{t_1+T} \sum_{k=1}^{K} s_k(t-\tau_k)c_1(t-\tau_1')\cos(\omega_c t+\theta')dt \quad (5)$$

In synchronous transmission, the starting of the first bit among users is identic. Assuming $\tau_k = 0$ dan $\theta' = 0$, we can obtain:

$$z_1(t) = \int_{t_1}^{t_1+T} \sum_{k=1}^{K} s_k(t)c_1(t)\cos(\omega_c t)dt \quad (6)$$

which can be equal as:

$$= \frac{A}{2}\int_{t_1}^{t_1+T}\left[d_1(t)c_1^2(t) + d_2(t)c_2(t)c_1(t) + \ldots + d_k(t)c_k(t)c_1(t)\right]dt \quad (7)$$

The codes among users have the orthogonaliy characteristic, if satisfy the following equation:

$$\int_0^T c_i(t)c_j(t) = \begin{cases} T \rightarrow i = j \\ 0 \rightarrow i \neq j \end{cases} \quad (8)$$

So when we integrate the correlator outputs, other users signal amplitude goes to zero in theoretical, except for the intended user signal. For example, when $d_1(t) = \pm 1$ and $\int_0^T c_1^2(t) = T$, hence:

$$z_1 = \pm\frac{AT}{2} \quad (9)$$

The set of PN sequence did not perfectly uncorrelated among them. These high values of cross correlation, lower the system performance. Cross correlation is defined as:

$$\int_0^T c_i(t)c_j(t)dt = \begin{cases} T \rightarrow i = j \\ R_{ij\neq 0} \rightarrow i \neq j \end{cases} \quad (10)$$

By substituting (10) into (6) we obtain:

$$z_1(t) = \frac{A}{2}\left[\pm T \pm TR_{12} \pm TR_{13} \pm \ldots \pm TR_{1k}\right] \quad (11)$$

the first term is the intended massage and the rest are the MAIs.

2.4 Multiuser Detection

Multiuser detection (MUD) is the MAI handling method in order to protect the system performance. MUD is needed for CDMA based multiple accesses working on multiuser environtment to cover the correlator or matched filter which only be able to solve the problem in single user environtment. In this research, MUD controls the information of spreading sequence owned by uplink receiver.

2.4.1 Decorrelator

The class of suboptimum MUD making the decision independently with respect to the intended receiver. Decorrelating processes on receiver do not affect the decorrelating process of other user and also not necessarily search for strongest amplitude.

In principle, it control the entire information of user's code cross correlation for interference rejection. The corresponding *cross correlation* value forms the *cross correlation* matrix, that is:

$$R = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{21} & 1 & \rho_{23} \\ \rho_{31} & \rho_{32} & 1 \end{bmatrix} \quad (12)$$

where $\rho_{i,j}$ is ith and jth user MAI,

$$\rho_{i,j} = \frac{1}{T_c}\int_{-T_c/2}^{T_c/2} C_i(t) \cdot C_j(t)\ dt \qquad (13)$$

The correlator output can also be written as:

$$y_n = RWd_n + n_n \qquad n = 1, 2, \dots, N \qquad (14)$$

Where
d_n = data vectors
n_n = Gaussian noise vector
R = spreading codes correlation matrix, one row for one user
W = diagonal matrix of user's signal energy (K x K)

The decision process is mathematically stated as:

` $$d_n = sign\ (y_n) \qquad (15)$$

Multiplying correlator output with R^{-1}, we obtain decorrelator output as:

$$= W \cdot d_n + R^{-1} \cdot n_n \qquad (16)$$

And finally the detector determines the bit sign using:

$$\grave{d}_n = sign\ (Z_n) \qquad (17)$$

2.4.2 Parallel Interference Cancellation (PIC)

Assume the users massage modulate the carrier sinusoida using BPSK, the transmitted signal is $x_k(t)$ with the amplitude A_k, bit duration is T_b and the *processing gain* N. Hence we can write the transmitted signal as:

$$r(t) = \sum_{k=1}^{K} x_k(t - \tau_k) + n(t) \qquad (18)$$

where τ_k represents the propagation delay for k^{th} user and n(t) represents the component produced by AWGN with PSD equal to $N_0/_2\ (W/Hz)$.

PIC MUD comprises several stages. Within each stage, $1 \le s \le S$, the j^{th} interferer to k^{th} interferer, $x_j^{(s)}(t)$ are regenerated. Then, the regenerated interferers relative to k^{th} user are cancelled from total received signal, so we obtain:

$$\grave{r}_k^{(k)}(t) = r(t) - \sum_{\substack{j=1 \\ j \neq k}}^{K} \grave{x}_{k,j}^{(s)}(t) \qquad (19)$$

The detector output for k^{th} user on s^{th} is given by:

$$Z_{k,i}^{(s)} = \int_{T_b+\tau_k}^{(i+1)T_b+\tau_k} \grave{r}_k^{(s)}(t) \cdot S_k(t - iT_b - \tau_k)\, dt \qquad (20)$$

where $S_k(t)$ is a *signature sequence* of k^{th} user. The s^{th} stage PIC of the i^{th} bit and k^{th} user is given by:

` $$b_{k,i}^{(k)} = sgn\ (Z_{k,i}^{(s)}) \qquad (21)$$

Sgn means signum function.

The MUD technique is more potential to be applied on base station rather than on mobile station due to its complex implementation and high cost.

The DS-CDMA MUD models herein is illustrated in Figure 1 as follows:

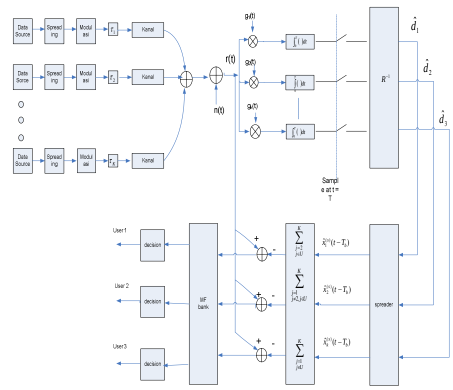


Figure 1. System Model

3. Simulation Results And Analysis

The simulation of the above transmitter and receiver is performed using Matlab. The results are given as follows:

3.1 Performance comparison with or without MUD

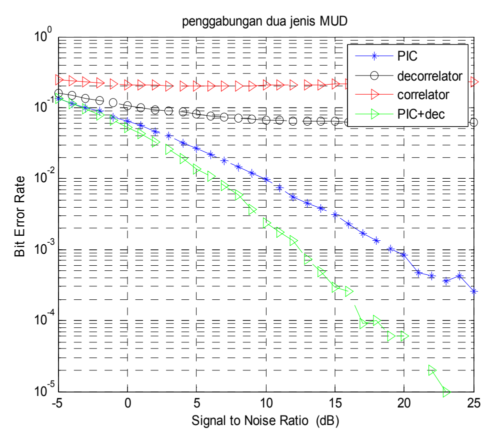
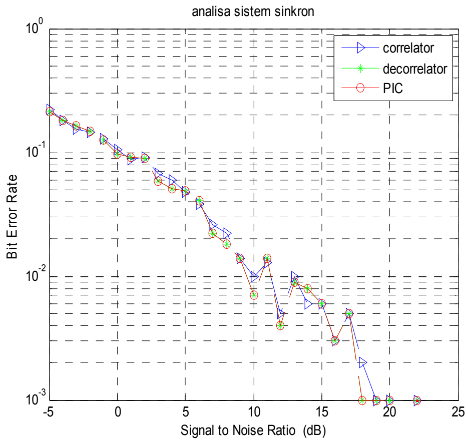


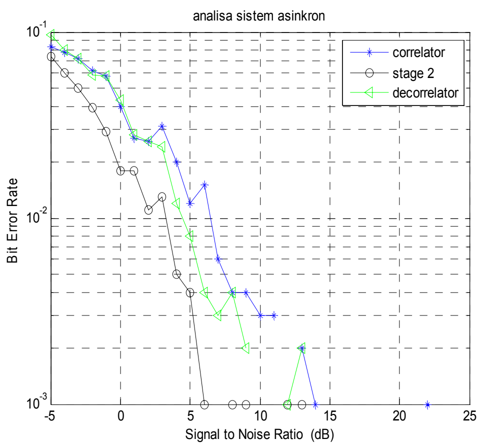
Figure 2. DS-CDMA Performance Comparison With and Without MUD

Figure 2 shows that the implementation of MUD combined decorrelator and PIC provides better BER performance than the others. For the constant SNR value, that is 23 dB, the combined decorrelator and PIC achieve 10^{-5} BER. On the other hand, single PIC achieves 10^{-4} BER and 10^{-1} BER for single decorrelator.

Decorrelator belongs into suboptimum MUD. According to the simulation result, the decorrelator is unable to provide better performance in asynchronous transmission. However, PIC still in asynchronous trasnsmission. PIC comprises several stages and the decisions of previous stages affect the performance of the next stages. However the single PIC performance is still lower compared to the performance of combined decorrelator and PIC MUD.



(a) Synchronous Trasmission



(b) Asynchronous Transmission

Figure 3. The Performance Comparison In Synchronous And Asynchronous Transmission

Fgure 3(a) shows that the combined decorrelator and PIC provides only a little bit BER improvement in synchronous transmission. However, this scheme shows more BER improvement in asynchronous transmission as shown in Figure 3(b).

Due to better interference rejection of decorrelator in the initial stage, the PIC improvement does not show any significant contribution. The scenario is quite different in asynchronous transmission when decorrelator unable to perfectly reject the interference in asynchronous transmission. On top of that, the second stage is inserted with PIC to reduce the problem as shown in Figure 3(b)..

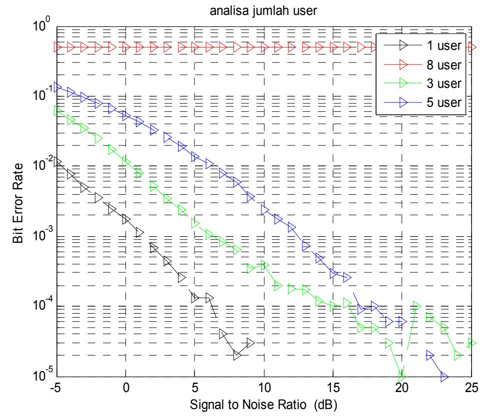


Figure 4. The Performance Comparison by Varying the Number of Users

The Figure 4 above shows that the increment of users deteriorates the BER performance. In single user condition without interference, 10^{-5} BER is obtained for 8 dB SNR requirement. For 2 and 4 interferer, 20 and 24 dB SNR is needed to reach the similar SNR. When 7 interferers are presence, the BER performance deteriorates badly. In other words, the increment of interferers should be handled with care.

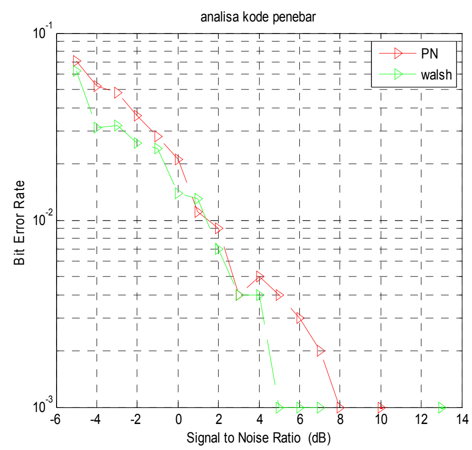


Figure 5. The Spreading Codes Performance Comparison

The selection of spreading codes also affects the system performance, as shown in Figure 5. The Walsh code performance surpasses PN m-sequence performance, denoted by 10^{-3} BER at 7 dB SNR, 3 dB less than m-sequence performance. The cross correlation property of Walsh code is better than m-sequence in synchronous transmission. It is the reason why it provides better performance in multiuser environtment.

4. Conclusion

According to the above resultas and discussion, we can conclude that:

1. Parallel interference Cancellation is a multistage multiuser detector. The combination of decorrelator and PIC is implemented through the insertion of decorrelator as an initial stage of PIC in order to obtain accurate estimation.
2. According to simulation result, it is shown that by combining PIC and decorrelator, we obtain better performance than by using single decorrelator or single PIC.
3. In synchronous transmission, when every user trasmit the massage at the same time, the combining method did not produce significant improvement compared with asynchronous transmission.
4. The spreading codes selections also affect the system performance of combined PIC and decorrelator.
5. The more the numbers of users transmit the massage, the less system performance due to the MAI.

To obtain more improvement, one can use the following considerations:

1. Using other way of combining, such as MMSE with PIC or SIC with PIC.
2. Implementing the combined MUD for the other systems, for example MC-CDMA, MC-DS-CDMA, etc.
3. Adding ADC and DAC for real time application trials.
4. Applying the combined MUD in DSP card.

References

[1] A.Bentrcia, A.U.Sheikh. and A.Zerguine."Real-Time Implementation Of The Decorrelator Detector".Telecommunication Research Lab, Department of Electrical Engineering.Saudi Arabia.

[2] CHO, Bong Youl. Nonmember and LEE, jae Hong."Nonlinear Parallel Interference Cancellation with Partial Cancellation for a DS-CDMA System".IEICE trans.Commun.,Vol.E83-B.2000.

[3] Correal, Neiyer S. Buehrer, Michael. And Woerner Brian D.,"ADSP-Based DS-CDMA Multiuser Receiver Employing Partial Parallell Interference Cancellation". IEEE journal in Communications, vol.17, No.4.1999.

[4] Haykin, Simon."Communication System". Wiley. USA.2001.

[5] L.Hanzou. L-L Yang.E-L Kuan. And K.Yen."Single-and Multicarrier DS-CDMA Multiuser Detection, Space-Time Spreading, Synchronisation and Standards".Wiley. England.2003.

[6] Mandayam, Narayan B. and Verdu, Sergio."Analysis of an Approximate Decorrelating Detector". The 33rd Annual Allerton Conference on Communication, Control, and Computing.1995.

[7] Nithyanandan,L. and Dananjayan,P.," Analytical BER Computation FOR Interference Cancellation Receiver In MC-CDMA System", Dept. of Electronics and Communication Engineering, Pondicherry Engineering College, Pondicherry-605014.India.

[8] Tikiya,V. Ramesh, A. chockalingan, A.,"Performance Analysis of a Parallel Interference Canceller on Rayleigh Fading Channels", IEEE.2004

DESIGN OF YAGI ANTENNA ON *WIRELESS LOCAL AREA NETWORK 2,4 GHz*

Tito Yuwono

Electrical Engineering Department, Faculty of Industrial Technology, Universitas Islam Indonesia
titoyuwono@yahoo.com, tito@fti.uii.ac.id

Abstract

Wireless Local Area Network is one of the capabilities of computer network technology to switch data wirelessly using air as the mediator. Antenna plays important role in transmitting and accepting the signal on the air. Based on above statement, it is easily found costly antennas appropriate with their ability. This could be emerged as one of the obstacles in developing the technology of WLAN system in Indonesia. Designing and analyzing of Yagi Antenna for WLAN 2,4 GHz system will be discusses in this paper. Yagi Antenna is chosen because its simple construction, good direction, and low cost. From designing process could be concluded that the Yagi antenna is deserved to be considered and applied because its good performance in data transmitting such as SOM, link quality and throughput.

Keywords:Yagi Antenna, WLAN 2,4 GHz

1. Basic Theory of Yagi's Antenna

Antenna is a means to accept and transmit the electromagnetic waves. There are various types of antennas, such as helix antenna, parabola antenna, sector antenna, smart antenna and yagi antenna. From all above antennas, some of them cover specific advantages, which are cheap construction and high direction. Hence, in this paper the performance of yagi antenna on WLAN will be further described and analyzed.

Yagi antenna, theoretically, is an antenna that consists of three elements. They are three elements that play important role in Yagi antenna, as follow: reflector, dipole and directory [1]. In implementation, Yagi antenna could be constructs from elements in both parallel and cylindrical shape.

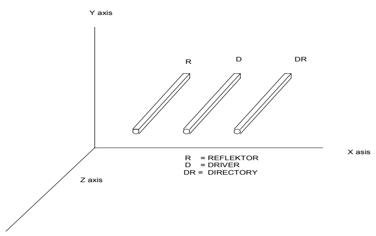


Figure 1. Yagi Antenna on *cartesius* coordinate [1][2]

Above figure shows that yagi's elements placed in line with Z Axis, while the boom (elements supporting) in line with X axis, the simplest Yagi antenna is a antenna with 2 elements consist of one radiator or driven element and one parasitic element as directory.

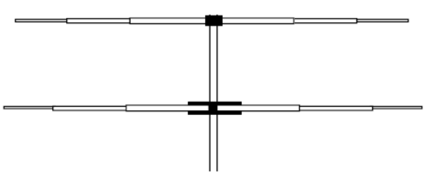


Figure 2. Yagi Antenna

Driver Element

Driver is the most important part of Yagi antenna because it generates the electromagnetic waves to be a transmitted signal. To make it a fine in transmitting radiation, dipole antenna is employed for driver.

Dipole antenna is an antenna in short linear shape, when it is transmitting, it can be considered has the same currents on its all length.

Figure 3. Dipole Antenna

In the making of Yagi antenna driver, dipole antenna that is often used is half-wave dipole antenna with minimum total length on transporter frequency is ½ λ, the using of this antenna is because of its low radiance resistance but high reactance, hence it suitable and efficient for the antenna that has wider wave length. It is shown by the radiation pattern of dipole ½ λ antennas.

The famous modification of dipole ½ λ antenna is folded dipole antenna. For supporting elements; it is preferred more than single dipole antenna because it is cheap, providing higher terminal impedance and has strong mechanical structure.

Each edge of folded dipole antenna will be united. Actually, the dipole antenna is divided into two parts each is ¼ λ, k. The radiation patterns will be exactly the same as single dipole antenna, but produce different radiation resistance because the transmission path is connected to center that has impedance equal to 73 Ω, however because of that combination, radiation resistance turn to be 4x 73Ω or 292 Ω on its feed point, folded dipole antenna is shown on below figure:

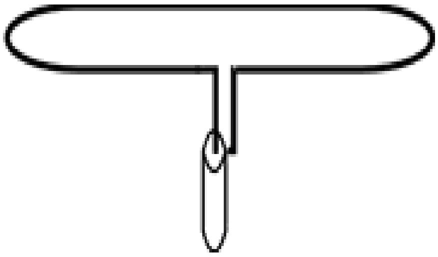


Figure 4. Folded dipole Antenna

Reflector Element

As it called, it is a reflector element, this element is located behind the dipole and was made longer than the dipole.

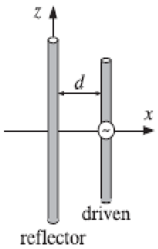


Figure 5. Positions of Driver and Reflector

Main purpose of putting the reflector behind is prohibiting the radiation from extent to the back part while its radiation strength will be strengthened to the opposite part. It also caused the antenna become more inductive.

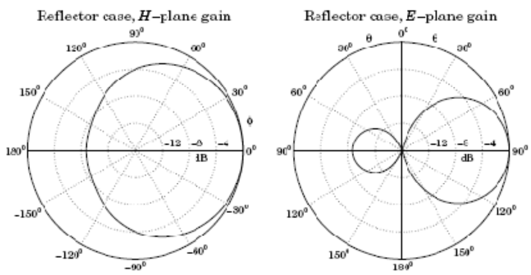


Figure 6. Radiation patterns caused by the reflector's influence

The dimension determination of reflector is concluded by [1] [4]:

$$l_{ref} = l_{dipole} + 5\% l_{dipole} \quad (1)$$

Which : l_{ref} = *reflector's length*

l_{dipole}= *driver element's length*

The reflector installment only conducted once, because the second and third practically won't affect its structure direction. While, the reflector element is placed behind driver element (dipole) with optimum distance 0, 15 – 0, 2 λ.

Directory Element

Directory Element is a directed element that situated in front of folded dipole antenna (driver), directory will force radiation to one direction. This element often called as parasitic element.

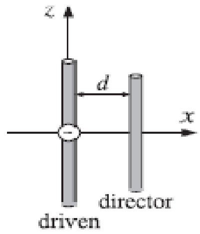


Figure 7. Position of Directory Element

Above figure shows that pattern 1 is radiation pattern that created by folded dipole antenna, with the addition of reflector and directory, antenna's radiation pattern will be changed and minimize to one direction but with longer transmission as shown by pattern 2.

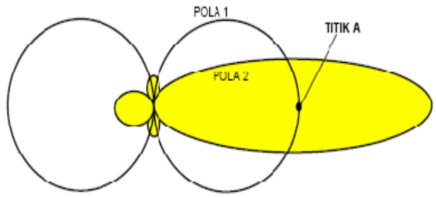


Figure 8. Directed Radiation Pattern

Addition of one or more directories is the most effective method in attaining bigger gain, the more directory elements the bigger gain will be achieved.

As reflector, directory elements also have rules in determining dimension and distance, whether it related with distance with driver or space between one directory to another. Because the performance of Yagi antenna will affected by this measurement.

In determining this dimension, directory should be made smaller than the size of dipole antenna, below equation could be applied to determine the dimension.

$$l_{directory} = l_{dipole} - 5\% l_{dipole} \quad (2)$$

Which are: $l_{directory}$= *directory's length*

l_{dipole}= *driver element's length*

Nearest directory to dipole antenna is the most affected direction to the strength, and the farthest direction, on the contrary. will provide smallest affect.

Several theories state about the equation that says about distance between directories, as follows:

$$d = \frac{36,6}{f} \quad (3)$$

Which are: d = distance between directory (m)
f = work frequency (MHz)

Gain of Yagi antenna

Gain of Yagi antenna is obtained from maximizing its important parasitic elements. In increasing its gain, Yagi antenna change the driver arrangement that won't provide significant affect on its gaining, the most effective way is by conducting right arrangement on element's dimension and distance.

Gain of Yagi antenna is the enhancement of dipole antenna as the driver, theoretically maximum gain of dipole antenna λ/2 are 10 log 1.66 or 2,2 dBi. In this case, the increasing of element factors' amount that rise the gain is showed by below equation:

$$G = 10 \log 1{,}66 \text{ x (nelemen)} \qquad (4)$$

Generally the gain of Yagi antenna will decrease tangibly if the length of reflector is smaller or on the contrary, directory is longer than dipole. Besides, there is another formula state that gain of Yagi antenna could be obtained from figure's analyzing; the equation used in figure's analyzing is described as follows:

$$G = 10 \log \frac{41253}{\theta_H \theta_E} \qquad (5)$$

Which are: G = gain towards isotropic antenna
θ_H = ½ angle of polarization horizontal strength
θ_E = ½ angle of polarization elliptical strength

2. The Design of Yagi Antenna

On this sub chapter, it will be described the design of Yagi antenna on WLAN with 2,4 GHz frequency.

To accomplish wave's length on work frequency, below calculation is provided:

$$\lambda = \frac{300.000.000}{240.000.000} = 0.125m = 12{,}5cm = 125mm$$

Above length is categorized as theoretical length, in fact it is not always match the calculation, usually the antenna's elements are shorter.

3.1 The Design of Driver's Element

As mentioned before, driver is the most important element, because the electromagnetic field will be radiated on air by this element. The design of driver is using the design of dipole antenna that has ½ wave's length, in order to discount the reactive component, the length of ½ dipole antenna should be:

$$0{,}475\lambda = 0{,}475x12{,}5cm = 5{,}93 \approx 5{,}9cm$$

The length of driver could be determined also with below formula:

$$l = \frac{142{,}65}{f} = 0{,}0593m = 5{,}93cm$$

1/4λ= 2,95cm

1/2 λ = 5,93 cm

1/4λ= 2,95cm

Figure 9. Dimension of Dipole antenna

3.2 The Design of reflector's Element

The length of Reflector is determined by:

$$l_{ref} = l_{dipole} + 5\% l_{dipole}$$

$$l_{ref} = 5{,}9 + (0{,}05)5{,}9 = 5{,}9 + 0{,}295 = 6{,}19 \approx 6{,}2cm$$

The placement or position from reflector will also provide gain, reflector's element is placed after driver's element (dipole) with optimum distance 0, 15 – 0, 2 λ. From research, theoretically, best placement is 0.15λ = 0, 15x12, 5cm = 1, 87 *cm,* will be described by below figure

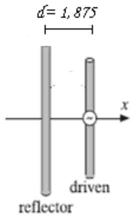


Figure 10. Optimum Placement of Reflector

3.3 The Design of Directory's Element

The length of directory could be determined by:

$$l_{directory} = l_{dipole} - 5\% l_{dipole}$$

l_d= 5,9 *cm* – 0.05*5,9*cm* = *5,9 – 0,295 = 5,6 cm*

As reflector, the placement also plays important role for gaining, distance between each directories are:

$$d = \frac{36{,}6}{f} = \frac{36{,}6}{2400} = 0{,}015m = 1{,}5cm$$

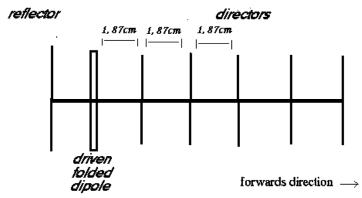


Figure 11. Space Between Directories

On the design of Yagi antenna, it is made 12 directories to obtain enough gain for direction oriented to one point.

The Experiments and Performance's measurement

The experiments are conducted on different distances, which are 20 m, 150 m and 500 m based on below figure.

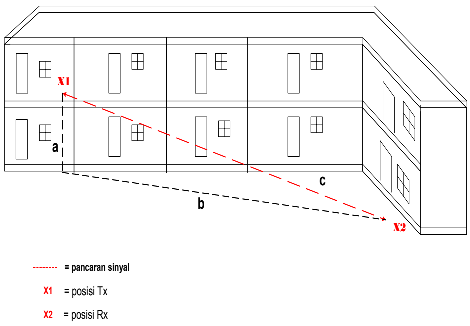


Figure 12. Placement Plan position Rx and Tx with distance of 20 m

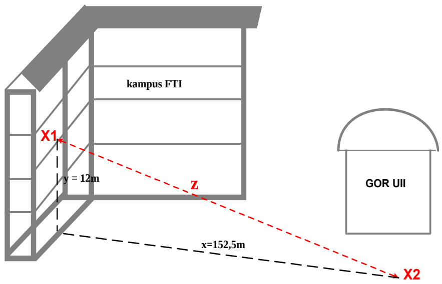


Figure 13. Placement Plan position Rx and Tx with distance of 150 m

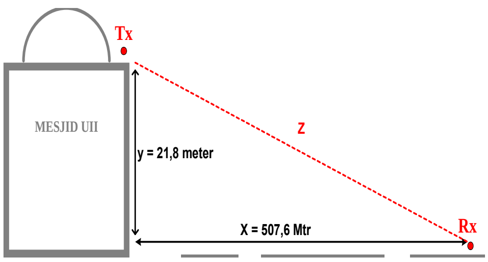


Figure 14. Placement Plan position Rx and Tx with distance of 500 m

Table 1. The experiments' Result

Object of Comparison	Distance of 20 m	Distance of 150 m	Distance of 500 m
SOM	59,72	42,22	15,64
Signal strength	-25	-68	-72
SNR Quality	Maximal = 75	Maximal = 32	Maximal = 32
Duration of data download 10.3 MB	23 second	48 seconds	1 minute 45 seconds

3. Conclusion

After designing and analyzing implementation on WLAN network, could be concluded that:

1. Yagi antenna is a modification between dipole ½ λ antenna with the addition of several elements which are: reflector and director in order to obtain more and focused gain.
2. Yagi antenna is one of recommended antenna on WirelessLAN network; beside its simple construction and higher level of gain which is 14 dB, Yagi antenna also contain an aimed direction.
3. From SOM result or fade margin, concluded that Yagi antenna have met the requirement in WLAN planning, because the value of SOM is bigger than 10 dB, appropriate with theory that states " if distance < 16 km, minimal value of SOM are 10 dB, this will affect wireless system connectivity
4. Yagi's antenna's gain level is influenced by several factors, such as:
   a. Material's quality
   b. Length of elements of driver, reflector and directory
   c. Space between each elements

References

[1] Schure S, *Antennas*, John F.Rider Publisher Inc, New York, 1989

[2] Laawson L, Yagi Antena design, American radio Relay League, Newington

[3] Onno W.Purbo *Internet Wireless dan Hot Spot*, Penerbit Elex Media Computindo, Jakarta

[4] William I, *The Radio Amateur Antenna handbook*, Radio Publication Inc, Wilton ,Conn 1999.

[5] Winarno Sugeng, *Instalasi WirelessLAN*, Penerbit informatika, Jakarta

[6] www.ARRL.com

[7] http://www.andrewhakman.dhs.org/yagi/

[8] http://en.wikipedia.org/wiki/Yagi_antenna

[9] www.ativa.net

[10] www.gemisis.net/wlan/wlanplan.php

ANALYZING INFLUENCE OF EARTH STATION ANTENNA POINTING TO RECEIVING SIGNAL PARAMETER AT DOWN LINK POWER BUDGET

Wahyu Pamungkas[1], Denny Pahlevie[2]

[1,2]Program Studi D3 Teknik Telekomunikasi, Akatel Sandhy Putra Purwokerto
[1]**wahyu_pamungkas@hotmail.com**, [2]**d_zippy06@yahoo.com**

Abstract

One problem caused by reduction energy of power link budget in satellite communications system is the alteration of earth station antenna towards satellite. Any mistake in antenna pointing would affect quality of transmission signal information at Rx station. At this research, friction of angle of pointing only happened at Rx station antenna, while Tx of station antennas leads precisely to satellite. This research will discussed about friction characteristic angle of pointing Rx station antenna that is azimuth and elevation parameters. This research worked by two satellite, TELKOM-1 and TELKOM-2 satellite with the used of channel coding Turbo Product Code (TPC) and errors correction at ¾. At first computing, it is done at two directions calculation power budget with the right pointing towards satellite, so that will be got an initial value in design power link budget desired. Continuously, at Rx earth station position will be simulated to any other position / friction toward satellite. With friction of this pointing, will be seen the influence of receiving downlink parameters power budget by watching result from antenna pattern. This variation of antenna pattern would cause reduction of power received and affected Bit Error Ratio, as result of pointing mistake. From calculation, it concludes that increase far friction of pointing to satellite, would influenced receiving value of signal parameters power budget of downlink.

Keywords: satellite, pointing antenna, azimuth, elevation, friction, power link budget, downlink.

1. Introduction

Basically, the idea of satellite technology has similarity with placing a relay station at upstairs and connected to the other relay station placed at earth. The relay station at sky has a function to strengthen signal from the earth station transmitter and directing signal to the earth station receiver. According to that reality, satellite cannot be possessed as a transmitter or receiver itself, but it can be a relay and signal amplifier from transmitter and receiver from earth station side (ground segment). Satellite Technology Development is not just improve quantity of satellite in orbits, the ability of satellite, size and weight, but it just to the antenna technology such as bigger Effective Isotropic Radiated Power (EIRP) make antenna at ground station has a smaller size, namely VSAT (Very Small Aperture Terminals)

For the GEO (Geosynchronous Earth Orbit) satellite, the position of satellite look like fix from ground segment, and it has easier antenna pointing because it doesn't need to move the antenna following satellite movement if place at LEO (Low Earth Orbit) or MEO (Middle Earth Orbit). The result from pointing process at Ground Segment will influence signal receive strength power, which is one of the main parameters of power budget at downlink side.

2. Research Methodology

1. Observation Study

Done by doing some observation and measuring some parameters that will be influenced by antenna pointing process. These data is taken from Telkom's Satellite Master Control Station at Cibinong, Bogor.

2. Formulation and Parameters From Power Budget (Downlink Direction)

Formulation and some parameters that will be discussed in this journal split into two categories, they are some parameters from power budget calculation that will influence pointing process and some parameters influenced by alteration pointing process.

a) Parameters having an effect on antenna pointing process
    - Noise parameters
    - Free Space Loss
b) Parameters that influenced by antenna pointing process
    - Transmit and receive power from antenna (Tx/Rx level)
    - Gain Antenna received
    - Gain to Noise Temperature
    - Carrier to Noise Ratio (C/N)
    - Eb/No (Energy Per Bit to Noise Density Ratio)

3. Basic Theory

Basic principle from satellite communication is the same with the other radio communication system that using some spectrum frequency. Satellite configurations consist of two parts, one part is called ground segment and another part is called space segment. Ground segment conduct as a transmitter and receiver. As a transmitter ground station, it will transmit some frequency directing to satellite at space segment, this kind of frequency named as uplink frequency. As a receiver ground station, it will receive signal that has been powered by satellite, and frequency from this signal named as downlink frequency. At space segment, satellite receive uplink frequency from ground station transmitter, amplify and re directing the signal to downlink side on ground segment.

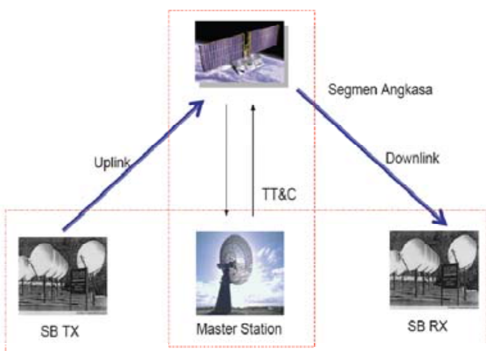


Figure 1. Basic Configuration of Satellite Communication System [3]

A. Ground Segment

The kind of ground segment as a starter part is transmitter ground station sending up link frequency to satellite direction. Ground segment as an end station is receiver ground station who will receive downlink frequency from satellite. As a transmitter ground station, base band information signal is converted into IF signal. Intermediate Frequency (IF) then converted to higher frequency by up converter and amplifies by HPA / SSPA (High Power Amplifier/Solid State Power Amplifier). Signal that has higher frequency and higher power then once again amplified by antenna and directing to satellite at space segment.

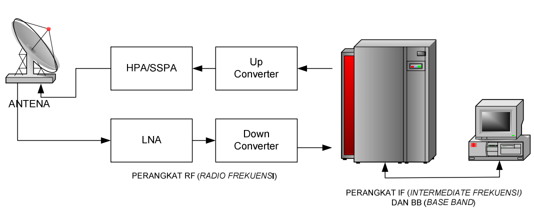


Figure 2. Basic Configuration at Ground Segment

B. Space Segment

Space segment or satellite side is a kind of space station that could only have a function as repeater station placing at outer space named orbit. Repeaters mean that space segment process repeating signal from ground segment, amplified by HPA (High Power Amplifier) and local oscillator from satellite. Band pass filter at space segment has a function to filtering some signal information and split noise from original signal.

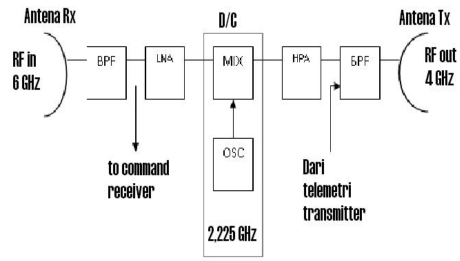


Figure 3. Basic Configuration from Space Segment

C. Parabolic Antenna

Parabolic antenna has some parameters to support its function, those kinds of parameters are:

1. Antenna Gain

Antenna gain is a parameter that describes some gain factors of antenna to electromagnetic signal from transmitter direction or receiver direction. Antenna gain is one of main parameters from satellite communication system, because it can have an effect on EIRP (Effective Isotropic Radiated Power) calculation. The antenna gain can be calculated from:

$$G = \eta . \left(\frac{\pi . D}{\lambda} \right)^2 \quad (1)$$

G = parabolic antenna *gain* (watt)
D = antenna diameter (m)
Λ = wavelength (m) = c/f
H = antenna efficiency (55% ≤ η ≥ 75%)
c = speed of light = $3x10^8$ m/s
f = working frequency (GHz)

If the formula is converted into dB, then it would be:

$$G = 20,4 \times 10 \log \eta \times 20 \log D \times 20 \log f \quad (2)$$

G = parabolic antenna *gain* (dB)

2. Antenna Beamwidth

Antenna beamwidth is a wide spread angel direction from antenna itself. Beamwidth calculated at 3 dB from top of the main lobe at the down direction. Beamwidth is also mean angle at main lobe that has right-left border 3 dB from the top of main lobe. The beamwidth that calculated as big as 3 dB from top of main lobe has a half total antenna gain used in satellite communication system.

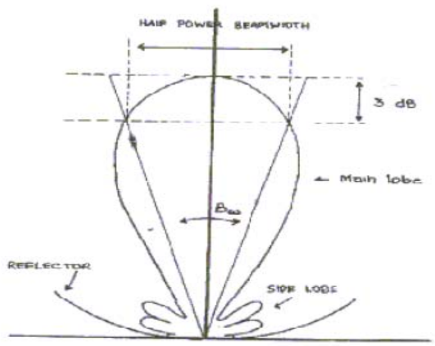


Figure 4. Basic Main Lobe and Side Lobe of Parabolic Antenna **[14]**

To find the beamwidth point of parabolic antenna, we used this formula below:

$$\theta_{3dB} = k \times \left(\frac{\lambda}{D}\right) = k \times \left(\frac{c}{f \times D}\right) \qquad (3)$$

θ_{3db} = beamwidth (degree)
k $\approx$ 70

3. Effective Isotropic Radiated Power (EIRP)

EIRP is a parameter that showing the real power radiated by parabolic antenna influenced by its antenna gain. These EIRP parameters can be calculated from:

$$EIRP_{SBTx} = P_{Tx} - L_{feeder} + G_{Tx} \qquad (4)$$

P_{Tx} = Output power from HPA at ground segment (dB)
L*feeder* = *loss* transmission line (*feeder*) (dB)
G_{Tx} = *gain* from Tx antenna (dB)

4. Power Flux Density Antenna

Antenna Power Flux Density is an antenna parameter that showing value of power signal from Tx side and Rx side, measuring in a unit area. The Value of power flux density antenna can be calculated from:

$$\Phi = EIRP - 10\log\left(\pi . R^2\right) \qquad (5)$$

Φ = *power flux density* (dB/m^2)
R = antenna diameter (m)

D. Pointing Antenna

The position of Tx ground segment and Rx Ground Segment has a main role in satellite communication system. In this case, the position of ground segment antenna must be having a right direction to satellite in order to reduce antenna pointing loss. Pointing angle direction from ground segment to space segment split into azimuth angle direction and elevation angle direction.

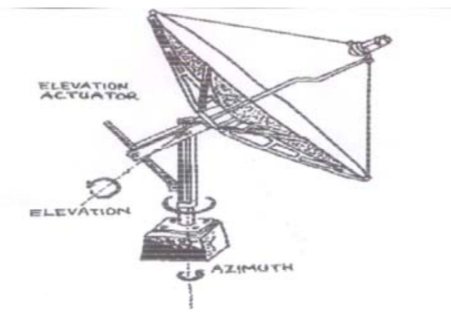


Figure 5. ***Azimuth and Elevation Angle***

1. Azimuth Angle

Azimuth angle direction is an angle between north direction lines of the earth to the position of satellite at latitude. To find the azimuth angle direction we must make some category to ground segment position as we see below:

a. Ground Segment Position at Northern equator; and
   Ground segment at the west side of satellite, we defined:
   A = 180° - A'
   Ground segment at the east side of satellite, we defined:
   A = 180 + A'
b. Ground Segment Position at southern equator; and
   Ground segment at the west side of satellite, we defined:
   A = A'
   Ground segment at the east side of satellite, we defined:
   A = 360 – A'

$$A' = \tan^{-1}\left[\frac{\tan\left|longSB - longSat\right|}{\sin latSB}\right] \qquad (6)$$

θs = *longitude* satellite
θ_L = ground segment *longitude*
$\theta 1$ = ground segment *latitude*
A = Azimuth angle direction

2. Elevation Angle

Elevation angle direction is an angle between a horizon lines at station to line of sight direction from ground segment to space segment. The direction of elevation angle is upside with starting point is derived from the horizon station. Elevation angle can be calculated as:

$$\cos\theta = \left(\mathrm{Re} + h\right)\sqrt{\frac{1 - \cos^2\phi_G \cos^2\Delta\lambda}{h^2 + 2\,\mathrm{Re}\left(\mathrm{Re} + h\right)\left(1 - \cos\phi_G \cos\Delta\lambda\right)}} \qquad (7)$$

$$E = \cos^{-1}\theta \qquad (8)$$

h = (±35786 km)
Re = earth radius (±6378)

φ = differences point between ground segment longitude and space segment longitude
Δ = latitude of ground segment
E = Elevation angle position

E. Slant Range

Slant range position of ground segment relative to space segment is a real distance from ground segment to space segment. It can be calculated by having a straight line direction from ground segment to space segment. This slant range determined the free space loss parameters.

$$D = \sqrt{h^2 + 2.R_E.\left(R_E + h\right).\left(1 - \cos\phi_G.\cos\Delta\lambda\right)} \quad (9)$$

D = slant range ground segment to spaces segment (km)
h = (±35786km)
R = earth radius (±6378km)
φ = differences point between ground segment longitude and space segment longitude
Δ = latitude of ground segment

F. Power Budget Calculation

At satellite communication system, to reach the best condition, we need to simulate in real time the quality of link from ground segment to space segment vice versa, before communication can take place. The result of quality link performance is very important to quality of service in satellite communication system. In order to simulate the quality of link performance, we need to calculate power budget from these parameters below:

1. Free Space Loss

Free space loss is a value describes how much signal power is loss due to propagation process from ground segment to space segment.

$$L_u = 10\log\left[\frac{4\times\pi\times F_u\times d_u}{c}\right]^2 \quad (10)$$

L_u = free space loss at uplink direction (dB)
F_u = uplink frequency (Hz)
d_u = slant range (km)
c = speed of light = $3\text{x}10^8$ m/s

2. Carrier Power (Rx Level)

This is the carrier power signal from ground segment antenna propagate to the space segment. This carrier power determines by EIRP from ground segment antenna, loss tracking antenna, free space loss and antenna gain at space segment side.

$$C_u = (EIRP_{SBT_x} - L) - L_u + G_{satR_x} \quad (11)$$

C_u = *carrier uplink receive by satellite antenna* (dB)
$EIRP_{SBTx}$ = EIRP ground segment Tx (dB)
L = *loss tracking* + *atmosphere attenuation* (1.2 – 1.5 dB)
L_u = *frees space loss uplink* (dB)
$GsatR_x$ = *gain* antenna satellite

3. *Gain to Noise Temperature* (G/T)

(G/T) is a parameter to compare the level of antenna gain at receiver end and noise temperature at receivers system.

$$\frac{G}{T}_{up} = G_{Rx} - L_R - L_{pol} - L_{FRx} - 10\log\left(T_{sys}\right) \quad (12)$$

$$T_{sys} = T_1 + T_2 \quad (13)$$

$$T_1 = T_A + \left(L_{FRx} - 1\right).T_F + T_R \quad (14)$$

$$T_2 = \frac{T_1}{L_{FRx}} = \frac{T_A}{L_{FRx}} + \left(1 - \frac{1}{L_{FRx}}\right).T_F + T_R \quad (15)$$

G/T = *gain to noise temperature* (dB/K)
G_{Rx} = *gain* antenna at receiver (dB)
L_R = *loss miss pointing* antenna
L_{pol} = *loss* polarization
T_A = temperature antenna satellite (K)
L_{FRx} = *loss feeder* (dB)
T_F = temperature feeder (K)
T_R = temperature at satellite equipment (K)

4. Carrier to Noise Ratio (C/N)

C/N is a parameter that compares carrier signal power to noise system.

$$\frac{C}{N} = EIRP - L + \frac{G}{T} - 10\log k - 10\log B - IB_o \quad (16)$$

L = *free space loss uplink* (dB)
G/T = *gain to Noise Temperature Ratio* at antenna satellite (dB)
k = *Boltzmann constant* = 1.3803×10^{-23} J/°K
B = *bandwidth frequency (*MHz)
IB_O = *back of input* = degradation of input value of signal power comparing to maximum input amplifier (dB)

5. *Carrier to Noise Ratio Total* (C/N_{total})

C/N total is a parameter who describes the quality of carrier power at the ends of system equipment to noise system.

$$\frac{C}{N}_{Total} = \left[\left(\frac{C}{N}_{Up}\right)^{-1} + \left(\frac{C}{N}_{Dn}\right)^{-1}\right]^{-1} \quad (17)$$

6. *Energy Per Bit to Noise Density Ratio (Eb/No)*

Eb/No (*Energy per Bit to Noise Density Ratio*) is a parameter who describes a comparison of energy per bit per square unit from output modulator to noise level of the system.

$$\frac{E_b}{N_o} = \frac{C}{N}_{Total} - 10 \log R \qquad (18)$$

R = speed of transmission (bps)

7. *Bit Error Ratio (BER)*

BER is a parameter to describe amount of information bit loss (error) compare to total information that has been transmitted.

4. Result and Analysis

Analysis and result in this research conduct by calculation of alteration pointing of ground segment antenna relatives to satellite. The value of alteration pointing of ground segment antenna is concluded with the value of receiving signal power at antenna. The calculation process is splitted into two parts, the calculation power budget from Telkom-1 satellite and the calculation power budget from Telkom-2 satellite. Those calculation based on assumption the location of transmitter is on 106^0 longitude and 6^0 latitude, that is at Telkom Master Control Station, Cibinong and receiver side position at 112^0 longitude and 7^0 latitude at Surabaya. Then, we can simulate calculation power budget by altering azimuth and elevation angle of each position ground segment receiver.

1. The result of power budget calculation by altering elevation angle at receiving antenna ground segment for Telkom-1 Satellite.

Power budget parameters result:

- BER = $1x10^{-9}$
- Eb/No = 4,9 dB
- C/N_{total} = 38,0133 dB
- C/N_{dn} = 58,598 dB
- C/N_{up} = 108,2116 dB
- G/T_{dn} = 30,7096 $dB/^0K$

Those result above is a fix position antenna pointing to Telkom-1 satellite, then comparing with the altering of any various elevation angle, it produce a variation number of C/N total, Eb/No and BER value showed by Table 1 below. When off axis value set on a furthermost value from its original, it is known that the variation value of C/N total could be from 23.63 dB to 19.38 dB rather than 38.01 dB as its original value.

Table 1. Result of Link Budget Calculation by Altering Elevation Angle at Receiving Antenna Ground Segment

Off-axis	Rx level	*Gain Rx*	G/T_{dn}	C/N_{dn}	C/N_{total}	Eb/No	BER
-0.8	-69.56942	27.50158	5.7098	33.5988	23.6383	-9.475	$<<1x10^{-1}$
-0.6	-60.56942	36.50158	14.7098	42.5988	30.5661	-2.5472	$<1x10^{-1}$
-0.4	-57.56942	39.50158	17.7098	45.5988	32.0805	-1.0328	$<1x10^{-1}$
-0.2	-49.56942	47.50158	27.7098	53.5988	35.8445	2.7312	$\pm1x10^{-2}$
0	-44.56942	52.50158	30.7098	58.5988	38.0133	4.9000	$1x10^{-9}$
0.2	-54.56942	40.50158	18.7098	46.5988	32.5772	-0.541	$<1x10^{-1}$
0.4	-69.56942	27.50158	5.7098	33.5988	25.6383	-7.475	$<1x10^{-1}$
0.6	-58.56942	38.50158	16.7098	44.5988	31.5823	-1.531	$<1x10^{-1}$
0.8	-79.56942	17.50158	-4.2902	23.5988	19.3738	-13.7395	$<<1x10^{-1}$

2. The result of power budget calculation by altering azimuth angle at receiving antenna ground segment for Telkom-1 Satellite.

By using the same power budget parameters above, it is shown in table below that the variation value of C/N total affected by alteration azimuth angle at receiving antenna is start from 23.84 to 35.84 dB from off axis positive value and start from 36.29 to 26.21 dB from off axis negative value rather than its original value at off axis 0, that's is 38.01 dB.

Table 2. Effect of Altering Azimuth Angle to Power Budget

Off-axis	Rx level	*Gain Rx*	G/T_{dn}	C/N_{dn}	C/N_{total}	Eb/No	BER
-0.8	-72,56942	24,50158	2,7098	30,5988	23,8537	-9,260	$<<1x10^{-1}$
-0.6	-55,56942	41,50158	19,7098	47,5988	33,0578	-0,0555	$<1x10^{-1}$
-0.4	-56,56942	40,50158	18,7098	46,5988	32,5723	-0,541	$<1x10^{-1}$
-0.2	-49,56942	47,50158	27,7098	53,5988	35,8445	2,7312	$\pm1x10^{-2}$
0	-44,56942	52,50158	30,7098	58,5988	38,0133	4,9	$1x10^{-9}$
0.2	-48,56942	48,50158	26,7098	54,5988	36,2890	3,1757	$\pm 1x10^{-5}$
0.4	-59,56942	37,50158	16,7098	44,5988	31,5823	-1,5310	$<1x10^{-1}$
0.6	-59,56942	37,50158	16,7098	44,5988	31,5823	-1,5310	$<1x10^{-1}$
0.8	-69,56942	27,50158	6,7098	34,5988	26,2165	-6,8968	$<<1x10^{-1}$

3. The result of power budget calculation by altering azimuth angle at receiving antenna ground segment for Telkom-2 Satellite.

Power budget parameters result:

BER = $1x10^{-9}$
Eb/No = 4.9 dB
C/N_{total} = 38.0133 dB
C/N_{dn} = 59.594 dB
C/N_{up} = 104.9718 dB
G/T_{dn} = 30.7098 $dB/^0K$
$EIRP_{Sb}$ = 119.6358 dB
P_{Tx} = 64.6124 dB

These result below show that when we altering and shift the position of azimuth angle antena at receiving ground segment, it would affect to Eb/No value and C/N total. The C/N total value is 38.01 dB when we set off axis 0 and swing to 25.45 dB (off axis value -0.8) and 19.92 dB (off axis value 0.8).

Table 3. Effect of Altering Azimuth Angle to Power Budget

Off-axis	Rx level	*Gain Rx*	G/T_{dn}	C/N_{dn}	C/N_{total}	Eb/No	BER
-0.8	-68,5755	27,5005	5,7086	33,5927	25,4487	-7,6646	$<<1x10^{-1}$
-0.6	-59,5755	36,5005	14,7086	43,5927	30,8014	-2,3119	$<<1x10^{-1}$
-0.4	-56,5755	39,5005	17,7086	46,5927	32,2696	-0,8437	$<1x10^{-1}$
-0.2	-48,5755	47,5005	25,7086	54,5927	35,9146	2,8013	$\pm 1x10^{-2}$
0	-43,5755	52,50158	30,7098	59,594	38,0133	4,9	$1x10^{-9}$
0.2	-53,5755	42,5005	20,7086	49,5927	34,0821	0,9688	$<1x10^{-1}$
0.4	-68,5755	27,5005	5,7086	33,5927	25,4487	-7,6646	$<<1x10^{-1}$
0.6	-57,5755	38,5005	16,7086	45,5927	31,7866	-1,3267	$<1x10^{-1}$
0.8	-78,5755	17,5005	-4,2913	24,5927	19,9247	-13,1886	$<<1x10^{-1}$

4. The result of power budget calculation by altering elevation angle at receiving antenna ground segment for Telkom-2 Satellite.

The same way to the 3rd process, power budget calculation has been done with those parameters above at the right position of antenna Rx station. By altering elevation angle with off axis value from -0.8 to 0.8 then produce C/N total from 24.28 dB to 38.01 dB at the right position. The biggest affect at this process is the value of BER decrease from 1 x 10^9 to << 1 x 10^{-1} which mean the amount of bit error increase rapidly.

Table 4. Effect of Altering Elevation Angle to Power Budget

Off-axis	Rx level	*Gain Rx*	G/T_{dn}	C/N_{dn}	C/N_{total}	Eb/No	BER
-0.8	-71,5755	24,5005	2,7086	31,5927	24,2841	-8,8292	$<<1x10^{-1}$
-0.6	-54,5755	41,5005	19,7086	48,5927	33,2164	0,1031	$<1x10^{-1}$
-0.4	-55,5755	40,5005	18,7086	47,5927	32,7461	-0,3672	$<1x10^{-1}$
-0.2	-48,5755	47,5005	25,7086	54,5927	35,9146	2,8013	$\pm 1x10^{-2}$
0	-43,5755	52,50158	30,7098	59,594	38,0133	4,9	$1x10^{-9}$
0.2	-47,5755	48,5005	26,7086	55,5927	36,3447	3,2313	$\pm 1x10^{-4}$
0.4	-58,5755	37,5005	15,7086	44,5927	31,2974	-1,8159	$<1x10^{-1}$
0.6	-58,5755	37,5005	15,7086	44,5927	31,2974	-1,8159	$<1x10^{-1}$
0.8	-68,5755	26,5005	5,7086	34,5927	26,0185	-7,0948	$<<1x10^{-1}$

5. Conclusion

1. In a general way, pointing error at satellite communication system would affect a great value of C/N total, Eb/No and BER.
2. The more value C/N change it would be affected the more changing value of EB/No and BER because of alteration azimuth and elevation angle of Rx ground segment antenna.
3. The alteration of elevation angle would affect more than the alteration of azimuth angle on BER value, producing much more error bit at transmission.
4. Biggest difference between C/N total when off axis 0 with alteration of azimuth and elevation angle done when pointing antenna directing to Telkom-1 satellite with 18.6395 dB differences point.

References

[1] Ahmed, Kamran. 2005. *Satellite Communication Lecture9.*

[2] *Anonymous*. *Vertex* RSI 11.1M *Cassegrain Antenna System*. 2009 http://www.satcomnow.com/satcom/images/full/VertexRS diakses pada tanggal 19 Mei 2009 pukul 14.28

[3] *CCIR*. 1988. *Satellite communications Handbook.* Geneva: *International Telecommunication Union.*

[4] Clarke, C Arthur, 1945. *Exta Terrestrial Relay*. Inggris: *Wireless World.*

[5] Dwiantoro, Tino. 2007. Modulasi, modul-4.

[6] Haykin, S. 2005. *Communication Link Budget Analysis*.

[7] Hranac, Ron. 2001. *QPSK and 16-QAM Digital Modulation*.

[8] Melinda, Rizal Munadi, dan M. Irhamsyah. 2004. *STUDI PERENCANAAN LINK BUDGET SISTIM KOMUNIKASI SATELIT PADA FREKUENSI KU-BAND DI NANGGROE ACEH DARUSSALAM.* Banda Aceh: Fak Teknik Universitas Syiah Kuala Banda Aceh

[9] Pamungkas, Wahyu. 2005. *Komunikasi Satelit*. Purwokerto: Diktat Kuliah Akademi Teknik Telkom Sandhy Putra Purwokerto.

[10] Pamungkas, Wahyu. 2005. *Modul Praktikum Komsat*. Purwokerto: Akademi Teknik Telkom Sandhy Putra Purwokerto.

[11] Simanjuntak, T.L.H. 2004. *Sistim Komunikasi Satelit*. Bandung: PT. Alumni.

COOPERATIVE SIGNAL DETECTION WITH DIFFERENT CHANNEL FADING

Nasrullah Armi[1], M Naufal M Saad[2], Muhammad Arshad[3]

[1,2,3]Universiti Teknologi PETRONAS
[1]armi@ppet.lipi.go.id, [2]naufal_saad@petronas.com.my, [3]muhammadarshad@comsats.edu.pk

Abstract

Cooperative sensing was proposed to combat noise fading and shadowing in cognitive radio networks. This technique will improve the performance of detection probability by sharing information of local detection among each cognitive radio user. This paper studies cooperative technique based on difference channel fading, which indicated by varying signal to noise ratio (SNR). Through computer simulation, we compare those results to non-cooperative scheme with signal to noise ratio and probability of detection as a metric. The results show that cooperative scheme based signal detection technique increases the performance of detection probability in compare with non-cooperative scheme. Furthermore, increasing number of cognitive radio users will improve detection probability in cognitive radio system.

Keywords: cognitive radio; cooperative spectrum sensing; probability of detection; probability of false alarm

1. Introduction

Today's wireless networks characterized by fixed spectrum assignment policy. As increasing demand for frequency spectrum and limited resource availability, FCC decided to make a paradigm shift by allowing more number of unlicensed users to transmit their signals in licensed bands to utilize the available spectrum efficiently.

Cognitive radio proposed as a means to overcome spectrum scarcity in wireless communication. A cognitive radio senses the spectrum environment over a wide range of frequency band and exploits this information to provide wireless links opportunistically that can meet the best demand of the user, but also of its radio environments.

The cognitive-radio devices have two important functionalities: spectrum sensing and adaptation. A secondary terminal first senses the spectrum environment in order to learn the frequency spectrum unoccupied by primary users. Once such a spectrum hole found, the secondary user adapts its transmission power, frequency band, modulation, etc., so that it minimizes the interference to the primary users. Even after starting the transmission, the secondary user should be able to detect or predict the appearance of a primary user so that it makes the spectrum available for the primary user.

Basically the primary user should not change their communication infrastructure due to these operations. Thus, these sensing (including the detection) and adaptation of the secondary users must be performed independently of the primary users. The goal of the spectrum sensing is to decide two hypotheses. It detects the presence or absence of the primary/licensed users and spectrum holes. Detection of the primary user is usually modeled as binary hypotheses as follows:

$$x(t) = \begin{cases} n(t) \\ h * s(t) + n(t) \end{cases} \quad (1)$$

where *x(t)* is the complex signal received by the cognitive radio, $s(t)$ is the transmitted signal of the primary user, $n(t)$ is the additive white Gaussian noise (AWGN) and h is the complex amplitude gain of the ideal channel.

Generally, spectrum sensing techniques can be classified into two main types, primary transmitter detection and interference temperature concept. It is characterized by two parameters: probability of detection, which is the probability of primary signal presence, and probability of false alarm is absence of primary signal probability. In addition, sensing techniques can be used both by cooperative and non-cooperative signal detection.

In non-cooperative signal detection, each cognitive radio individually decides whether the primary user is present or not and acts accordingly. These detectors generally use a matched filter detector, cyclostationary feature detector, and energy detector. Each of these detectors has its own pros and cons. While the energy based detector requires no prior knowledge of the signal and less complex than the other detectors. However, it has a limit on the required amount of signal SNR (called SNR wall) for the detector to work effectively. A matched filter and cyclostationary detector require prior knowledge of the signal that used to detect.

Meanwhile, the cooperative signal detection was proposed to overcome hidden node, noise fading and shadowing problem. Several nodes are collaborated in cognitive networks, then cognitive radio base station collects these local signal detection and conducts the final decision in accordance with the decision rules whether primary signal presence or absence in the cognitive radio networks. Most of the previous research focuses on comparison between non-cooperative and cooperative scheme with probability of detection and probability of false alarm as a metric. However, in this paper we present cooperative and non-cooperative comparison with signal to noise ratio (SNR) and detection probability to evaluate the system performance. It also presents the effects of increasing collaborated user number to detection probability by varying signal to noise ratio. The several of SNR values indicate different fading and shadowing which is experienced by CR users.

This paper is organized as follows. In section II cooperative spectrum sensing is briefly described. It introduces some cooperative sensing techniques that is commonly used in cognitive radio system. Section III, system model of non-cooperative and cooperative based signal detection is explained. Probability of signal detection for both of non-cooperative and cooperative cognitive users is formulated. Then, performance evaluation by presenting simulation results is discussed in section IV. Finally, we draw the conclusion in section V.

2. Cooperative Spectrum Sensing

One of the most critical issues of spectrum sensing is hidden terminal problem, which occurred when cognitive radio user is shadowed. In this case, they can not reliably sense the presence of primary user due to low SNR of received signal.

Cooperative signal detection was proposed to combat noise fading, shadowing, and hidden node problem in spectrum sensing algorithm. This technique has been widely used to detect the primary user with high speed and accuracy. In general, cooperative signal detection is performed as follows [1] [2] [3]:

- Each of cognitive radio users performs local signal detection independently and then makes a binary decision.
- All cognitive radio users forward their binary decision to a common receiver which is access point in cognitive radio base station in cellular network
- The common receiver fuses those binary decisions and makes a final decision to infer the absence or presence of primary user in the observed band.

In figure 1 shows the system model of cooperative signal detection where only one cognitive radio user could detect signal of primary user. The other cognitive radio users are not able to distinguish existence of the primary signal due to fading and shadowing effects. In this case, cooperative signal detection can improve the probability of signal detection. The cooperative signal detection among cognitive users is theoretically more accurate and convenient.

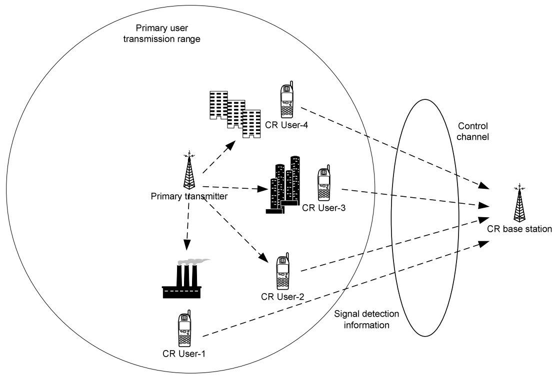


Figure 1. Cooperative Signal Detection

Some cooperative sensing techniques are introduced as follows [4]:

a) Centralized Sensing

In centralized sensing, a central unit collects sensing information from cognitive devices, identifies the available spectrum, and broadcasts this information to other cognitive radios or directly controls the cognitive radio traffic.

The hard sensing results are gathered at a central place which known as access point. The goal is to mitigate the fading effects of the channel and increase detection performance. Resulting detection and false alarm rates are given for the sensing algorithm. The sensing results are combined in a central node, termed as master node. Hard and soft information combining methods are investigated for reducing the probability of missed opportunity.

In the case of a large number of users, the bandwidth required for reporting becomes huge. Only the cognitive radios with reliable information are allowed to report their decisions to access point. Hence, some sensors are censored. Censoring can be implemented by simply using two threshold values instead of one [5]. Analytical performance of this method is studied for both perfect and imperfect reporting channels.

b) Distributed Sensing

In the case of distributed sensing, cognitive nodes share information among each other but they make their own decisions as to which part of the spectrum they can use. Distributed sensing is more advantageous than centralized sensing in the sense that there is no need for a backbone infrastructure and it has reduced cost.

An incremental gossiping approach termed as GUESS (gossiping updates for efficient spectrum sensing) is proposed in [6] for performing efficient coordination between cognitive radios in distributed collaborative sensing. The proposed algorithm is shown to have low-complexity with reduced protocol overhead. Incremental aggregation and randomized gossiping algorithms are also studied for efficient coordination within a cognitive radio network. A distributed collaboration algorithm is proposed in [6]. Collaboration is performed between two secondary users. The user closer to a primary transmitter, which has a better chance of detecting the primary user transmission, cooperates with far away users.

c) External Sensing

Another technique for obtaining spectrum information is external sensing. In external sensing, an external agent performs the sensing and broadcasts the channel occupancy information to cognitive radios. External sensing algorithms solve some problems associated with the internal sensing where sensing is performed by the cognitive transceivers internally. Internal sensing is termed as collocated sensing in [7]. The main advantages of external sensing are overcoming hidden primary user problem and the uncertainty due to shadowing and fading. Furthermore, as the cognitive radios do not spend time for sensing, spectrum efficiency is increased. The sensing network does not need to be mobile and not necessarily powered by batteries. Hence, the power consumption problem of internal sensing can be addressed as well.

3. System Model

3.1. Non-cooperative based signal detection

In this technique, each cognitive user decides the presence of primary signal. It has any kind of cooperation. Each cognitive user independently detects the channel, and if it is not utilized, they would vacate the channel without informing the other cognitive radio users.

Non-cooperative technique is fallible in signal detection. As seen in figure 2, cognitive users experience shadowing detect channel incorrectly

In signal detection, energy detector is the most common way of spectrum sensing due to its low computational and implementation complexities[8][9]. In addition, it is more generic in compare with the other techniques as receiver does not need any knowledge of primary user signal. The block diagram of signal detector is described in figure 3.

First, the input signal $y(t)$ is filtered with a Bandpass Filter (BPF) in order to limit the noise and to select the bandwidth of interest. The noise in the output of the filter has a band-limited, flat spectral density. Furthermore, as shown in the figure.3, there is energy detector consisting of a squaring device and a finite time integrator. The output signal V from the integrator is as follows:

$$V = \frac{1}{T}\int_{t-T}^{t} \left| y(r) \right|^2 dr \qquad (2)$$

Finally, this output signal V is compared to the threshold in order to decide whether a signal is present or not. The threshold is set according to statistical properties of the output V when only noise is present.

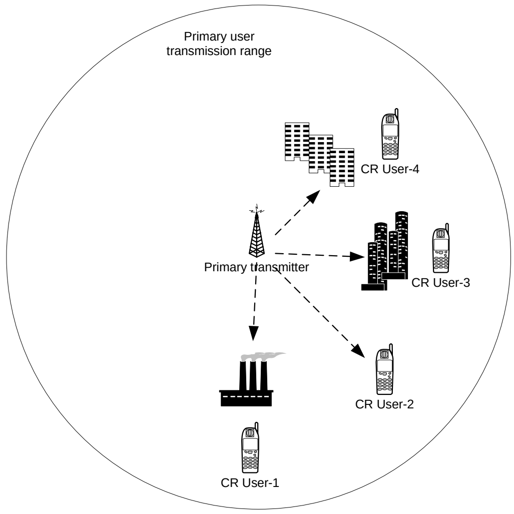


Figure 2. Non-Cooperative Signal Detection

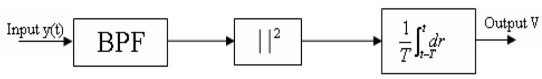


Figure 3. Energy Detection

The performance metric used for the simulations is receiver operating characteristics (ROC). It is completely specified by the values of probability of false alarm P_f and probability of detection P_d. In signal detection theory, ROC is used for measuring the performance as a trade off between selectivity and sensitivity. The probability of detection (or true positive) P_d is given as a function the probability of false alarm (or false positive) P_f.

Probability of false alarm can be computed using central chi-square (or gamma) PDF with N degrees of freedom.

$$P_f = P\{Y > \lambda | H_o\} = \frac{\Gamma\left(N/2, \dfrac{\lambda}{2\sigma^2}\right)}{\Gamma(N/2)} \qquad (3)$$

where Γ(.,.) and Γ( ) are the incomplete and complete gamma function respectively. In energy detector, this function is represented by the following operation [1]: 1) sampling the received signal and passing through an FFT device to obtain the signal spectrum, 2) the peak of the spectrum is then located and windowed, and finally, 3) the signal energy is then collected in the frequency domain and binary decision is created by comparing this energy to threshold value. Then, N is degrees of freedom, σ^2 is noise variance of communication, Y is a decision statistic, λ is the decision threshold, and H_0 stand for the hypothesis: no signal transmitted. On the other hand probability of detection can be computed using non-central chi-square PDF with N degrees of freedom.

$$P_d = P\{Y > \lambda | H_1\} = Q_{N/2}\left(\sqrt{\frac{S^2}{\sigma^2}}, \sqrt{\frac{\lambda}{\sigma^2}}\right) \quad (4)$$

where $Qm(.,.)$ is Marcum Q-function, and H_1 stand for the hypotheses: signal transmitted. The parameter $S^2 = \sum_{i=0}^{N/2} A_i$ is called the non-centrality parameter of the distribution and A is signal amplitude. These equations are valid for simply energy detector.

3.2. Cooperative based signal detection

In order to improve performance of signal detection, the cognitive radio base station collects the local signal detection information from each cognitive radio users and performs the final decision in accordance with decision rule.

There are 3 decisions rules in fusion for cooperative sensing. The cognitive radio base station receives the decision from n other users and decide H_1 if any of the total n individual decision is H_1. This fusion rule is known as OR rule. AND rule decide H_1 only if all decision is H_1 and MOST rule decides based on the most of decision results. However, this paper will only adopt OR rule for fusion decision in cognitive radio base station.

Signal detection and false alarm probability for cooperative sensing are formulated as follows:

$$C_d = 1 - \prod_1^n (1 - P_{d,k}) \quad (5)$$

$$C_{fa} = 1 - \prod_1^n (1 - P_{fa,k}) \quad (6)$$

where $P_{d,k}$ and $P_{fa,k}$ are detection probability and false alarm probability of the k-th cognitive user, respectively.

4. Result and Discussion

In this simulation, we adopt OR rule for cooperative based signal detection and assume number of collaborated users in the networks. Each CR user experiences different channel fading by varying SNR values. The threshold value is defined through simulation. The cognitive radio base station receives information from each user and makes the final decision whether the primary user is present or not.

Figure 4 describes the performance of signal detection probability for both of non-cooperative and cooperative users. It is shown that detection probability increases around 0.3 at SNR=0dB when cooperative based signal detection is employed in cognitive radio networks. The different values among the scheme become lower when SNR has higher values. As describes in the figure, when SNR value is around 10dB, detection probability becomes equal relatively for both of cooperative and non-cooperative users. It concludes that cooperative signal detection was proposed as solution for primary signal detection when SNR has lower values. In other word, cognitive radio user experiences fading and shadowing.

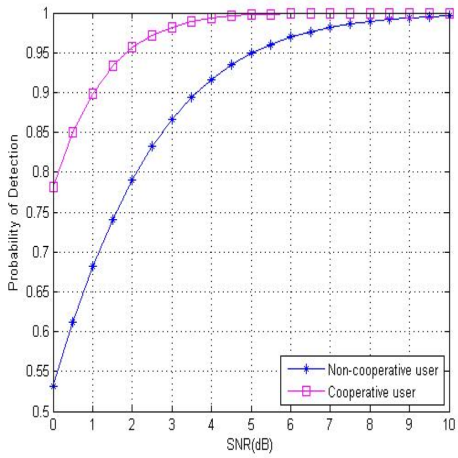


Figure 4. Probability of detection for non-cooperative and cooperative users

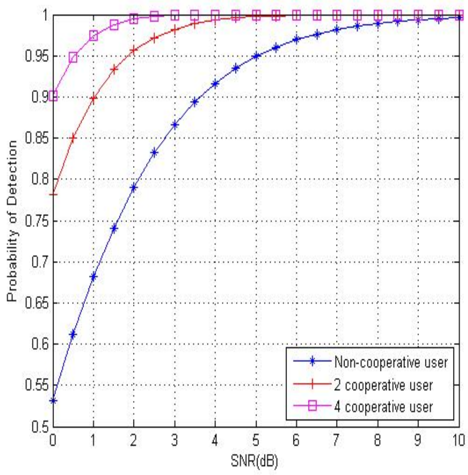


Figure 5. Probability of detection for non-cooperative, 2 and 4 collaborated users

The following simulation is investigating the effects of increasing collaborated user number in cognitive radio network. We assumed that 2 and 4 number of cognitive radio users is collaborated to detect primary user signal and send their local detection to fusion center. Then, fusion center decides whether primary signal is present or not.

Figure 5 shows that increasing number of collaborated users give better performance of detection probability. With the same SNR values, four collaborated user improves detection performance in compare with 2 users and single user (non-cooperative scheme). Furthermore, detection probability among them will be equal relatively when SNR greater than 10dB. This simulation concludes that increasing number of collaborated users in cognitive radio system can improve detection probability performance.

Furthermore, performance improvement of detection probability by increasing number of collaborated user shows in the following results. The simulation is performed by varying number of cooperative users. We use SNR values 0dB, 1dB, and 2dB, respectively, and the decision threshold (λ) = 12. The results confirm that increasing number of collaborated user can improve detection probability in cognitive radio networks. Hence, we conclude that primary user signal can be more accurately detected by increasing collaborated users in cognitive radio network in compare to non-cooperative scheme.

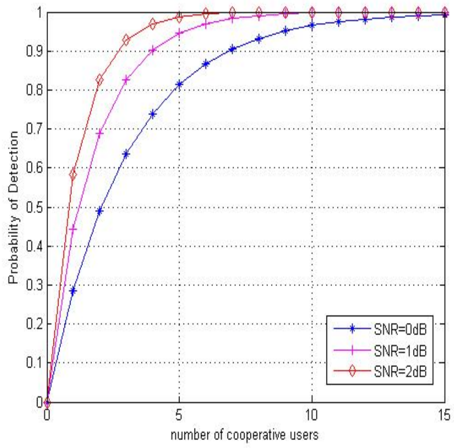


Figure 6. Number of cooperative user affects signal detection probability in cognitive radio networks (λ=12)

5. Conclusion

Cooperative based signal detection is one technique to improve detection probability of spectrum sensing in cognitive radio. This technique was proposed to combat noise fading and shadowing which probably occur in primary signal detection. Furthermore, the number of users which is collaborated in cognitive radio networks affects performance of signal detection. Increasing them can improve the probability of signal detection. Moreover, cooperative scheme could be effectively implemented in signal detection when primary signal experiences fading and shadowing.

References

[1] E. Hossain and V.K. Bhargava, "Cognitive wireless communication networks," New York, NY, Springer Science Business Media, LLC, 2007.

[2] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," IEEE Journal on Selected Areas in Communications, vol.23, no.2, February 2005.

[3] K.Hamdi and K.B. Letaief, "Cooperative communication for cognitive radio networks," in the 8th Annual Postgraduate Symposium, The Convergence of Telecommunications, Networking and Broadcasting, Liverpool John Moores University, June 2007.

[4] Tevfik Yücek and Hüseyin Arslan,"A survey of spectrum sensing algorithms for cognitive radio applications," IEEE Communication Surveys and Tutorials, Vol.11,No.1, First Quarter 2009

[5] Dong Chen, Jiandong Li, Jing Ma, "Cooperative Spectrum Sensing under Noise Uncertainty in Cognitive Radio" Proceeding of IEEE Conference, 2008

[6] N. Ahmed, D. Hadaller, and S. Keshav, "GUESS: gossiping updates for efficient spectrum sensing," in Proc. *International workshop on Decentralized resource sharing in mobile computing and networking*, Los Angeles, California, USA, 2006, pp. 12–17.

[7] Ganesan, G.; Li, Y, "Cooperative spectrum sensing in cognitive radio networks," in *Proc. IEEE Int. Symposium on New Frontiers in Dynamic Spectrum Access Networks*, Baltimore, Maryland, USA, Nov. 2005, pp. 137–143.

[8] D.Cabric, A.Tkachenko, and R.Brodersen, "Spectrum sensing measurements of pilot, energy, and collaborative detection," in Proc.IEEE Military Commun.Conf., Washington DC, USA, Oct,2006, pp. 1-7.

[9] D.Cabric, S.Mishra, and R.Brodersen, "Implementation issues in spectrum sensing for cognitive radios," in Proc. Asilomar Conf. in Signal, System, and Computers, vl.1, Pacific Grove, California, USA, Nov.2004, pp. 772-776.

DESIGN OF DISPERSION SHIFTED FIBER (DSF) TO INCREASE THE PERFORMANCE OF OPTICAL FIBER COMMUNICATION SYSTEM

Muhamad Tomi Haetami, Mamat Rokhmat[1], Akhmad Hambali[2]

[1]Engineering Physics Study Program, Institut Teknologi Telkom
[2] Telecommunication Engineering study Program, Institut Teknologi Telkom
[1]mrm@ittelkom.ac.id, [2]hbl@telkom.co.id

Abstract

Optical fiber communication system performance is limited by the attenuation and dispersion factor. Both factor must be minimized in order to get a good performance. The attenuation of fiber have a minimum value at 1,31 µm wavelength, while the dispersion of optical fiber have a minimum value at 1,55 µm wavelength. There are two option that must be chosen. If we choose one of the wave length then the performance of optical fiber is not optimum. One way to decrease the dispersion at optical fiber is using *Dispersion Shifted Fiber (DSF)*. DSF is optical fiber that have minimum dispersion at 1,55 µm wavelength. The wavelength that have a minimum value of dispersion is shifted by DSF. So we have optical fiber with minimum dispersion and attenuation. In this paper, DSF is made by constructing three layer of cladding in single mode fiber, and the design is done by manipulating the refracted index profile of core and each layer. In the result of the design, we have DSF with the value of dispersion is -0,05819 ps/nm.km in 1,55 µm wavelenght. In this paper we also investigate the dispersion characteristic, cut off wavelenght and field distribution.

Keywords: Optical communication system, Dispersion, DSF, Single mode, Tripple Cladding

1. Introduction

Dispersion and attenuation are the factor that can be limited optical fiber communication system. There are so many way to solve the broadening of the pulse caused by dispersion in optical fiber. One of them is using modified optical fiber or it is called *Dispersion-Altered Fibers*, like *Dispersion-Flattened Fibers* (DFF), *Dispersion-Shifted Fibers* (DSF), and *Dispersion-Compensating Fibers* (DCF) [1]. By using *Dispersion Shifted Fiber (DSF)*, wavelenght that have minimum dispersion is shifted to 1,55 µm wavelength, so we can get fiber with minimum value of dispersion and attenuation [2].

Dispersion-shifted fibers is optical fiber that have a minimum value of dispersion in wavelenght of minimum attenuation, where in SiO_2 – GeO_2 fiber is placed in 1,55 µm wavelength [3]. DSF have a minimum dispersion in very thin range of wavelenght so DSF is verry compatible for single channel transmission. The evolution of the pulse in optical fiber can be described by *Nonlinier Schrödinger Equations* (NLS) [5]:

$$i\frac{\partial A}{\partial Z}+\frac{i}{2}\alpha A-\frac{1}{2}\beta_2\frac{\partial^2 A}{\partial T^2}+\gamma|A|^2A=0 \quad (1)$$

This equations is depending on loss coefficient α, dispersion β_2, dan nonlinear parameter γ.

The aim of this research are trying to find and design the DSF that have dispersion less closed to zero in the wavelenght about 1,55 µm, and to investigate the fiber parameter of the design result.

The first step in this research is trying to find a model or characteristic equation that can be describe the evolution of the pulse. The second step is making a computer program to find the numerical result from the characteristic equations.

2. Theorical Background

Multiple cladding optical fiber is made to result optical fiber with optimum dispersion slope. In this paper, the general analyze of multiple cladding optical fiber is focused in four layer cylindrical (one core and triple cladding) of dielectric structure. Special equation is made in general for some probability of triple cladding optical fiber. r_i is radius of each layers and n_i is refracted index of each layers. If i = 1, we have the core of the fiber, and i = 2,3, and 4 are *layer* cladding. The most outer *layer* (i = 4) is assumed have infinity radius. The wave in optical fiber must obey the equation

$$\nabla_t^2\psi+(k_0^2n^2-\beta^2)\psi=0 \quad (2)$$

The method to find the solutions of above equation is using cylindrical and separation variable analyze, until we have a characteristic equation [5]. Characteristic equation is a function of optical fiber parameters include radius of each layers and refraction index, azimuth constant ν, wavelenght λ, and propagation constanti β. The form of characteristic equation is

$$f(r_i,n_i;i=1,...,4,\nu,\lambda,\beta)=0\,. \quad (3)$$

If radius of each layers, refractive index each layers, and ν are given then this equation only depends on β dan λ variable. This equation can be solved numerically by using *Bisection Methods*. If we give the *range* of λ variable then the value of β can be get. This condition represents some modes of waveguide. This mode can be represented with $LP_{\nu m}$, where *LP* is Linear Polarization. And then *m* ≥ 1 is the order of mode that the value of field is maximum/minimum in the radial direction. In the other side, integer *ν*≥0 has a meaning that the value of field is maximum/minimum in the azimuth direction.

Table 1The composition and refracted index of fiber materials

Material	Composition	Refracted index (n)
M1	SiO_2 (pure silica)	1,44402
M2	13,5 m/o GeO_2 + 86,5 m/o SiO_2	1,46598
M3	7,0 m/o GeO_2 + 93,0 m/o SiO_2	1,4554
M4	4,1 m/o GeO_2 ; 95,9 m/o SiO_2	1,45031
M5	9,1 m/o GeO_2 ; 83,2 m/o B_2O_3 ; 7,7 m/o SiO_2	1,45505
M6	4,03 m/o GeO_2 ; 9,7 m/o B_2O_3 ; 86,27 m/o SiO_2	1,44767
M7	0,1 m/o GeO_2 ; 5,4 m/o B_2O_3 ; 94,5 m/o SiO_2	1,44454
M8	13,5 m/o B_2O_3 ; 86,5 m/o SiO_2	1,44217
M9	13,5 m/o B_2O_3 ; 86,5 m/o SiO_2 (*Chilled*)	1,44174
M10	3,1 m/o GeO_2 ; 96,9 m/o SiO_2	1,4487
M11	3,5 m/o GeO_2 ; 96,5 m/o SiO_2	1,44951
M12	5,8 m/o GeO_2 ; 94,2 m/o SiO_2	1,4529
M13	7,9 m/o GeO_2 ; 92,1 m/o SiO_2	1,45624
M14	3,0 m/o B_2O_3 ; 97,0 m/o SiO_2	1,44225
M15	3,5 m/o B_2O_3 ; 96,5 m/o SiO_2	1,44165
M16	3,3 m/o GeO_2 ; 9,2 m/o B_2O_3 ; 87,5 m/o SiO_2	1,4435
M17	2,2 m/o GeO_2 ; 3,3 m/o B_2O_3 ; 94,5 m/o SiO_2	1,44621
M18	*Quenched* SiO_2	1,44439
M19	13,5 m/o GeO_2 ; 86,5 m/o SiO_2	1,46552
M20	9,1 m/o P_2O_5 ; 90,9 m/o SiO_2	1,45894
M21	13,3 m/o B_2O_3 ; 86,7 m/o SiO_2	1,43855
M22	1,0 F ; 99,0 m/o SiO_2	1,43942
M23	16,9 m/o Na_2O ; 32,5 m/o B_2O_3 ; 50,6 m/o SiO_2	1,50771

Equation 3 is used to design and to analyze *dispersion-shifted fiber* and optimization of *dispersion-shifted fiber* by using numerical methods. The transmission properties that will be evaluated are dispersion characterization *D*, propagation constant, cut off wavelenght, and field distribution at λ = 1,55 μm, and the propagation of single pulse in the result of DSF design. Next, a computer program is made from the characteristic equation to find the numerical solution [6]. The input of this program are material composition, radius, wavelength, and number of mode. In this program, propagation constant is calculated as a function of wavelength by using root searching methods. In this design we use materials based on silica like in table 1. The refracted index of the materials can be found by substituting Sellmeier coefficient for each materials in λ = 1,55μm.

Table 2. Parameters and the dispersion of design results

Fbr	n_1	n_2	n_3	n_4	r_1	r_2	r_3	D
A	M13	M1	M6	M4	3,2	3,8	4,3	-0,058
B	M13	M1	M6	M4	3,3	3,8	4,3	1,198
C	M13	M1	M6	M4	3,4	3,8	4,3	2,45
D	M13	M1	M6	M4	3,2	3,9	4,3	-0,517
E	M13	M1	M6	M4	3,2	4	4,3	-0,97
F	M13	M1	M6	M4	3,2	3,8	4,4	-0,778
G	M13	M1	M6	M4	3,2	3,8	4,5	-1,483
H	M3	M1	M6	M4	3,2	3,8	4,3	1,117
I	M5	M1	M6	M4	3,2	3,8	4,3	1,611
J	M13	M16	M6	M4	3,2	3,8	4,3	-0,427
K	M13	M14	M6	M4	3,2	3,8	4,3	-1,297
L	M13	M1	M17	M4	3,2	3,8	4,3	-0,995
m	M13	M1	M7	M4	3,2	3,8	4,3	-1,959
N	M13	M1	M6	M11	3,2	3,8	4,3	-0,058
O	M13	M1	M6	M10	3,2	3,8	4,3	-0,058

A normalization of propagation constant is represented by

$$b = \frac{\bar{\beta}^2 - n_4^2}{n_{\max}^2 - n_4^2} \quad (4)$$

where n_{max} is the highest value of refracted index in the fiber that will be designed. A normalization of propagation constant must be in the range $n_4 < \bar{\beta} < n_{max}$, so *b* is always having a value in range of 0 and 1. In b expression the characteristic equation can be rewritten by

$$f(\lambda, b) = f_1(\lambda, b) - f_2(\lambda, b) = 0 \quad (5)$$

Equation 5 is solved in order to get *b* by using bisection Methods [7]. This methods have an error $\varepsilon = 10^{22}$. In this research, we use 1μm - 1,7μm wavelength, because that range of wavelength have a minimum value of attenuation for fiber based on silica.

3. Design and Results

In this paper the design of DSF is using fiber with one core and triple *cladding,* where the refracted index profile is $n_2 < n_3 < n_4 < n_1$. The design is done by manipulating on the radius of core and radius of each cladding. Then we try to find a fiber that have a minimum dispersion at λ = 1,55µm. The dispersion is second derivative of propagation constant over wavelenght. Table 2 are some example of the design results.

3.1 Effect of radius fiber to Dispersion

Figure 1-3 show the effect of radius of core , *cladding* 1, and *cladding* 2 over the value of total dispersion. The facts are:

(i) If the radius of the core is increasing, then the dispersion is increasing too. It can be seen at Figure 1.

(ii) If the radius of the cladding 1 is increasing, then the dispersion is decreasing. It can be seen at Figure 2.

(iii) If the radius of the cladding 1 is increasing, then the dispersion is decreasing. It can be seen at Figure 3.

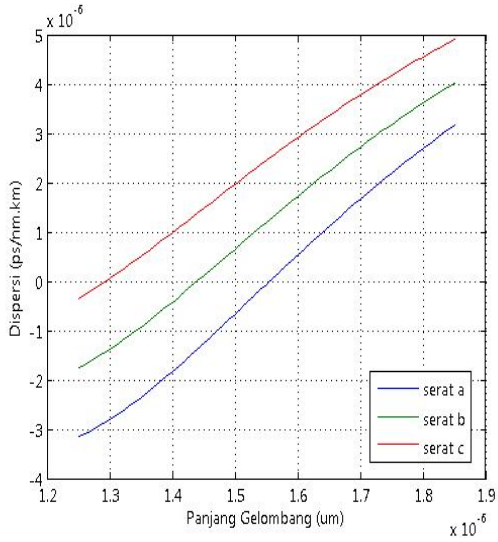


Figure1. The effect of core radius

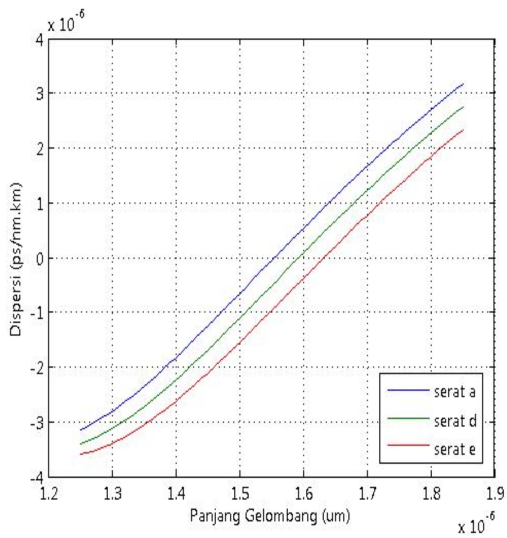


Figure 2. The effect of *cladding* 1 radius

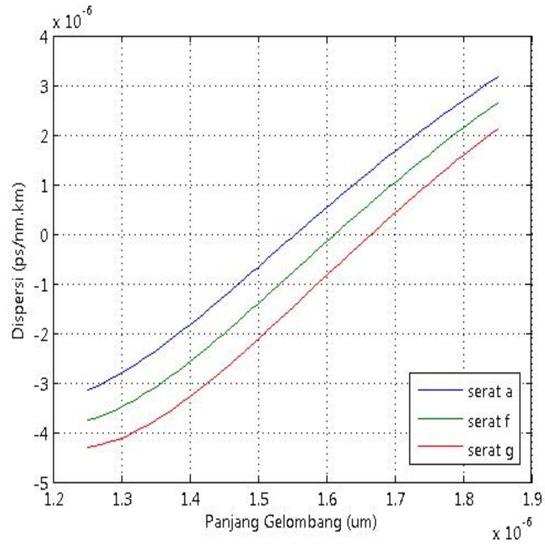


Figure 3. The effect of *cladding* 2 radius

3.2 Effect of refracted index to Dispersion

Figure 4-7 show the effect of refracted index of core , *cladding* 1, and *cladding* 2 over the value of total dispersion. The facts are:

(i) If the refracted index of fiber core is increasing then the total dispersion is decreasing. It can be seen at Figure 4.

(ii) If the refracted index of *cladding* 1 is increasing then the total dispersion is increasing. It can be seen at Figure 5.

(iii) If the refracted index of *cladding* 2 is increasing then the total dispersion is increasing. It can be seen at Figure 6.

(iv) The change of refracted index in *cladding* 3 does not impact to the value of total dispersion. It can be seen at Figure 7.

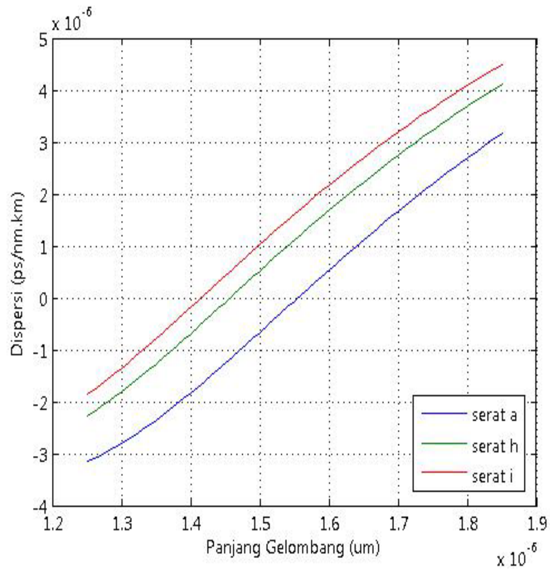


Figure 4. Effect of refracted index of the fiber core to total dispersion

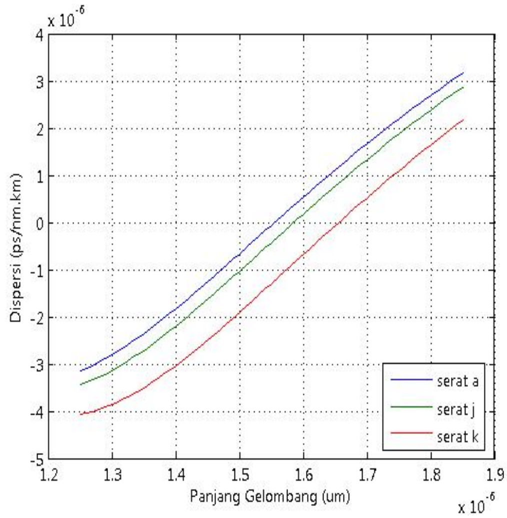


Figure 5. Effect of refracted index of the cladding 1 to total dispersion

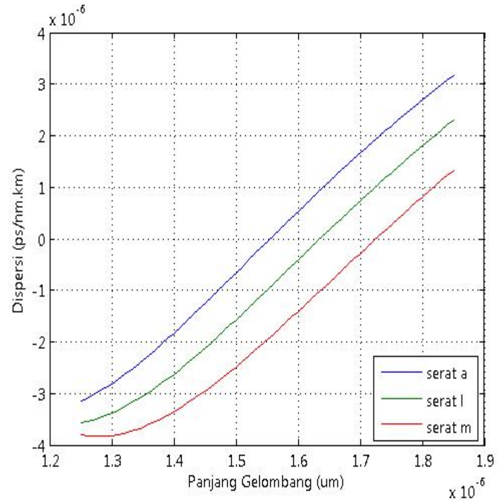


Figure 6. Effect of refracted index of the cladding 2 to total dispersion

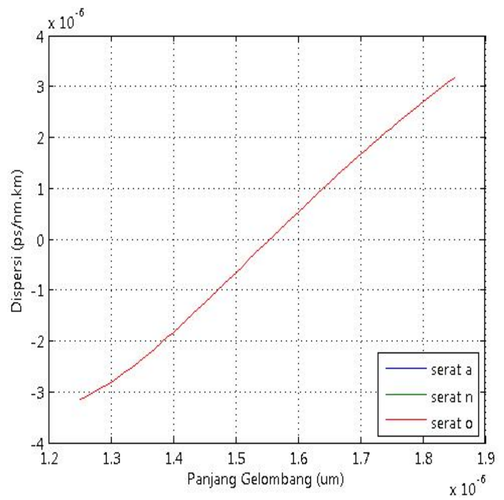


Figure 7. Effect of refracted index of the cladding 3 to total dispersion

From above figure, it is easy to see that the change of refracted index of the fiber can impact to the total dispersion.

4. Fiber parameter Analysis of the design result

In this paper, fiber parameter Analysis of the design result are dispersion characteristic, normalized of propagation constant b, and distribution of radial field at λ=1,55 µm.

4.1 Dispersion Characteristic

The choice of material and geometry of radius at each fiber layers can influence the design result. By doing manipulation at the value of core radius (r_1),cladding 1 radius (r_2), *cladding* 2 radius (r_3), refracted index of the *core* (n_1), refracted index of *cladding* 1 (n_2), refracted index of *cladding* 2 (n_3), and refracted index of *cladding* 3 (n_4), we can find that the "a" fiber is closed to the fiber that become to the aim of this research.. "a" have a dispersion equal to -0,05819 ps/nm.km, and the slope of *D* equal to 0,012 ps/nm^2.km at λ = 1,55µm. Dispersion characteristic over wavelenght for LP_{01} mode of "a" fiber is shown at Figure 8.

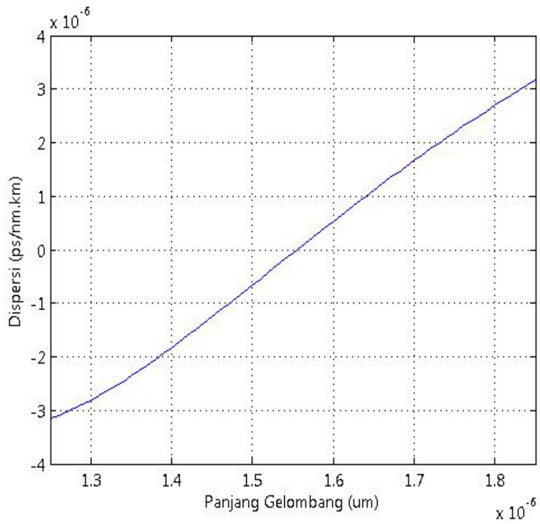


Figure 8. Graph of Dispersion over wavelenght for LP_{01} mode of "a" fiber

4.2 Fiber Optimalization

Dispersion-shifted fiber ("a" fiber) have a value of dispersion (i.e. second order) closed to zero, so this dispersion does not influence the propagation of the pulse. However, probably there is the third order of dispersion that can influence the propagation of the pulse. We can take the series Taylor of $\beta(\lambda)$ about $\lambda=\lambda_0$,

$$\beta(\lambda) = \beta(\lambda_0) + (\lambda - \lambda_0)\left.\frac{d\beta}{d\lambda}\right|_{\lambda=\lambda_0} + \frac{(\lambda-\lambda_0)^2}{2!}\left.\frac{d^2\beta}{d\lambda^2}\right|_{\lambda=\lambda_0} + \frac{(\lambda-\lambda_0)^3}{3!}\left.\frac{d^3\beta}{d\lambda^3}\right|_{\lambda=\lambda_0} + \ldots$$

where λ_0 = 1,55µm. $d\beta/d\lambda$ represents *delay* group, and $d^2\beta/d\lambda^2$ second order of dispersion. In general, third order of dispersion is neglected, but if $d^2\beta/d\lambda^2$

is zero in operational wavelenght then third order of dispersion must be used in the pulse propagation. In conventional *dispersion-shifted fiber*, $d^2\beta/d\lambda^2$ is disappear at λ = 1,55µm. However, in *dispersion-shifted fiber optimalization*, second and third orde of dispersion must be disappear at λ_0=1,55µm. We have "p" fiber as the fiber optimalization. The dispersion parameters of "p" fiber is shown in Table 3.

Table 3. Dispersion Parameters and Fiber Optimalization

Fiber	n_1	n_2	n_3	n_4	r_1	r_2	r_3	D
P	M20	M8	M18	M4	2,93	3,5	4,5	-0,034

At λ_0=1,55µm, second orde of dispersion and third order of dispersion (the slope of *D*) closed to zero. Total dispersion less than 1 ps/nm.km at the range of wavelength 1,41µm < λ < 1,69µm. Dispersion Characteristic for mode LP_{01} mode of "p" fiber is shown in Figure 9.

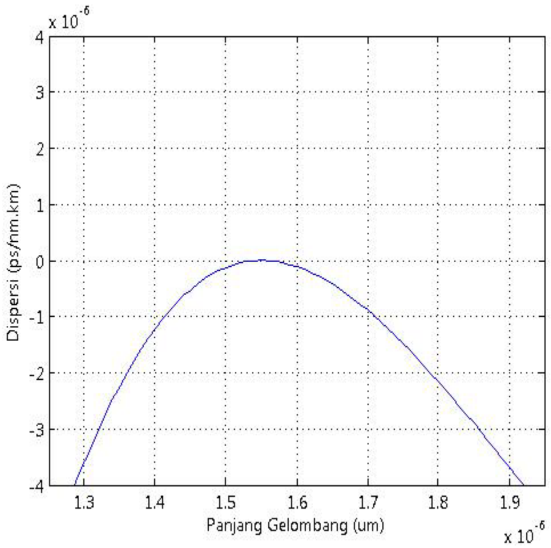


Figure 9. Graph of Dispersion over wavelenght for LP_{01} mode of "p" fiber

4.3 Normalized propagation constant

From above experiment, we found that "a" fiber is closed to the aim of this research. Figure 10 describes normalized constant of propagation (b) as a function of wavelenght for two lowest orde of mode LP_{01} and LP_{11} for "a" fiber. From this figure we see that *cutoff* for LP_{01} mode occurs when λ > 2µm, and *cutoff* for LP_{11} mode occurs when λ =0,96µm. The other mode have s *cutoff* frequency less than LP_{11} mode. "a" fiber have a single mode at range of wavelenght about 1,0µm < λ < 2,0µm, it can be seen in figure 11. However, for fiber optimalization of LP_{11} mode have a cutoff at λ = 0,87µm, and LP_{01} mode have a cutoff at 1µm < λ < 2µm.

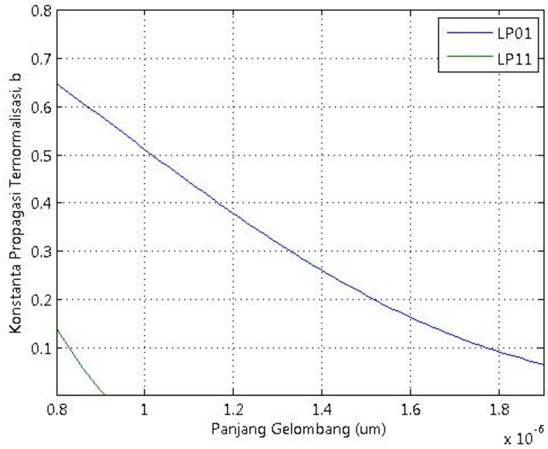


Figure 10.Graph of Normalized Propagation Constant for LP_{01} Mode and LP_{11} Mode of "a" Fiber

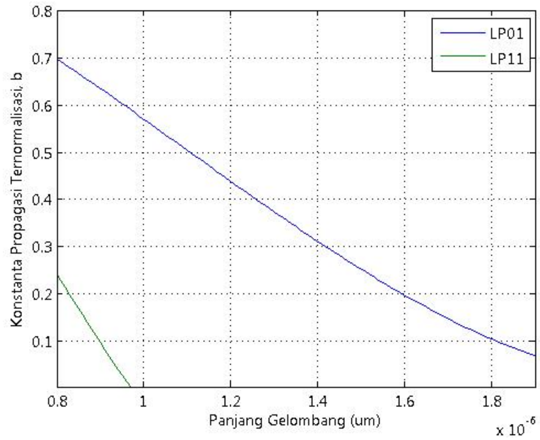


Figure 11. Graph of Normalized Propagation Constant for LP_{01} Mode and LP_{11} Mode of "p" Fiber

4.4 Field Distribution

Field distribution is evaluated in a fix of wavelenght and mode. The result of maximum field is normalized and plotted as a function of radius of fiber. The result of maximum field at λ = 1,55 µm and LP_{01} mode for "a" fiber "p" fiber can be seen at figure 12 and figure 13.

The field in fiber is maximum in the layer that have highest refracted index layer. In "a" fiber and "p" fiber, the core have highest refracted index so in this layer, the field is maximum. Information of normalized field is needed for finding the thick of *cladding* 3 (r_4) radius. This *Cladding* is assumed have an infinity radius in a model that we use to analyze *triple-cladding fiber*. However, in real manufacture, *cladding* 3 must have finite radius. The way to find this radius is the field distribution in cladding 3 must below 10^{-4}% from maximum field that occurs in the core. According to this criteria, we found that the cladding 3 radius for the "a" fiber is 34,09 µm, and for the "p" fiber is 24,62 µm.

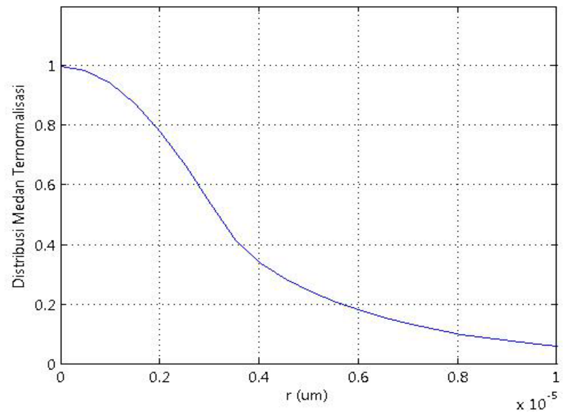


Figure 12. Graph of field distribution for LP_{01} mode of "a" fiber at λ = 1,55µm

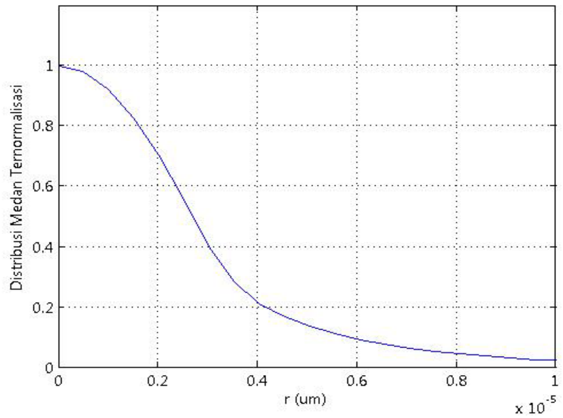


Figure 13. Graph of field distribution for LP_{01} mode of "p" fiber at λ = 1,55µm

4.5 Propagation of the pulse

Dispersion effect influences the propagation of the pulse. In "a" fiber, the dispersion is -0,05819 ps/nm.km at 1,55 µm wavelenght. The propagation of the pulse in "a" fiber is shown in Figure 14. If the widht of the pulse is given equal to 1.10^{-11} s then in 5 km distance, the width of the pulse still. This situation is the same, if we apply this pulse to the "p" fiber (can be seen in Figure 15).

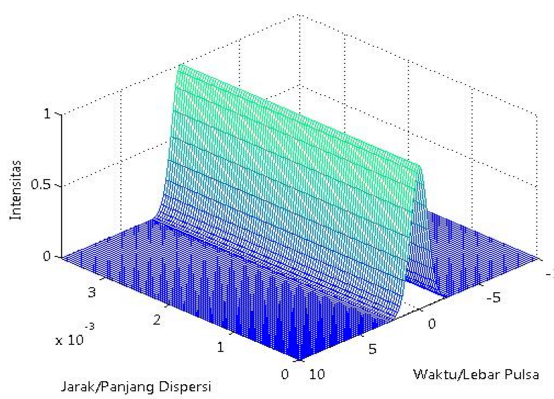


Figure14. Evolution of pulse propagation in "a" fiber at λ = 1,55 µm

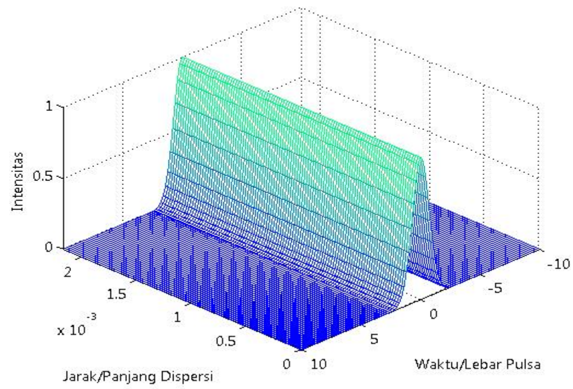


Figure15. Evolution of pulse propagation in "p" fiber at λ = 1,55 µm

5. Conclusion

According to this research, we have conclusion that :

a. Geometrical fiber (radius and refracted index of each layers) can influence the total dispersion.
b. The design result of Dispersion Shifted Fiber (DSF) Triple-Clad have the core radius r_1=3,2µm, cladding 1 radius r_2=3,8µm, cladding 2 radius r_3=4,3µm, material of the core is n_1=M13, cladding 1 is n_2=M1, cladding 2 is n_3=M6, and cladding 3 is n_4=M4, with the total dispersion is -0,05819 ps/nm.km and the slope of the dispersion is 0,012 ps/nm^2.km at 1,55 µm wavelenght.
c. The design result of Optimalized Dispersion Shifted Fiber (DSF) Triple-Clad have the core radius r_1=2,93µm, r_2=3,5µm, r_3=4,2µm, and material of the core is n_1=M20, cladding 1 is n_2=M18, cladding 2 is n_3=M8, and cladding 2 is n_4=M4 with total dispersion less than 1 ps/nm.km at 1,41µm < λ < 1,69µm.
d. Single mode characteristic of optimalized DSF occurs at 1,0µm < λ < 2,0µm.
e. Normalized of filed distribution is maximum at the core of the fiber.

References

[1] Barake, Taha M. 1997. *A Generalized Analysis of Multiple-Clad Optical Fibers with Arbitrary Step-Index Profiles and Applications*: Tesis. Blacksburg: Virginia Polytechnic Institute and State University.

[2] Hattori, Harroldo T. 1998. *Low Nonlinearity Optical Fibers for Broadband and Long-Distance Communications*: Tesis. Blacksburg: Virginia Polytechnic Institute and State University.

[3] Keiser, Gerd. 2000. *Optical Fiber Communications 3rd Edition*. Singapura: McGraw-Hill.

[4] Agrawal, P.Govind. 1989. *Nonlinear Fiber Optics*. New York : Academic Press.

[5] Iskander. 992. *Electromagnetic FieldsWaves*.New Jersey: Prentice Hall.

[6] Kosasih, P. Buyung, Ph.D. 2006. *komputasi Numerik Teori dan Aplikasi*. Yogyakarta: Penerbit ANDI

[7] Hanselman, Duane. Littlefield, Bruce. 2000. *MATLAB Bahasa Komputasi Teknis*. Yogyakarta: Penerbit ANDI.

DESIGN AND REALIZATION PULSE GENERATOR FOR GROUND PENETRATING RADAR APPLICATION USING TIMING SWITCHED METHOD

Frangky Ariadi Simanjuntak[1], Heroe Wijanto[2], Iswahyudi Hidayat[3]

[1,2,3]Fakultas Teknik Elektro dan Komunikasi, Institut Teknologi Telkom

Abstract

Precocity of electromagnetic technology in several recent years thinks out a new method at exploration of geophysics, that is Ground Penetrating Radar (GPR). Ground Penetrating Radar (GPR) is a method of geophysics by using electromagnetic techniques that are designed to detect the buried object in subsurface or underground and evaluate the depth of the objects. GPR can also be used to know the conditions and characteristics of subsurface without drilling or excavation. Pulse generator is an important component to produce the appropriate pulse for GPR. GPR system consists of transmitter and receiver. Transmitter is antenna which is connected to pulse generator. Receiver is antenna which connected to LNA and ADC, and then connected to processing unit and display. Signal of GPR is generated with synchronization of sinusoidal generator's output by square pulse generator's output using timing switched method. Pulse postprocessing process is using band pass filter and amplifier. Output pulse generator produces pulse monocycle at frequency (200 ± 15) MHz, pulse width (5 ± 1)ns, PRI 67,2 ns, and PRF 14,88 MHz. It produces a stable enough pulse monocycle although there are some ripples that is caused square pulse generator is not yet stable, distortion is caused inter connection block, and using single layer PCB. Nevertheless, pulse generator can use in GPR because it has waveform appropriate for GPR specification.

1. Introduction

Ground Penetrating Radar is the new method on geophysics of exploration. GPR is intensive method that used in ground to mapping structure of underground, on archeology, pipe, wire, mineral underground. Detection process in GPR is using electromagnetic wave. Position and shape of the object was known by reflected signal that transmitted by the object. Impulse radar that used to detect the object is produce by pulse generator where pulse characteristic of the using signal is depending on maximum dept of the object. Pulse generator is a transmitter on GPR system that supplied by high voltage supply.

2. Theorem

2.1. Ground Penetrating Radar

In principles, GPR system is depending on reflection signal. GPR systems have transmitter, pulse generator and receiver.

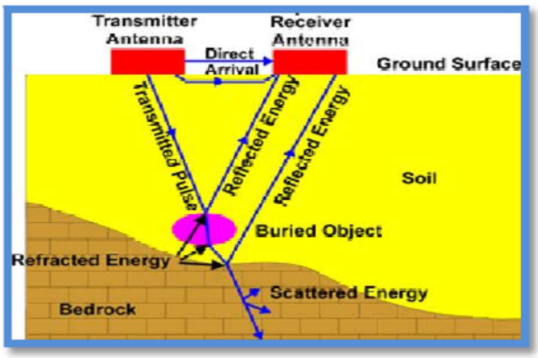


Figure 1. Illustration of GPR System

In GPR system, the Antenna is connecting to pulse generator that operate by timing, switched. In the receiver, antenna is connecting to LNA and ADC (analog to Digital converter) and than connect to data processing than the output of data processing will display.

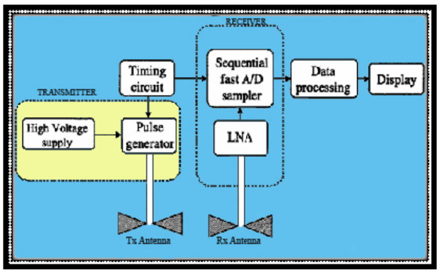


Figure 2. Diagram Block of GPR

2.2. Pulse Generator

Pulse generator is a initial tools in GPR, that have a function as generate impulse signal to radiate to underground. Impulse signal have monocycle with pulse width in nanosecond. Pulse generator have five block, sinusoidal generator, pulse generator, mixer, Band Pass Filter and Amplifier. Every block will be integrated and will do matching impedance.

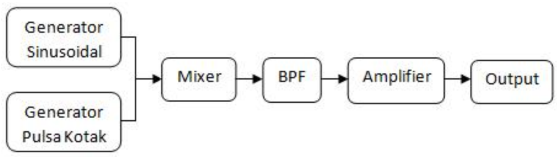


Figure 3. Diagram Block of Pulse Generator

2.3. Timing Switched Method

Timing Switched Method is a method to produce impulse signal in Ground Penetrating Radar that have monocycle waveform. Principe of this method is there is source that as a carrier and another source as timing switched. The two sources will be mix and produce a signal with delay because of timing control by timing switched. With another way, timing switched have a function as on/off second source.

3. Design and Realization Pulse Generator

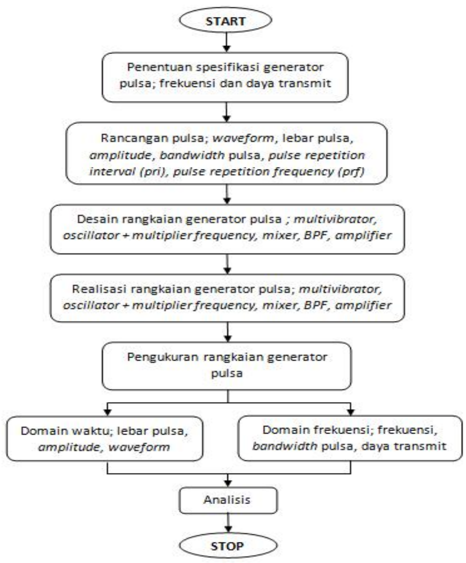


Figure 4. Flowchart Design and Realization of Pulse Generator

3.1. Design of Monocycle Pulse

GPR will realize at frequency (fc) 200 MHz:

- Pulse width : $t_C = \frac{1}{f_C} = \frac{1}{200.10^6} = 5ns$
- Depend on UWB specification: $B_f = \frac{B}{f_C} = \frac{B}{200.10^6} > 0.2$ and $B > 0.2x200.10^6 > 40MHz$.
  using bandwidth 50 MHz.
- f_L = 175 MHz and f_U = 225 MHz, f_C = 200 MHz
- Fractional *bandwidth* = $\frac{B}{f_C} = \frac{50MHz}{200MHz} = 0.25 > 0.2$
- Expectation of output power is 20mW or equal to 13dBm.
- *Peak pulse voltage* 1 V with current 260 mA produce *peak pulse power* = 260 mW.
- From[1]:

$$mean\_pulse\_power = \frac{t_c}{T} = \frac{5.10^{-9}}{T} \times 260mW = 20mW.$$

space between pulse $pri = T = 65ns.$

- *Pulse repetition frequency* $= \frac{1}{65\times10^{-9}} = 15.38MHz.$
- *Time Resolution two way* $(\Delta t) = \frac{1}{B} = \frac{1}{50.10^6} = 20ns.$

3.2. Specification of Pulse Generator

Table 1. Specification of Pulse Generator

Input and Output Impedance	50 Ohm
Type of HF connector	BNC sockets
Waveform	Monocycle
Center Frequency	200MHz
Bandwidth	>40MHz
Voltages	± 1Volt
PRF	15.38MHz
PRI	65ns
Mean Pulse Power	13dBm
Power Spply	12 Volt

3.3. Design and Realization of Sinusoidal Generator

Implementing signal generator in this research is using crystal oscillator to generate sinusoidal signal.

Crystal oscillator is expected to be work at 200MHz, but because of there is only crystal oscillator 59,475 MHz in the market, in this research we design multiplier frequency using IC MC1496. To generate frequency 200 MHz, we use two doubler frequency to produce 197.9 MHz (closed to 200 MHz). In this research we use DC power supply 12 Volt. Soft ware that use in this research is Altium DXP.

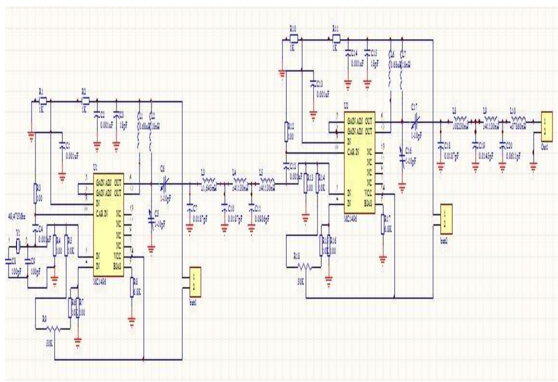


Figure 5. Multiplier Frequency Circuit

3.4. Design and Realization of Square Pulse Generator

Multivibrator is used to generate square pulse. IC 555 is used to design multivibrator with width pulse 5ns and period 65ns.

For multivibrator circuit we use:

C1 = 1nF and *T = T1 +T2 = 65ns* and we get:

*T1 =0.7*R1*C1* (1)

T1 = 5 ns

*R1=T1/(0.7*C1)* (2)

R1=7.14 Ohm

*T2=0.7*R1*C1* (3)

T2=60 ns

*R2= T2/(0.7*C1)* (4)

R2= 85.71 Ohm

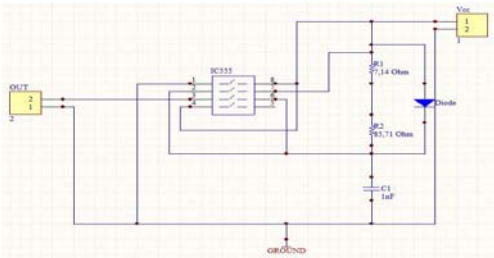


Figure 6. Multivibrator Circuit

3.5. Design and Realization of Mixer

IC MC1496 is used to design mixer circuit. This IC has a function as balanced modulator.

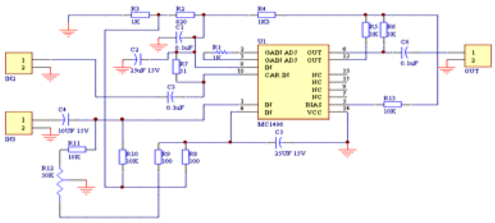


Figure 7. Mixer Circuit

3.6. Design and Realization of Amplifier

Realization amplifier for GPR in this research are Class A Amplifier using transistor 2N 5179. Specification for the component are:

1.	C1 :	(3-35) pF	7.	C10 :	1 nF
2.	C2 :	(2-10) pF	8.	C11 :	0,1 µF
3.	C3 :	(1-5) pF	9.	C12 :	0,02 µF
4.	C4 :	(2-10) pF	10.	R1 :	91 Ohm
5.	C5 :	(2-10) pF	11.	R2 :	10 KOhm
6.	C6 :	1,2 nF	12.	L1 :	10 mH
7.	C7 :	1,2 nF	13.	L2 :	10 mH
8.	C8 :	0,1 µF	14.	L3 :	1 µH
9.	C9 :	1,2 n	15.	L4 :	10 mH

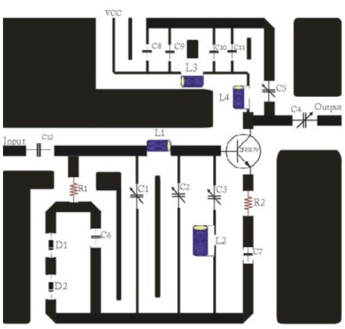


Figure 8. Amplifier Circuit.

3.7. Design and Realization of BPF

Realization of BPF in pulse generator is Narrow band pass Filter Butterworth with π type.

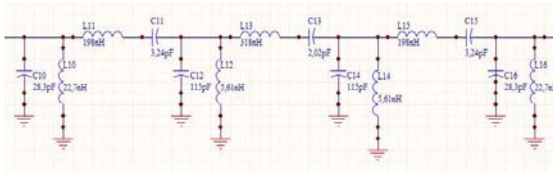


Figure 9. BPF Circuit

For design this filter, we used:

- Rs = Rl = 50 Ohm
- Using order Butterworth (n) = 7, component normalized *LPF Butterworth* :
  - C10 = C16 = 0,445
  - L11 = L15 = 1,247
  - C12 = C14 = 1,802
  - L13 = 2
- Transform LPF to NBPF
  - C component pararel to L component
    - C10 = L10 = 0,445
    - C12 = L12 = 1,802
    - C14 = L14 = 1,802
    - C16 = L16 = 0,445
  - L component serial with C component
    - L11 = C11 = 1,247
    - L13 = C13 = 2
    - L15 = C15 = 1,247
- Denormalized NBPF

$$\text{Cparalel} = \frac{Cn}{2\pi * R * B} \quad (5)$$

$$\text{Lparalel} = \frac{R * B}{2\pi * fl * fu * Ln} \quad (6)$$

$$\text{Cseri} = \frac{B}{2\pi * fl * fu * R * Cn} \quad (7)$$

$$\text{Lseri} = \frac{R * Ln}{2\pi * B} \quad (8)$$

With the formula (5),(6),(7),(8) we get:

- C10 = C16 = 28,330 x 10^{-12} F = 28,330 pF
- L10 = L16 = 22,708 x 10^{-9} H = 22,708 nH
- C11 = C15 = 3,241 x 10^{-12} F = 3,241 pF
- L11 = L15 = 198,446 x 10^{-9} H = 198,446 nH
- C12 = C14 = 114,719 x 10^{-12} F = 114,719 pF
- L12 = L14 = 5,608 x 10^{-9} H = 5,608 nH
- C13 = 2,021 x 10^{-12} F = 2,021 pF
- L13 = 318,31 x 10^{-9} H = 318,31 nH

3.8. Design and Realization of Matching Impedance Circuit

In this research IMC that used is matching impedance multi element (low Q). IMC circuit is realized by cascade IMC L-sections

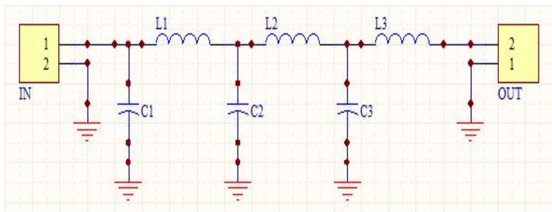


Figure 10. IMC Circuit

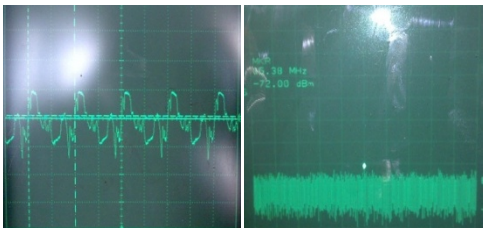
$$Q = 2 \quad (9)$$
$$Qs = \frac{Xs}{Rs} \quad (10)$$
$$Qp = \frac{Rp}{Xp} \quad (11)$$
$$L = \frac{Xl}{2\pi * f} \quad (12)$$
$$C = \frac{1}{2\pi * f * Xc} \quad (13)$$

- IMC *Multivibrator – Mixer*
  Rs = 92,86 Ohm
  Rl = 18 MOhm
  - L Right first level
    Rv_1 = 464,29 Ohm
    Xs = Xl = 185,71 Ohm
    Xp = Xc = 232,14 Ohm
    L_1 = 1,92 µH
    C_1 = 22,28 pF
  - L Right second Level
    Rv_2 = 2,32 KOhm
    Xs = Xl = 928,57 Ohm
    Xp = Xc = 1,16 KOhm
    L_2 = 9,61 µH
    C_2 = 8,92 pF
  - L Right third level
    Xs = Xl = 4,64 KOhm
    Xp = Xc = 9 KOhm
    L_3 = 3,69 µH
    C_3 = 0,09 pF
- IMC *Multiplier Frequency I -II*
  Rs = 6,8 MOhm
  Rl = 34 KOhm
  - L Right Fist level
    Rv_1 = 170 Kohm
    Xs = Xl = 13,6 Mohm
    Xp = Xc = 85 Kohm
    L_1 = 21,645 mH
    C_1 = 0,0187 pF
  - L Right second level
    Rv2 = 170 KOhm
    Xs = Xl = 340 KOhm
    Xp = Xc = 85 KOhm
    L2 = 541,13 µH
    C2 = 0,0187 pF
  - L Right Third level
    Xs = Xl = 340 KOhm
    Xp = Xc = 17 KOhm
    L3 = 541,13 µH
    C3 = 0,0936 pF
- IMC *Multiplier Frequency II – Mixer*
  Rs = 8 MOhm
  Rl = 44 KOhm
  - L Right first level
    Rv1 = 220 Kohm
    Xs = Xl = 16 Mohm
    Xp = Xc = 110 Kohm
    L1 = 25,465 mH
    C1 = 0,0145 pF
  - L Right Second level
    Rv2 = 220 KOhm
    Xs = Xl = 440 KOhm
    Xp = Xc = 110 KOhm
    L2 = 700,28 µH
    C2 = 0,0145 pF
  - L Right Third level
    Xs = Xl = 440 KOhm
    Xp = Xc = 22 KOhm
    L3 = 700,28 µH
    C3 = 0,0723 pF
- IMC *Mixer* – BPF
  Rs = 17 MOhm
  Rl = 50 Ohm
  - L Right First level
    Rv1 = 250 Ohm
    Xs = Xl = 34 MOhm
    Xp = Xc = 125 Ohm
    L1 = 27,056 mH
    C1 = 6,366 pF
  - L Right Second Level
    Rv2 = 250 Ohm
    Xs = Xl = 500 Ohm
    Xp = Xc = 125 Ohm
    L2 = 1,592 µH
    C2 = 6,366 pF
  - L Right Third level
    Xs = Xl = 500 Ohm
    Xp = Xc = 25 Ohm
    L3 = 1,592 µH
    C3 = 31,831 pF

4. Analysis

4.1. Measurement of Multivibrator

Frequency	: 15,15 MHz
Period	: 66 ns
Duty Cycle	: 18,18 %
Vp-p	: 42,50 mV
Vrms	: 11,72 mV
Vav	: 5,53 mV



(a) (b)
Figure 11. Square Signal
(a) Time Domain, (b)Frequency Domain

Frequency multivibrator is 15.15MHz closer to design, shifting to 0.23 MHz because of timer IC is not better at frequency 15.38MHz, and limitation of component.

4.2. Measurement Multiplier Frequency

Frequency	: 200,63 MHz
Period	: 5 ns
Duty Cycle	: 49,5 %
Vp-p	: 125,30 mV

Vrms : 45,27 mV
Vav : 20,18 mV

(a) (b)

Figure 13. Monocycle Pulse Signal
(a)Time Domain, (b)Frequency Domain

Frequency in realization is matching to design, 200 MHZ. There is shifting 0.5 MHz at PRF because of shifting frequency at multivibrator, doesn't precision at matching between block, and soldering techniques doesn't better.

4.3. Measurement of Amplifier

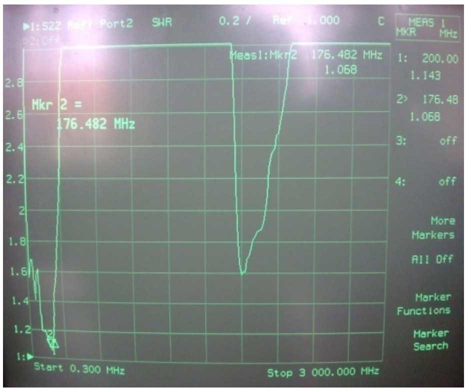


Figure 15. Measurement of VSWR and Bandwidth Amplifier

From measurement, VSWR at 200MHz is 1.143(≤1.5) and Γ equal to 0.067. Bandwidth signal measured from different upper frequency and lower frequency at VSWR ≤1.5 is 101,21 MHz. From measurement of output power of amplifier and output BPF, gain is equal to 14 dB.

Table 2 Measurement to Every Block

Blok	Osiloskop	*Spectrum Analyzer*
Multivibrator	Vp-p = 42,50 mV, periode = 66 ns, *duty cycle* = 18,18 %	Frequency15,15 MHz, *bandwidth* 1,4 MHz, daya -72,00 dBm
Multiplier frequency	Vp-p = 125,30 mV, periode = 5 ns, *duty cycle* = 49,50 %	Frequency *bandwidth* 51,4 MHz, daya -63,00 dBm
Multivibrator, multiplier frequency, mixer, dan *BPF*	Vp-p = 118,75 mV, prf = 14,88 MHz, pri = 67,2 ns, periode = 13,44 ns *duty cycle* = 20 %	Frequency 200,67 MHz, *bandwidth* 50,96 MHz, daya -56,83 dBm
Multivibrator, multiplier frequency, mixer, BPF, dan *amplifier*	Vp-p = 250,63 mV, prf = 14,88 MHz, pri = 67,2 ns, periode = 13,44 ns *duty cycle* = 20 %	Frequency 200,67 MHz, *bandwidth* 50,70 MHz, daya -42,83 dBm

Table 3 Comparison Design and Realization

No.	Parameter	Design	Realization
1.	Waveform	Monocycle	Monocycle
2.	Frekuensi	200 MHz	200,67 MHz
3.	Bandwidth	> 40	50,70 MHz
4.	PRF	15,38 MHz	14,88 MHz
5.	PRI	65 ns	67,2 ns
6.	Power	13 dBm	-42,83 dBm
7.	Tegangan	± 1 Volt	250,63 mV

4.4. Power and Bandwidth

Power level in realization is not matching to design, where power level in realization is -42.83dBm, and power level in design is 13 dBm. Gain in Amplifier is 14 dB. Voltages in realization isn't matching to design where voltages design is 1 Volt and 250.63mV in realization. Bandwidth design is matching to design because of there is shaping Right L third level of matching impedance circuit that used between block.

PRI and PRF realization is closer to design, but there is shifting point because of multivibratorhave hifting period 65ns to 66 ns, frequency 15.38MHz to 15.15MHz. Shifting in period and frequency is effected by IC timer that used is not good, doesn't precision of component.

5. Conclusion and Suggestion

5.1. Conclusion

1. From the simulation, multi vibrator block as pulse generator give unstable pulse wave because of IC timer work as ineffective at frequency 15,38 MHz. Multiplier frequency block as sinusoidal generator can generate sinusoidal wave with ripple.
2. Pulse generator can produce waveform monocycle at 200MHz, but there is a ripple in the signal, because of unstable pulse wave that produce by multi vibrator, unmatched time multi vibrator and multiplier frequency, distortion from jumper, and using single layer PSB.
3. Amplifier can give good performance at frequency 200MHz, with VSWR 1,143 and gain 14dB, because of using high gain transistor (max 20dB), silicon transistor NPN type 2N 5179.
4. From measurement , we get bandwidth 50,70 MHz suitable with requirement of UWB (>40 MHz), PRI 67,2 ns, and PRF 14,88MHz.
5. Using high frequency component is effect to performance of generator pulse that designed.
6. Soldering technique for high frequency can effect to performance of generator pulse

5.2. Suggestion

1. Using another technique of timing switched like PLL to generate sinusoidal wave, or using PIN diode to generate monocycle pulse.

2. Realization pulse generatorfor another frequency.
3. Using double layer PCB.

6. Reference

[1] Daniels, David. (2004). *Ground-penetrating radar. - 2nd ed. - (Radar, sonar, navigations & avionics).* London: The Institution of Electrical Engineers.

[2] Bluhm, Hansjoachim. (2006). *Pulsed Power Systems Principles and Applications.* Germany: Springer-Verlag Berlin Heidelberg.

[3] Daniels, David. (2004). *Surface penetrating radar.* London: The Institution of Electrical Engineers.

[4] Baum, Carl E., Alexander P. Stone, J. Scott Tyo. (2007). *Ultra-Wideband, Short-Pulse Electromagnetics 8.* USA: Springer Science+Business Media, LLC.

[5] Boylestad, Robert. and Louis Nashelsky. (1992). *Electronic Devices And Circuit Theory*. New Jersey: Prentice Hall International Editions.

[6] Collin, Robert E. (1992). *Foundation for Microwave Engineering.* USA : McGraw Hill.

[7] David M, Pozar. (2003). *Microwave Engineering*. JWS. Singapore.

[8] Vendelin, George D. (1990). *Microwave Circuit Design*. Singapore.

[9] Team Dosen Elkom. *Handout Elektronika Komunikasi (TE3623).* Teknik Elektro STT Telkom. Bandung: STT Telkom.

CONTENS OF GOLD AND SILVER ESTIMATION USING ORDINARY KRIGING METHOD (CASE STUDY : PT ANEKA TAMBANG)

I Ketut Gede Adhy Pradnyana[1], Sri Suryani2, Rimba Whidiana[3]

[1,2,3] Institut Teknologi Telkom
adhypradnyana@yahoo.com[1], ssi@ittelkom.ac.id[2], rmb@ittelkom.ac.id[3]

Abstract

One activity performed within eksploration phase is obstetrical estimation of mineral at one particular block. This paper studied about estimation process by analysing and manufacturing data of preexist mining point to get content of mineral in other point. The estimation result is got by conduct the statistically approach using Ordinary Kriging method. Today, the estimation method is done by analysing land sample at points around mine block. The result is estimation about the womb of mineral in that points. Ordinary Kriging is a method that is used to estimate mineral content at a particular mine block. The method is determined by three factor, those are x absisca, y ordinate, and direction factor. The later result is accurate information about womb of gold and silver which not yet been exploited at an one particular block mine. Other, it will be yield a contour Figure about mine's stream pattern of gold ore and silver ore. The result of this paper is useful as consideration in feasibility study plan to exploit a new mine point.

Key words : Estimation, Semivariogram, Ordinary Kriging.

1. Introduction

1.1 Background

In the mining area there are several stages to go through, which included prospecting (general investigations), exploration, evaluation (feasibility study), mine plant design & finance (plant design and cost), development (preparation), exploitation (mining), and processing. At these stages, exploration is a basic step that determines the success of a mining activity. Exploration is a step to estimate / assess the location and amount of the deposits reserve.

Estimation methods which still in use today is by analyzing the womb of the soil samples at points around the block the mine. The result is an estimate of the level of minerals content in that point.

Ordinary Kriging will be examined as a method of estimation which is determined by three factors as mentioned earlier, with gold and silver mines as a sample material to be used. This method find approximate solutions for gold and silver levels by analyzing and processing mining point data of pre-existing content for minerals in the other point. This method is determined by three factors, namely factors abscissa x, the ordinate y and direction factor. In this paper, will be discussed about the determination of drilling points that have not been exploited. In addition the contour map will be shown from the spread of gold and silver content in a exploration block.

1.2 Problem Definition

Problem that will be described and investigated are :

1. How do we get accurate information about the womb of minerals in the points that have not been exploited in a mining block.
2. How do we know the mining material flow pattern .

1.3 Research Objectives

The writing of this paper to design and implement an application system that aims to:

1. Provide accurate information about the womb of gold and silver ore in the points that have not been exploited in a mining block by using method of Ordinary Kriging
2. Produce contours of the womb of gold and silver ore in a mining block.

1.4 Problem Solving Method

1. Literature study, conducted by looking for information and references from books, magazines, articles and internet related to this topic.
2. Observations through data collection and information about the two-dimensional coordinates of the point / position of existing mining and gold and silver levels contained

2. Basic Theory

2.1 Spatial Data

Spatial data is a measurement data includes information of a location [3]. It is to be input in the process of estimation of gold and silver content. Let s_i, $i = 1, \ldots, n$, is a location with coordinates (x_i, y_i). Thus, $Z(s_i)$ is a value of measurement data on site content or Z coordinates s_i. Spatial data is one of the dependent data models, because It collected from different locations spatial indicate dependence between the measurement data with the location.

2.2 Eksperimental Semivariogram

Semivariogram experimental is result of measurements We defined it as average value of the sum of squared differences from the womb of two or more points have different distance vector for h units [4]. Semivariogram value of the experimental data obtained by the formula:

$$\gamma^*(h) = \frac{1}{2\,N(h)} \sum_{i=1}^{N(h)} [Z(s_i + h) - Z(s_i)]^2 \qquad (1)$$

where:

$\gamma^*(h)$: experimental semivariogram
H : distance between two points
s_i : coordinat sample
$Z(s_i)$: data at location s_i
$N(h)$: number of pairs (s_i , $s_i + h$) wihch has distance h

Semivariogram PLOT $\gamma^*(h)$ to distance h is a the experimental semivariogram plot. Experimental semivariogram graph is used as the basis of theoretical semivariogram model for estimation processes.

2.3 Theoretical Semivariogram Model

The pattern of experimental semivariogram is obtained form data irregulary, so it cause dificutly in interpretation and can not be used directly in prediction processes. So a prediction or estimation processes need a theoretical semivariogram model.

Theoretical semivariogram is a value obtained which it based on model parameter, range and sill in accordance with the experimental semivariogram graph form [8]. Parameters in the theoretical semivariogram is [1]:

1. *Range,* is the maximum distance where there is still a correlation between the data.
2. *Sill,* is the semivariogram value not changed to h is not limited. *Sill* value is generally close to the data variation.

There are some models used in the process of fitting theoretical semivariogram model, namely [8]:

A. Nugget Effect

General form is :

$$\gamma(h) = \begin{cases} 0, & h = 0 \\ c, & h > 0 \end{cases} \qquad (2)$$

where :
$\gamma(h)$ = theoretical semivariogram
c = sill.
h = distance between two points

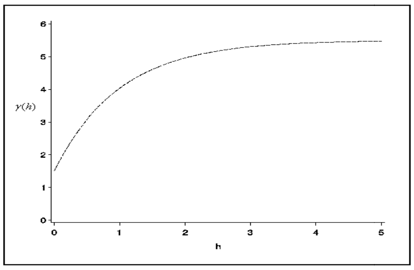


Figure 2.1 Graph of theoretical semivariogram Nugget Effect

This model is used to describe the phenomenon of discontinuity at the base point semivariogram. This phenomenon arises because the distance is smaller than the shortest distance between two points is not known the semivariogram model.

B. Spherical Model

General form is :

$$\gamma(h) = \begin{cases} c\left[\dfrac{3h}{2a} - \dfrac{1}{2}\left(\dfrac{h}{a}\right)^3\right] & , h \le a \\ c & , h > a \end{cases} \qquad (3)$$

where :
$\gamma(h)$ = theoretical semivariogram
c = sill.
h = distance between two points
a = range

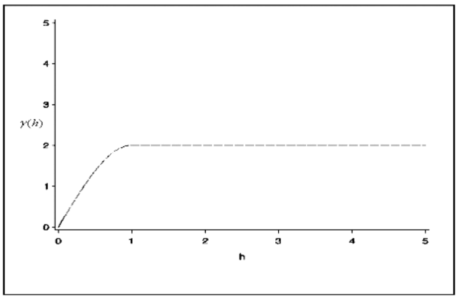


Figure 2.2 Graph of theoretical semivariogram Spherical Model

Spherical model is a mostly used model. This model has a simple polynomial expression and form in accordance with the frequently observed semivariogram: semivariogram values increase almost linearly towards at a particular value.

C. Exponential Model

General form is :

$$\gamma(h) = c\left[1 - \exp\left(-\frac{h}{a}\right)\right] \qquad (4)$$

where :
$\gamma(h)$ = theoretical semivariogram
c = sill.

h = distance between two points
a = range
Practical range of this model is 3a, because at this distance can be achieved 95% of the limit value of semivariogram

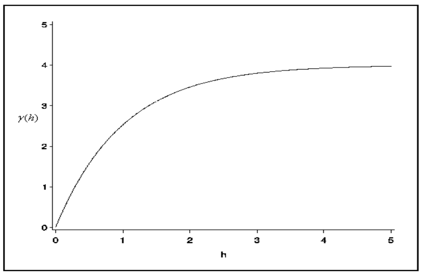


Figure 2.3 Graph of theoretical semivariogram Eksponensial Model

D. Linear Model

General form is :

$$\gamma(h) = h \quad (5)$$

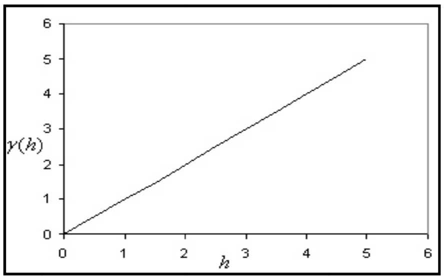


Figure 2.4 Graph of theoretical semivariogram Linear Model

where :
γ(h) = theoretical semivariogram
h = distance between two point

E. Gaussian Model

General form is :

$$\gamma(h) = c\left[1 - \exp\left(-\frac{h^2}{a^2}\right)\right] \quad (6)$$

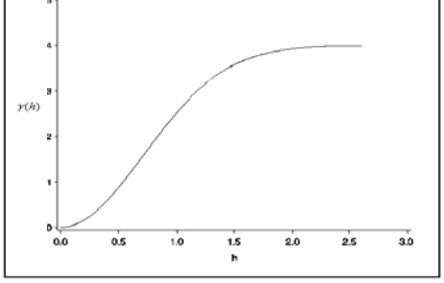


Figure 2.5 Graph of theoretical semivariogram Gaussian Mode

where :
γ(h) = theoretical semivariogram
c = sill.
h = distance between two points
a = range
Practical range of this is 1,73a. Gaussian model describes a continuous event.

2.4 Validation of Theoretical Semivariogram Model

Validation of the model is a testing process of the model. Theoretical semivariogram model which obtained should be tested at first time so it can be used as the estimate processes. The validation of this model is based on statistical test involving the residual statistics and the statistical Q_1 Q_2 [6].

Statistic Q_1: $$Q_1 = \frac{1}{n-1}\sum_{k=2}^{n} \varepsilon_k \quad (7)$$

Statistic Q_2: $$Q_2 = \frac{1}{n-1}\sum_{k=2}^{n} (\varepsilon_k)^2 \quad (8)$$

$$\varepsilon_k = \frac{\delta_k}{\sqrt{\sigma_{OK}^2}} \quad (9)$$

$$\delta_k = Z(s_k) - \hat{Z}(s_k) \quad (10)$$

Where :
$Z(s_k)$ = actual content of k point
$\hat{Z}(s_k)$ = value of estimate at k point
${\sigma_k}^2$ = variance of estimate at k point
n = number of sample point

Value of $Z(s_k)$, $\hat{Z}(s_k)$, ${\sigma_k}^2$ obtained by taking a point from the sample, then perform kriging process to that point by using one or more (n -1) other sample points as the sample data. The process was carried out at all sample points.
Model accepted if:

1. $$|Q_1| < \frac{2}{\sqrt{n-1}} \quad (11)$$
2. $$L < Q_2 < U \quad (12)$$

Value of L and U variable are obtained from the percentile tables.

2.5 Ordinary Kriging Method

Ordinary Kriging Method estimate value at s_0 point using data values from a sample of n $_\alpha$ s

neighbors and combine them linearly with weighting ω_α [8].

$$\hat{Z}(s_o) = \sum_{\alpha=1}^{n} \omega_\alpha^{OK} Z(s_\alpha) \quad (13)$$

Where ::

$\hat{Z}(s_o)$= value of estimate at s_0 point

ω_α^{OK} = value of the weight data (from OK system)

$Z(s_\alpha)$= Value of minerals content in the sample point

n = Number of samples involved in the estimation process

It is assumed that the existing data is part of the realization of *intrinsic random* functions $Z(x)$ the semivariogram $\gamma(h)$.

Variance of the estimate is:

$$\sigma_{OK}^2 = \mu_{OK} + \sum_{\alpha=1}^{n} \omega_\alpha^{OK} \gamma(s_\alpha - s_o) \quad (14)$$

Where :

σ_{OK}^2 = variance of estimate

μ_{OK} = lagrange parameter (from OK system)

$\gamma(s_o - s_o)$ = theoretical semivariogram of point to be estimated with itself $\gamma(s_\alpha - s_o)$= theoretical semivariogram of point to be estimated with sample point (from OK system)

ω_α^{OK} = value of the weight data (from OK system)

n = Number of samples involved in the estimation process

By minimizing the variance of the estimate of the restriction on weight, is obtained the system Ordinary Kriging (OK):

$$\begin{pmatrix} \gamma(s_1 - s_1) & \cdots & \gamma(s_1 - s_n) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma(s_n - s_1) & \cdots & \gamma(s_n - s_n) & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} \omega_1^{OK} \\ \vdots \\ \omega_n^{OK} \\ \mu_{OK} \end{pmatrix} = \begin{pmatrix} \gamma(s_1 - s_0) \\ \vdots \\ \gamma(s_n - s_0) \\ 1 \end{pmatrix} \quad (15)$$

where :

ω_n^{OK} = Weight of which will be filled on the value

μ^{OK} = Lagrange parameter.

The left side of the above system describes the difference between sample data points, while the right shows the difference between each sample data point to estimate the point .

Ordinary Kriging is an appropriate interpolator in the sense that if s_0 identical with the location data, the estimate value is identical to the data value at that point.

$$Z^*(s_0) = Z(s_\alpha), \quad jika\, s_0 = s_\alpha \quad (16)$$

where:

s_α = point of sample data

s_0 = point which will estimate

$Z(s_\alpha)$ = Value of the content of the sample data points

$Z^*(s_0)$ = Estimate value content in a point

Some factors that affect the accuracy of the kriging estimation, namely [1];

1. The number of samples and data quality of any point or location.
2. Sample position in the reserves or deposits.
3. Sample distance of a point to be estimated.
4. Fluctuations or the spread of the sample (regular or irregular).

3. Analysis

3.3 Sample Data

The sample data is used in this paper obtained from measurements of gold and silver content of the sample points contained in block III vein B Tambang Udang Handak L 500 PT Aneka Tambang, Pongkor Bogor. The data is used as a sample of the data point sample position and levels of gold and silver contained.

Table 4.1 Data of levels gold and silver sample

No	X	Y	LEVELS	
			Au(gr/ton)	Ag(gr/ton)
1	11428	9522	4.083666	190.0893
2	11420	9520	4.197484	190.1953
3	11439	9526	4.151658	190.1623
4	11425	9534	4.070711	190.0693
5	11408	9517	4.063246	190.0593
6	11405	9529	4.122474	190.1244
7	11452	9528	4.083666	190.0864
8	11445	9502	4.154919	190.1494
9	11434	9549	4.118322	190.1144
10	11412	9493	4.238747	190.2483
11	11462	9530	4.202485	190.2063
12	11412	9555	4.170294	190.1693
13	11400	9552	4.282843	190.2793
14	11415	9482	4.063246	190.0792
15	11462	9481	4.252982	190.2495

3.4 Calculation experimental semivariogram

Because simple position is irregular, so calculation process experimental semivariogram carried out with GSLIB software.

3.5 Fitting Theoretical Semivariogram Model

From several trials, the valid theoretical semivariogram model parameters, range, and a sill are obtained in statistical tests Q1 and Q2. Based on 2:11 equations in chapter 2, the estimated model is valid if the absolute value of Q1 is less than 0.5345 (the result of $\frac{2}{\sqrt{n-1}}$, N = 15) and smaller Q2 values of 1.86 (U parameters from percentile table, attachment C) and greater than 0.402 (L parameter of the percentile table, attachment C)

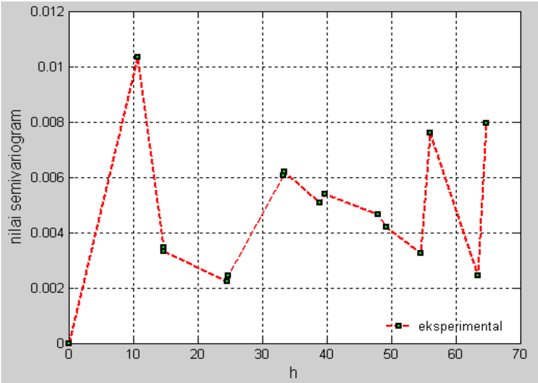
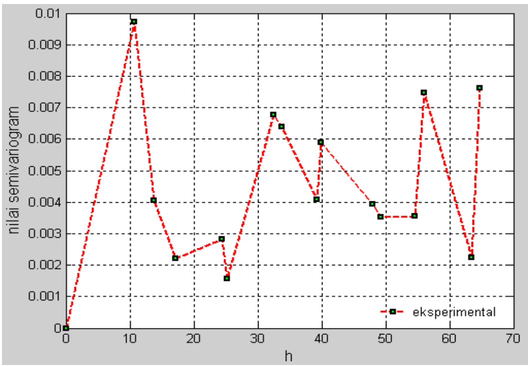


Figure 4.1 Graph Gold Experimental Semivariogram



Figurer 4.2 Graph Silver Experimental Semivariogram

3.6 Estimation process

Estimation process of Ordinary Kriging is done to find value of correlation coefisien as a parameter for estimation accuracy of gold and silver sample.

A. Gold

Estimation is done with input as follows :
Smallest X Range :11400
Largest X Range : 11460
Smallest Y Range : 9500
Largest Y Range : 9560
X Grid : 1
Y Grid : 1
Contour generated along with legend of estimation process of gold above is shown at the Figure 4.3.

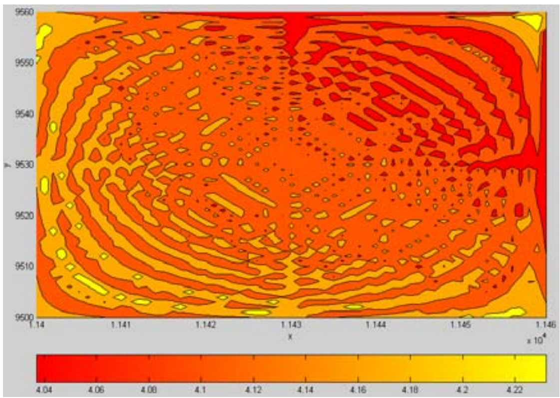


Figure 4.3 Contour of estimation gold contains

B. Silver

Estimation is done with input as follows :
Smallest X Range :11400
Largest X Range : 11450
Smallest Y Range : 9500
Largest Y Range : 9550
X Grid : 1
Y Grid : 1
Contour generated along with legend of estimation process of gold above is shown at the Figure 4.4.

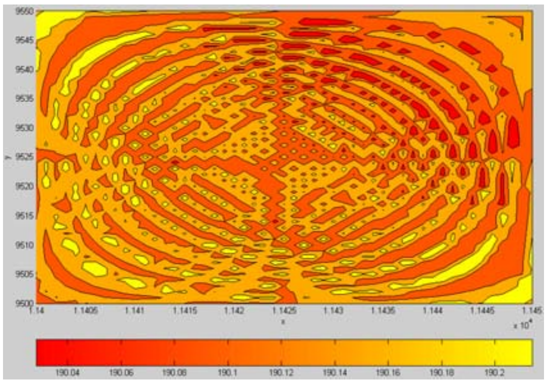


Figure 4.4 Contour of estimation silver contains

3.7 Analysis of result case test

From contour the estimation result of gold mines, sampling should be done on several areas following :

1. Around 11400-11404 abscis and 9541-9550 ordinate
2. Around 11442 -11450 abscis and 9500-9504 ordinate
3. Around 11400-11412 abscis and 9501-9517 ordinate

While from contour the estimation result of silver mines, sampling should be done on several areas following :

1. Around 11400-11405 abscis and 9545-9550 ordinate

2. Around 11402 -11412 abscis and 9535-9546 ordinate
3. Around 11435-11447 abscis and 9502-9511 ordinate
4. Around 11444 -11450 abscis and 9500-9505 ordinate

5. Conclusion and Suggestion

5.1 Conclusion

1. By using the Ordinary Kriging method, can be obtained information about womb of gold and silver on several mining areas.
2. The estimation result is obtained from contour maps that help to determine the location of mines with the most womb of gold and silver.
3. From contour the estimation result of gold mines, sampling should be done on several areas following :
   Around 11400-11404 abscis and 9541-9550 ordinate, 11442 -11450 abscis and 9500-9504 ordinate, 11400-11412 abscis and 9501-9517 ordinate.
4. While from contour the estimation result of silver mines, sampling should be done on several areas following :
   Around 11400-11405 abscis and 9545-9550 ordinate, 11402 -11412 abscis and 9535-9546 ordinate, 11435-11447 abscis and 9502-9511 ordinate, 11444 -11450 abscis and 9500-9505 ordinate.

5.2 Suggestion

For further development, it's needs to add features / modules in software for calculation of the experimental semivariogram is easier so that the estimation process can be used for all types of data

References

[1] Armstrong, Margaret. "Basic Linear Geostatistics". Berlin: Springer-Verlag.1998

[2] Cressie, N.A.C.. "Fitting Variogram Models by Weighted Least Squares, Mathematical Geology". New York: John Wiley & Sons.1985

[3] Cressie, N.A.C.. "Statistic for Spatial Data, Resived Edition". New York: John Wiley & Sons.1993

[4] Deutsch, CV., Journel, A.G.. "GSLIB Geostatistical Software Library and User Guide. New York: Oxford University Press. 1992

[5] Hanselman, Duane dan Bruce Litttlefield. "MATLAB, Bahasa Komputasi Teknik". Yogyakarta: Andi. 2000

[6] Kitanidis, P.K.. "Introduction to Geostatistics: Aplication to Hydrogeology". New York: Cambridge University Press. 1999

[7] Pressman, Roger S.. "Software Engineering A Practitioner's Approach Fifth Edition". Singapore: McGrow-Hill Book Co.1997

[8] Wackernagel, Hans. "Multivariate Geostatistics, An Introduction with Applications". Berlin: Springer-Verlag.1998

[9] Walpole, Ronald E. dan Raymond H Myers. "Ilmu Peluang Statistika untuk Insinyur dan Ilmuwan, edisi ke-4". Bandung: Penerbit ITB. 1995

[10] Stewart, Robert dan Tom Purucker."SADA Technical Support: Using Secondary Sampling Strategies". Tennessee:University of Tennessee. 2000.

SCHEDULING REPLACEMENT OF HYDRAULIC PUMP COMPONENT SOFTWARE USING WEIBULL DISTRIBUTION (CASE STUDY : PT INCO, TBK)

Putu Sagita Purwaningsih[1], Sri Suryani[2], Dade Nurjanah[3]

[1,2,3] Institut Teknologi Telkom
[1] gitaiscute@yahoo.com, [2] ssi@ittelkom.ac.id, [3] dade_nur@plasa.com

Abstract

To days, manufacture industry experiencing of rapid growth. Machine is developed to increase quality and product quality, by in consequence be needed a treatment in an optimal fashion and protected from timeworn component. Research object in this paper is hydraulic pump. Inspection activity need conducted to monitor pressure pump. During the inspection time is done periodically as according to company policy, the question is : " Whether the inspection schedule have effective?" it's necessary to make an analysis to give effective schedule proposal as reasonable and with clarification scientifically . Weibull distribution is an applied method to determine correctly time/schedule replacement for the hydraulic pump component. This distribution is according to calculation of malfunction of component / wear out component. The output of this paper is the recommendation of replacement time of hydraulic pump component and alsot he recommendation of the replacement reverse.

Keywords :Manufacture Industry, Hydraulic Pump, Weibull Distribution

1. Introduction

1.1 Background

One example of a manufacturing machine is hydraulic pump (Vane-pump type) that has important components such as casing (in this case eccentric casing), Vane, and the rotor. Vane serves as an oil seal so that oil pressure can be increased. Vane worked with centrifugal force is caused by rotor rotation. The failure of such components function as Vane cause the hydraulic pump engine does not operate smoothly even total paralysis on the hydraulic pump machine. One indication of distinguishing-mark of the Vane is the pressure outlet of pump which is not produced in accordance with the specifications. Each component has a standard hydraulic pump with minimum allowable pressure. In the operating state of each component of the pump, pressure must not exceed the minimum threshold. Another indication is the knowing generated flow pump. The resulting flow monitoring is rarely done for pump hydraulic system.

The examination is applied during manually which is still controlled by the company's perspective based on the time interval has been setting. The weakness of this condition is not considering the probability of occurrence of a condition that indicates engine damaged will occur. In addition, the frequency (intervals) checks that are not appropriate or too often will need high cost, time and *loss of benefits*. This has encouraged the author to design and implement a system that can optimize the inspection program for the treatment policy is more efficient and effective, using the Weibull distribution, because the distribution is appropriate for the calculation of component malfunction. Despitefully, the authors also discuss the cost analysis is required in the replacement components.

1.2 Problems Definition

In this paper, problems will be discussed are as following :

1. How does the system manage replacement schedule of machine manufacture component.
2. How does the company know minimum cost of component replacement and total loss due to damage of machine manufacture component

1.3 Research Objectives

The writing this paper is to design and implement an application system that aims to :

1. Knowing the optimal replacement schedule time of the machine manufacture component before the damage occurred by using method of Weibull distribution.
2. Knowing average cost of component replacement in the long term with future worth analysis.

1.4 Problems Solving Method

1. Literature study is done by looking for information and references from books, magazines, article, and internet related to this topic.
2. Observation through collect data and information about pressure and performance of machine
3. Testing and system analysis using weibull analysis
4.

2. Basic Theory

2.1 Manufacture Definition

Manufacturing industry is all economic activities that produce goods and services that are not classified as primary products. The primary products are products that are classified as raw material, generated by the exploitation of natural resources of agricultural, forestry, marine and mining, with a possible cover-processing products start to form and the technical specifications and standards commonly traded as a primary products.

2.2 Hydraulic Pump

In the nickel-producing industry, *nickel* ore smelting process using *electric furnaces. Electric furnace,* as the name suggests, uses electrical energy to produce a *spark* (stepping fire) to produce a high temperature. This high temperature is used to melt *nickel* ore.

Furnaces operate continuously for 24 hours. During *furnace* operation, the hydraulic system and pump will also operated. The pump to supply pressurized fluid is used in hydraulic systems.

The pump serves the pressure increment of the fluid to be channeled into the hydraulic system. So the initial fluid pressure equal to atmospheric pressure is increased match to the desired pressure. For the hydraulic system of the *electric furnace,* maximum pressured is about 100 kPa / cm ^{2} or about 1470 psi.

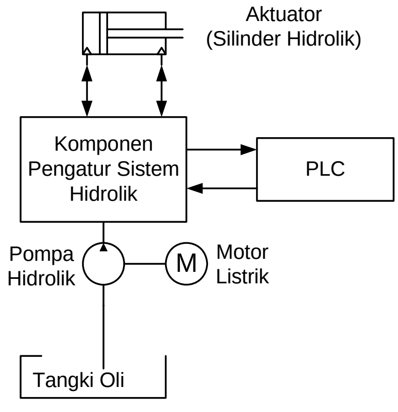


Figure 2.1 Schematic Hydraulic System Settings

2.3 Weibull Distribution

Modern technology has enabled for us to design complex systems depending on the reliability of the various components that build the system. Suppose that a fuse will burn, the steel column will bend or sensing equipment heat-failure. The same components in the same environment will be damaged in different time and unpredictable. In addition to the exponential distribution and gamma distribution, other distributions have been used extensively in recent years to handle the problem is Weibull distribution, introduced by a Swedish physicist Walloddi Weibull in 1939.

In this paper, Weibull distribution is used because the distribution is suitable for the calculation of *components wear out.* Weibull distribution is suitable to describe the component failure where the failure rate is determined by time.

2.3.1 Probability Density Function

Continuous random variable T has Weibull distribution with α and β parameter, probability density function is defined as :

$$f(t) = \alpha\beta t^{\beta-1} e^{-\alpha t^{\beta}} \quad (1)$$

where :

β = weibull slope (beta)

η = characteristic life (eta)

T = variable(tekanan)

2.3.2 Damage rate Function (λ)

Damage rate Function is a function that shows the relationship between component age and failure frequency per unit on T age. Damage rate Function as follows :

$$\lambda(T) = \frac{\beta}{\eta}\left(\frac{T}{\eta}\right)^{\beta-1} \quad (2)$$

2.3.3 *Reliability* Function R(T)

Reliability is defined as the amount of the probability of a device, machine, the system can perform specific functions without failure a specified time interval period in which the operation has been determined.

Reliability function is a function that shows the relationship between the age of the components with the possibility that components will work at a certain time starting from t = 0. Reliability function is shown by the following equation:

$$R(T) = e^{-\left(\frac{T}{\eta}\right)^{\beta}} \quad (3)$$

2.3.4 Probability of Failure

Probability of failure is a possibility failed state that has been formulated achieved. The magnitude of this probability will increase if the age of equipment or system is used longer.

$$POF = 1 - R_{(T)} \quad (4)$$

POF : value of failure probability in system
$R_{(T)}$: Reliability of system

2.3.5 End of Useful Life

End of Useful Life is the final estimate of the age of the equipment or system. Determining the age of POF was considered the maximum allowed aspect and a minimum pressure of the device or system, in which case the minimum pressure hydraulic pump components.

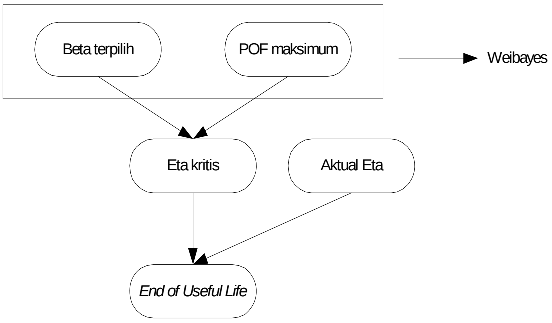


Figure 2.2 Process of Determining End of Useful Life

2.4 Future Worth Analysis

$$F = P\,(1+i)^n \quad (5)$$

or

$$F = P(F/P,i,n) \quad (6)$$

where :
F = cost of the proposed replacement
P = cost of early replacement
I = interest rate
N = period (months)

2.5 Goodness of Fit

Goodness of fit is a test to determine whether the distribution is in accordance with the existing data. In this paper, method of goodness of fit is linear regression model. Linear regression is a forecasting technique that is used to predict *future events* based on past data (historical data). If the values that appear is a straight line form, it can be concluded that the failure of the existing data in accordance with the distribution Weibull [4]. The formula of the linear regression method are:

$$\hat{Y} = a + bX \quad (7)$$

where

$$a = \frac{\Sigma x_i^2\, \Sigma y_i - \Sigma x_i \Sigma x_i y_i}{n\, \Sigma x_i^2 - (\Sigma x_i)^2}$$

$$b = \frac{n\, \Sigma x_i y_i - \Sigma x_i \Sigma y_i}{n\, \Sigma x_i^2 - (\Sigma x_i)}$$

2.6 Determining Parameters of Weibull Distribution

The reason of use Weibull distribution in this analysis is the distribution has a eta and beta parameter for the calculation of characteristic values that are used in determining the estimated and useful life time (the achievement of the minimum pressure allowable). Beta parameter indicate the rate of decline in pressure hydraulic pump components. While eta parameter indicates the average pressure of each pump component that must be controlled to the lowest limit is called the critical eta. β parameter in the Weibull distribution are estimated by $\hat{\beta}$ = b and η by $\hat{\eta}$ = $e^{-a/b}$. Values of a and b obtained from equation 7

3. Analysis

3.1 Description of Data

PT International Nickel Indonesia Tbk (PT Inco) is one of the lowest-cost nickel producers in the world. Data collection for this paper is a hydraulic pump system that a crucial component in nickel mining.

The pump serves to raise the pressure of the fluid to be channeled into the hydraulic system. The production process that occur is a continuous circuit. Failure in one component such as Vane (the propeller) and bearing in the rotor will cause the entire array of disrupted production. To support its function, each component has a pressure different to support the overall pump operation.

The selected research object is vane-pump type. In normal operation, the vane-pump components often damaged in vane (the propeller) and rotor bearing. In the normal operating conditions and run continuously, Vane usuallly reach 3 years and rotors for 4 years.

One indication of distinguishing-mark of the vane is when the pressure pump outlet which produced is not in accordance with the specifications. Here is the criteria for the operation of the vane and rotor components of hydraulic pump

Table 3.1 Criteria for the Operate Component Hydraulic Pump

Component	Minimum Pressure(bar)	Normal Operate (bar)
Vane	70	150
Rotor	105	180

The replacement of the pump components is done when the minimum conditions is achieve due to the conditions under minimum pressure, the pump function is failed.

3.1.1 Data of Pressure Hydraulic Pump Component

Data of pump pressure is data obtained from routinely report that conducted by PT Inco. Here are the data for the vane and rotor components pressure of the hydraulic pump inspection period once every 6 months.

Table 3.2 Data of Pressure Vane Component (bar)

VANE	MONTHS					
	6	12	18	24	30	36
1	124	119	111	95	80	80,5
2	127	117	115	97	83	84
3	124,5	117	105	100	83,5	80
4	125	121	109,5	106	85	81
5	127	118,5	111	101	91	78

Table 3.3 Data of Pressure Rotor Component (bar)

ROTOR	MONTHS							
	6	12	18	24	30	36	42	48
1	175.5	173	170.5	159	145	137	129.5	116
2	175	173	170	164	145.5	139	133	120
3	174.5	173	168	164	147	139.5	134	120
4	173	170	167	158	140	132	129.5	116.5
5	175.5	173,5	170	162	145.5	137	130	118

3.1.2 Data of Operational Cost

Data of cost is needed:

Table 3.4 List of Cost Vane Inspection

Employees Salaries	Rp 20.000,- (per our)
Number Workers	1 person
Equipment Cost	Rp 200.000,- (for use of the pressure gauge)
Old Check	1 our
Total Cost	Rp 220.000 * 6 = **Rp 1.320.000**

Table 3.5 List of Cost Rotor Inspection

Employees Salaries	Rp 20.000,- (per our)
Number Workers	1 person
Equipment Cost	Rp 200.000,- (for use of the pressure gauge)
Old Check	1 our
Total Cost	Rp 220.000 * 8 = **Rp 1.760.000**

Table 3.6 List of Cost Vane Replacement

Employees Salaries	Rp 30.000,- (per our)
Number Workers	2 person
Equipment Cost	Rp 200.000,- (for use of the pressure gauge)
Old Replacement	2 our
Material	Rp 2.000.000
Total Cost	**Rp 2.320.000**

Table 3.7 List of Cost Rotor Replacement

Employees Salaries	Rp 30.000,- (per our)
Number Workers	3 person
Equipment Cost	Rp 200.000,- (for use of the pressure gauge)
Old Replacement	1 our
Material	Rp 1.150.000
Total Cost	**Rp 1.440.000**

3.2 Result of Case Test

Based on calculation above, we obtain graph output which show replacement recommendation for each pump component: vane and rotor.

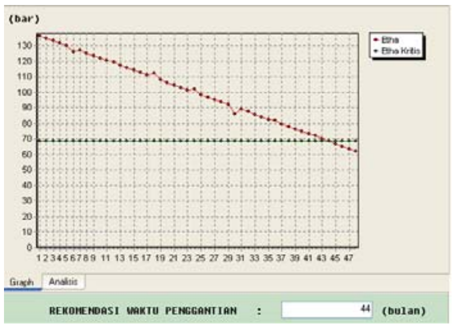


Figure 3.1 Time recommendation of system replacement of vane

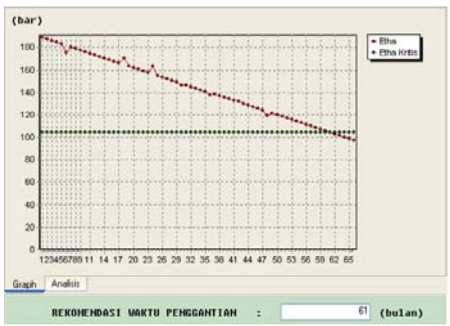


Figure 3.2 Time recommendation of system replacement of rotor

Two pictures above shows that the recommendations when replacement vane components is optimal for 44 months and 61 months of the rotor components.

The value of eta, beta and eta critical to vane and rotor components shown in the two table as follows :

Table 3.8 Value of eta, beta and critical eta of vane component

Month	Beta (β)	Eta (η)	Critical eta
6	84.9783	126.2207	68.7720
12	67.9546	119.3498	
18	30.6328	112.0338	
24	23.8486	101.8038	
30	20.2018	86.5000	
36	36.7315	81.7623	

Table 3.9 Value of eta, beta and critical eta of vane component

Month	Beta (β)	Eta (η)	Critical eta
6	155.6635	175.3491	103.9904
12	108.3352	173.2787	
18	110.3625	169.8483	
24	57.1337	162.7732	
30	50.8249	145.9831	
36	44.9934	138.3755	
42	55.4882	132.3518	
48	61.3577	119.0365	

Beta parameter indicate the rate of decline in pressure hydraulic pump components. While eta parameter indicates the average pressure of each pump component that must be controlled to the lowest limit is called the critical eta.

4. Conclusion and Suggestion

4.1 Conclusion

1. As a decision support system, this system can recommend the optimal replacement time for the hydraulic pump components in the month-44 to vane component and the month-61 to the rotor component.
2. Significant cost savings for 45.9% in the process of replacement components.

4.2 Suggestion

From the evaluation of results and analysis by using Weibull distribution, shows that this distribution can help in the calculation of failure component. For further development, may be used also to determine flow monitoring indications hydraulic pump component usage.

References

[1] Hines, Montgomery. *Probability and Statistics in Engineering and Management Science* , Third Edition. New Jersey : John Wiley and Sons,1990.

[2] http://www.dprin.go.id/kebijakan/05KPINBab1 . pdf

[3] Miller, E Freund. *Probability and Statistics for Engineers,* Third Edition. New Jersey: Prentice Hall International Editions, 1985.

[4] Newman, Donald G. *Engineering Economic Analysis,* Third Edition. California: Engineering Press, 1988

[5] Pressman, Roger S. *Software Engineering A Practitioner's Approach*, Fourth Edition. New York : McGrow-Hill International Editions, 1997.

[6] Pacheco,Xavier and Steve *Teixera. Delphi Developers's Guide.* Los Angeles:Sams Publishing,1994

[7] Supranto J."Statistik". Jakarta:Erlangga, 1994.

[8] Walpole Ronald E , Myers. "Pengantar Statistika". Jakarta: PT Gramedia Pustaka Utama, 1993.

[9] www.barringer1.com

[10] www.inco.com

SPACE-TIME ANALYSIS OF PRODUCT–SUM SEMIVARIOGRAM MODEL (CASE STUDY : OIL PRODUCT)

[1] **Sri Suryani**, [2] **Sutawanir Darwis**

[1] Institut Teknologi Telkom
[2]Institut Teknologi Bandung
[1] ssi@ittelkom.ac.id, [2]sdarwis@math.itb.ac.id

Abstract

Space-time analysis is development of spatial analysis in time variable. Analog to spatial analysis, the main tool of space-time analysis is semivariogram which mesures variability based on lag, direction, and time. One of semivariogram model of space-time analysis is product-sum model, which is a combination of product of spatial semivariogram model with time semivariogram model and sum of spatial semivariogram model with time semivariogram model. In this paper, disscused two theorem which are used to estimate the k parameter and the validation of product-sum model. The analysis of two theorem based on sill parameter of semivariogram model of spatial, time, and spatial-time. This parameter are estimated visually. For some location which the sample is not taken, space-time simple kriging method is used to estimate these parameter. This method is an expantion of spatial simple kriging to time. This paper investigates semivariogram model of product-sum, space-time simple kriging method, and their application on oil product. It shows that product-sum semivariogram model can be used in space-time kriging method with assumptions that k parameter is bounded and covarian matrix is definite positive. The result of this paper is an accurate information about content of oil product in points which not yet been exploited at one particular block mine. Other, it will be yield a contour of mine stream pattern and reserve of oil in block mine . The result of this paper is useful as consideration in feasibility study plan to exploit a new mine point.

Key words : space-time analysis, product-sum semivariogram model, space-time simple kriging

1. Introduction

1.1 Background

Many of the statistical tools that provide qualitative information about the various natural phenomena, but generally do not utilize spatial data information. Spatial data is locations influence data. Then geostatistics appear and develop statistical study of natural phenomena by using spatial data information which called spatial analysis. But at that time, spatial analysis can only be used in parameter estimation process based on location , regardless of the time.

Then in 1992, Christakos developed a method in geostatistics with utilizing information of spatial and time data on natural phenomena, and is called spatial-time analysis. Spatial-time analysis can apply for geological phenomena, environmental pollution, meteorology.

As in the spatial analysis, the main tool in spatial analysis – time is the semivariogram, which measures the variability in the distance, direction and time. Semivariogram model is the empirical model obtained from the data. One of the semivariogram model on spatial-time analysis is the product-sum model. Semivariogram model is derived from the covariance model product – sum introduced by De Cesare et al. (2001). (2001). Model semivariogram product – sum. Semivariogram of product – sum model is a combination of multiplication spatial semivariogram model with time semivariogram model and the sum of spatial semivariogram model with time semivariogram model.

In 2002, Posa et.al used product-sum semivariogram model and the spatial-time analysis to estimate the content of NAO_2 on air pollution in the district.

This paper examined product-sum semivariogram model and simple kriging estimation methods of spatial – time and its application in petroleum production. In the oil industry, the estimate can be used to determine the contour maps and physical data to modeling reservoir. Later, it can be used in the development of new wells to increase oil production in the future.

1.2 Problem Definition

Problem that will be described and investigated are as follows :

1. How do we analysis of *spatial-time for product-sum semivariogram* for oil production

2. How do we get accurate information about the content of minerals in the points that have not been exploited in a mining block.
3. How do we know the mining material flow pattern.

1.3 Research Objectives

The purpose of this study is to examine the product-sum semivariogram model as one of spatial semivariogram model – time, the estimated parameters, validate the model and apply it to estimate oil production at some locations for the future.

1.4 Problem Solving Method

1. Literature study, conducted by looking for information and references from books, magazines, articles and internet related to this topic.
2. Observations through data collection and information about the two-dimensional coordinates of the point / position of existing mining and oil product contained

2. Basic Theory

2.1 Spatial-Time Process

The process Z (s, t) with s states location and t express time is called the spatial–time. The natural phenomena which can be analyzed are oil product, environmental pollution, velocity of wind, meteorology, etc.

Analysis of the spatial-time process consist of structural analysis, parameter estimation and model validation. Structural analysis Phase is the process of fitting the experimental semivariogram model, where contain three stages i.e. fitting model for data of time, spatial, time and spatial combination. The next stage of the parameter estimates determined visually by the three models of experimental semivariogram.

Lets $Z = \{ Z(s,t), (s,t) \in D \times T \}$ is a second order stationary process of spatial-time, where $D \subset \Re^d \; and T \subset \Re_+$. Formula of expectation, covariance, and semivariogram are :

1. $E[Z(s,t)] = m(s,t) = m$ (1)
2. $C_{s,t}(h_s, h_t) = Cov(Z(s+h_s, t+h_t) - Z(s,t))$ (2)

   where $(s, s+h_s) \in D^2$, $(t, t+h_t) \in T^2$
3. $\gamma_{s,t}(h_s, h_t) = \dfrac{Var(Z(s+h_s, t+h_t) - Z(s,t))}{2}$ (3)

2.2 Covariance Model

Here are some covariance model of spatial–time process :

1. Metric Model

$$C_{s,t}(h_s, h_t) = C\left(a^2 \left| h_s \right|^2 + b^2 h_t^2\right) \text{ , } a, b \in \Re \quad (4)$$

where h_s and h_t : lag of spatial and time
This model is introduced by Dimitrakopoulos and Luo, (1994).

2. Product Model

$$C_{s,t}(h_s, h_t) = C_s(h_s)\, C_t(h_t) \quad (5)$$

where C_s = spatial covariance and C_t = time covariance.
Product model is introduced by De Cesare et al., (1997).

3. Linier Model

$$C_{s,t}(h_s, h_t) = C_s(h_s) + C_t(h_t) \quad (6)$$

This model is introduced by Rouhani and Hall, (1989).

4. Nonseparable Model

$$H(\omega, h_t) = \rho(\omega, h_t) K(\omega) \quad (7)$$

where $H(\omega, h_t) = (2\pi)^{-d} \int e^{-ih_s^T \omega} C_{st}(h_s, h_t) dh_s$
and fullfill two condition

– $\forall \omega \in \Re^d, \rho(\omega,.)$ is autokorelation function continuously
– $K(\omega) > 0$ dan $\int K(\omega) d\omega < \infty$

Nonseparable model is introduced by Cressie and Huang, (1999).

5. Product–sum Model

$$C_{s,t}(h_s, h_t) = k_1\, C_s(h_s) C_t(h_t) + k_2\, C_s(h_s) + k_3 C_t(h_t) \quad (8)$$

where $k_1 > 0$, $k_2 \geq 0$, dan $k_3 \geq 0$

2.3 Model of Experimental Semivariogram

Before the theoretical semivariogram model is determined, first time experimental semivariogram fitting is performed. Experimental semivariogram is derived from the known data. there are three of experimental semivariogram models to formulate a product – sum semivariogram model i.e.:

1. Experimental semivariogram of time

$$\hat{\gamma}_{s,t}(0,r_t)=\frac{1}{2|N(r_t)|}\sum_{N(r_t)}[Z(s,t+h_t)-Z(s,t)]^2 \quad (9)$$

where $|N(r_t)|$ # pairs (s, t + h_t) which has distance h_t

2. Experimental semivariogram of spatial

$$\hat{\gamma}_{s,t}(r_s,0)=\frac{1}{2|N(r_s)|}\sum_{N(r_s)}[Z(s+h_s,t)-Z(s,t)]^2 \quad (10)$$

where $|N(r_s)|$ # pairs (s + h_s, t) which has distance h_s

3. Experimental semivariogram of spatial–time

$$\hat{\gamma}_{s,t}(r_s,r_t)=\frac{1}{2|N(r_s,r_t)|}\sum_{N(r_s,r_t)}[Z(s+h_s,t+h_t)-Z(s,t)]^2 \quad (11)$$

where $|N(r_s,r_t)|$ # pairs (s + h_s, t + h_t) which has distance h_s and h_t

2.4 Product-Sum Semivariogram Model

Product-sum covariance model combines the multiplication covariance model of spatial and time, spatial model and time model, which is defined as follows

$$C_{s,t}(h_s,h_t)=k_1\, C_s\,(h_s)\, C_t\,(h_t)+k_2\, C_s\,(h_s)+k_3\, C_t\,(h_t) \quad (12)$$

where Cs and C_t are covariance model of spatial and time, where $k_1>0$, $k_2\geq 0$, dan $k_3\geq 0$. If $h_s=0$ and $h_t=0$, from equation (12) will be obtained value of combination of sill , namely

$$C_{s,t}(0,0)=k_1\, C_s\,(0)\, C_t\,(0)+k_2\, C_s\,(0)+k_3\, C_t\,(0) \quad (13)$$

While combination of sill value is determined visually from respon curve of spatial-time exprimental semivariogram.

Analog with spatial analysis, on product–sum model, combination of sill is included as semivariogram component. By using the relationship between the semivariogram and covariance , we obtain three equations :

1. $\gamma_{s,t}(h_s,h_t)=C_{s,t}(0,0)-C_{s,t}(h_s,h_t)$ (14)

2. $\gamma_{s,t}(h_s,0)=C_{s,t}(0,0)-C_{s,t}(h_s,0)$ (15)

3. $\gamma_{s,t}(0,h_t)=C_{s,t}(0,0)-C_{s,t}(0,h_t)$ (16)

By using these three equations, the covariance model of product – sum, and the combination of sill above, by De Iaco et al., (2001), model (12) developed a product-sum semivariogram model as follows:

$$\gamma_{s,t}(h_s,h_t)=(k_2+k_1 C_t(0))\gamma_s(h_s)+(k_3+k_1 C_s(0))\gamma_t(h_t) - k_1\gamma_s(h_s)\gamma_t(h_t) \quad (17)$$

Where γ_s dan γ_t are semivariogram model of spatial and time, while $C_s(0)$ dan $C_t(0)$ are sill for spatial-time semivariogram model. By using the properties $\gamma(0)=0$, from equation (17) we can obtain two equation as follows :

1. $\gamma_{s,t}(h_s,0)=(k_2+k_1 C_t(0))\gamma_s(h_s)=k_s\gamma_s(h_s)$ (18)

   if $\gamma_s(h_s)$ is estimator for $\gamma_{s,t}(h_s,0)$, then

   $(k_2+k_1 C_t(0))=k_s$ 19)

2. $\gamma_{s,t}(0,h_t)=(k_3+k_1 C_s(0))\gamma_t(h_t)=k_t\gamma_t(h_t)$ (20)

   if $\gamma_t(h_t)$ is estimator for $\gamma_{s,t}(0,h_t)$, then

   $(k_3+k_1 C_s(0))=k_t$ (21)

From equation (13), (19), and (21) are obtained values of k_1, k_2 and k_3 parameter, namely :

$$k_1=\frac{k_s C_s(0)+k_t C_t(0)-C_{s,t}(0,0)}{C_s(0)C_t(0)} \quad (22)$$

$$k_2=\frac{C_{s,t}(0,0)-k_t C_t(0)}{C_s(0)} \quad (23)$$

$$k_3=\frac{C_{s,t}(0,0)-k_s C_s(0)}{C_t(0)} \quad (24)$$

2.5 Determining k Parameter and Validation Model

Validation model of product-sum semivariogram can be seen from value of k parameter, where values of k_1, k_2 and k_3 parameters come from equation (22), (23) and (24). By using equation (17) until (21), $\gamma_{s,t}(h_s,h_t)$ can be expressed, follows :

$$\gamma_{s,t}(h_s,h_t)=\gamma_{s,t}(h_s,0)+\gamma_{s,t}(0,h_t)-k\gamma_{s,t}(h_s,0)\,\gamma_{s,t}(0,h_t) \quad (25)$$

where

$$k = \frac{k_1}{k_s k_t} = \frac{k_s C_s(0) + k_t C_t(0) - C_{s,t}(0,0)}{k_s C_s(0) k_t C_t(0)} \quad (26)$$

Lets Z is a spatial-time process that follows the second order stationer and covariance model follows the form in (12) and continuous in spatial – time. If semivariogram model as in equation (25) and the value of k as in (26), then $k_1 > 0$, $k_2 \geq 0$, dan $k_3 \geq 0$ if and only if k satisfy

$$0 < k \leq \frac{1}{\max\{\mathrm{sill}(\gamma_{s,t}(h_s,0));\ \mathrm{sill}(\gamma_{s,t}(0,h_t))\}} \quad (27)$$

and

$$\gamma_{s,t}(h_s,h_t) = \gamma_{s,t}(h_s,0) + \gamma_{s,t}(0,h_t) - k\,\gamma_{s,t}(h_s,0)\,\gamma_{s,t}(0,h_t)$$

is valid semivariogram model, with

$$k = \frac{C_s(0) + C_t(0) - C_{s,t}(0,0)}{C_s(0)\,C_t(0)}, \quad (28)$$

where by Myers et.al, k_s and k_t are assumed equal 1, then equation (25) ekivalent with equation (12). Because k bounded and equation (28) then $k_1 > 0$, $k_2 \geq 0$, dan $k_3 \geq 0$

2.6 Estimation Method of Kriging Simple for Spatial–Time

In general, interpolation problems are estimation process for parameter of some location that is not sampled. Estimation method which estimating variables regionally at a point, area or on a volume to minimize variance, so there are smoothing of the variability in the location variables were observed. One of the kriging estimation methods, namely simple kriging method of spatial – time which is an extension of simple kriging spatial methods over time. Let

$$Z(s,t) = Y(s,t) + m(s,t) \quad (29)$$

where $m(s,t)$ mean of Z, $E[Y(s,t)] = 0$, and

$$Z^*(s,t) = m(s,t) + \sum_{i=1}^{n} \lambda_i \left(Z(s_i,t_j) - m(s,t)\right) \quad (30)$$

is linear estimator. This estimator not bias and has minimum variance. In order to estimator not bias, then expectation value of error estimation is zero, ie

$$\begin{aligned} E\left[Z^*(s,t) - Z(s,t)\right] &= m(s,t) + \sum_{i=1}^{n} \lambda_i \left(E\left[Z(s_i,t_j)\right] - m(s,t)\right) \\ &\quad - E\left[Z(s,t)\right] \\ &= m(s,t) + \sum_{i=1}^{n} \lambda_i \left(m(s,t) - m(s,t)\right) \\ &\quad - m(s,t) \\ &= 0 \end{aligned}$$

Because the mean of $Y(s,t)$ nol, then automatically estimator not bias. So there is no requirement on the sum of weights. Variance of the error estimate is

$$\begin{aligned} Var\left(Z^*(s,t) - Z(s,t)\right) &= E\left[\left(Z^*(s,t) - Z(s,t)\right)^2\right] \\ &= E\left[\left(Z^*(s,t)\right)^2 + \left(Z(s,t)\right)^2 - 2Z^*(s,t)\,Z(s,t)\right] \\ &= \sum_{i=1}^{n_s}\sum_{j=1}^{n_s}\sum_{k=1}^{n_t}\sum_{l=1}^{n_t} \lambda_{ij}\lambda_{kl}\,C\left(s_i - s_j, t_k - t_l\right) + \\ &\quad C(s-s, t-t) - 2\sum_{i=1}^{n_s}\sum_{k=1}^{n_t} \lambda_i \lambda_k C\left(s_i - s, t_k - t\right) \end{aligned} \quad (31)$$

Because there is no requirement on the sum of weights, then the Lagrange multiplier is not needed. Therefore, kriging system can be stated as follows:

$$C\left(s_i - s_j, t_k - t_l\right)\bar{\lambda} = \bar{C}\left(s_i - s, t_k - t\right) \quad (32)$$

While variance of kriging simple of spatial–time is as follows

$$\sigma^2_{SK} = C_{s,t}(0,0) - \bar{\lambda}'\,\bar{C}\left(s_i - s, t_k - t\right) \quad (33)$$

3. Result Analysis

3.1 Data

The data which used in this paper is secondary data of oil production from wells in 9 locations 14-month period at a center of oil drilling. Coordinates of the location and the average of oil production based on location and time are shown on the following two tables:

Table 3.1 Data of average for oil production from 9 wells location

wells	X	Y	Production average
66	100	100	451.7857
163	200	100	480.2857
59	300	100	557.8571
68	100	200	500.2143
70	200	200	506.3571
56	300	200	498.7857
77	100	300	525.5
144	200	300	495.1429
99	300	300	468.7143

Table 3.2 Data of average for oil production at 14-month period

Time	Production average
1	512
2	551.8889
3	529.4444
4	504.4444
5	472.1111
6	521.1111
7	547.1111
8	517.1111
9	508.6667
10	453.7778
11	439.2222
12	496.7778
13	469.1111
14	453.5556

3.2 Calculation of Descriptive Data

Descriptive of data as follows :

Table 3.3 Descriptive of data

Descriptives of data	Value
Minimum	222
Maximum	763
Mean	498.31
Median	499.50
Variance	9313.175
Standard deviation	96.5
Skewness	–0.08
Kurtosis	0.708

Table 3.4 Descriptive of data after transformation with log

Descriptives of data	Value
Minimum	2.6223
Maximum	2.7325
Mean	2.6887
Median	2.6985
Variance	0.008
Standard deviation	0.09

From plot of oil production , that is shown on figure 3.1, data randomly and stationary.

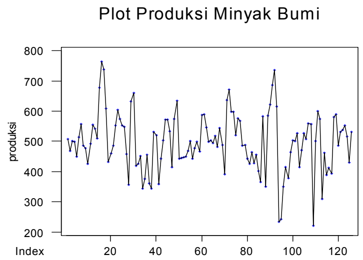


Figure 3.1 Plot of oil production

3.3 Calculation of Experimental Semivariogram

Product-sum semivariogram model is combination of two semivariogram model those are time and spatial semivariogram. But to estimate parameter of product–sum semivariogram model, it needed sill value of time semivariogram, spatial semivariogram, and spatial-time semivariogram. Calculation experimental semivariogram use software GSLIB (Geostatistical Software Library), and Matlab. Result of calculation for three experimental semivariogram are shown on three table as followt :

Table 3.5 Experimental semivariogram of time

Lag of time	Semivariogram
0	0
1	0.00058
2	0.00107
3	0.00109
4	0.00105
5	0.00082
6	0.00094
7	0.00149
8	0.00169
9	0.00168
10	0.00146
11	0.00141
12	0.00185
13	0.00107

Table 3.6 Experimental semivariogram of

spatial to B–T direction

Lag of spatial	Semivariogram	N(h)
0	0	9
100	0.0007	6
200	0.00228	3

Table 3.7 Experimental semivariogram of spatial–time

Lag waktu \ Lag spatial	0	100	200
1	0	0.00833	0.0073
2	0	0.00857	0.00833
3	0	0.00853	0.01003
4	0	0.0084	0.00963
5	0	0.00765	0.0095
6	0	0.00787	0.01083
7	0	0.01027	0.01223
8	0	0.01068	0.0147
9	0	0.01027	0.01777
10	0	0.01027	0.01407
11	0	0.00573	0.01053
12	0	0.00553	0.0161
13	0	0.00268	0.0189

3.4 Theoretical Semivariogram

Based on the three plots of the experimental semivariogram, Theoretical semivariogram model is chosen that approach :

1. Theoritical model of time semivariogram is spherical

$$\gamma(h)=\begin{cases}0.00146\left(1.5\left(\frac{h_t}{10}\right)-0.5\left(\frac{h_t}{10}\right)^3\right) & , 0\le h_t<10\\ 0.00146 & , h_t\ge 10\end{cases}$$

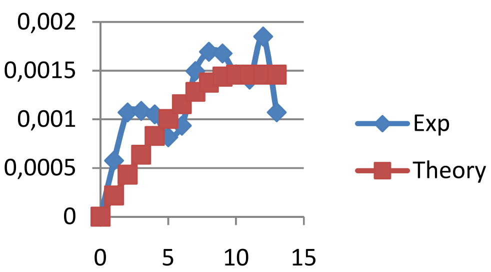


Figure 3.2 Plot of experimental and theoretical semivariogram for spatial

2. Theoretical model of spatial semivariogram is linier with sill

$$\gamma(h)=\begin{cases}\left(\frac{0.00228}{200}\right)h_s & , 0\le h_s<200\\ 0.00228 & , h_s\ge 200\end{cases}$$

:

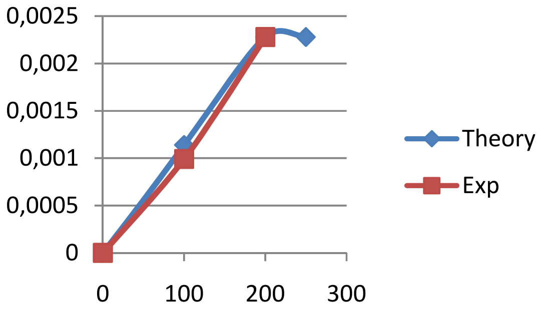


Figure 3.3 Plot of experimental and theoretical semivariogram for time

3. Theoretical model of spatial-time semivariogram with sill = 0.00268 is obtained by interpolation in matlab, as follows :

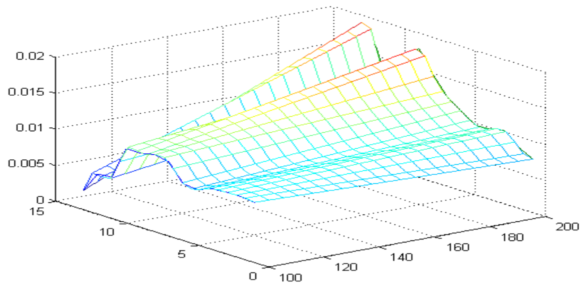


Figure 3.4 Plot of experimental semivariogram for spatial–time

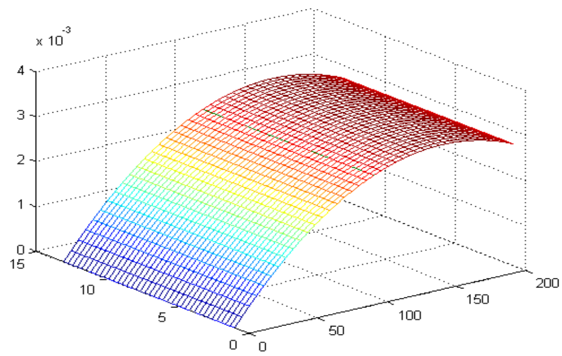


Figure 3.5 Plot of theoretical semivariogram for spatial–time

3.5 Defining Product-sum Semivariogram Model

After we determine three theoretical semivariogram model above, the product – sum model can be defined. By using the value of the third Sill semivariogram model, namely Time sill = 0.00146, spatial sill = 0.00228, and spatial – time sill = 0.00268. The value of k parameter in the model product-sum semivariogram obtained by using equation (35) is 318,433. Product – sum model obtained is

$$\gamma(h_s,h_t)=0.00146\left(1.5\left(\frac{h_t}{10}\right)-0.5\left(\frac{h_t}{10}\right)^3\right)+\left(\frac{0.00228}{200}\right)h_s-$$
$$318.433\left(0.00146\left(1.5\left(\frac{h_t}{10}\right)-0.5\left(\frac{h_t}{10}\right)^3\right)\right)\left(\left(\frac{0.00228}{200}\right)h_s\right) \quad (3.1)$$

where h_s is lag of spatial and h_t is lag of time.

3.6 Cross Validation Model

Validation of product–sum model above use theory. Value of k in model (3.1) is 318.433, while value of $\frac{1}{\text{maks}\{C_s(0),C_t(0)\}}$ is 483.596. So

- $C_{s,t}(0,0) \geq \text{maks}\{C_s(0),C_t(0)\}$ is $(0.00268 \geq \text{maks}\{0.00228, 0.00146\})$
- $k \in \left(0, \frac{1}{\text{maks}\{C_s(0),C_t(0)\}}\right]$ is 0 < (k = 318.433) ≤ 483.596
- Covariance matrix $\left(C_{s,t}(h_s,h_t)\right)$ is definite positive

Validity model can be seen from good of ness model. Plot between $Z^*(s,t)$ *and* $Z(s,t)$, by Minitab can be approached by linear regression model $Z^*(s,t)=0.993\,[Z(s,t)]$. Cause value of b = 0.993, then the model approach regression model $Z^*(s,t)=Z(s,t)$. Calculation by Minitab is shown on table3.8 below :

Table 3.8 Result of regression

Coefisien of b	Standart deviation	T	P–value
0.99	0.006	178.68	0.00

The following table presents an estimate of oil production in 9 months up to 14 months, where 1-month until the 8-month is used as a calibration. In this table, it shows that approximately 60% of the estimated production from 9 to 14 months close to the actual value.

Table 3.9 Result of estimation vs actual production

Coordinate	T	Production	Estimation	Var
(100,100)	9	361	579.43	1.004
(200,100)	9	531	566,76	1.004
(300,100)	9	605	586.14	1.004
(300,100)	10	575	488.09	1.017
(200,200)	10	519	870.76	1.034
(100,100)	11	531	535.30	1.016
(200,300)	11	222	239.49	1.02
(200,100)	12	545	603.95	1.03
(200,300)	12	502	559.11	1.02
(200,200)	13	489	524.69	1.03
(100,300)	13	428	474,24	1.02
(200,200)	14	392	428.75	1.03
(200,300)	14	375	439.14	1.02

3.7 Estimation of Oil Production by Simpel Kriging of Spatial–Time

Estimation of oil prodution on some new wells location and future time can be determined by simpel kriging of spatial–time. Here is the estimate of oil production from 49 locations of new wells to the15-month and 16-month with a 7x7 grid size based on the 14 months period, as follows :

Table 3.9 Result estimation of oil production in new wells location

Coordinate	T	Production	Variance
(120,120)	15	504.55	1.015
(145,120)	15	429.15	1.012
(170,145)	15	277.52	1.016
(170,270)	15	365.09	1.015
(270,120)	16	510.50	1.011
(145,145_	16	416.68	1.01
(195,270)	16	725.77	1.017
(270,270)	16	361.33	1.01

In Figure 3.7 shows that the young blue wider than the other regions. Because on the contour, the blue color indicates that oil production below 237,633. When seen in B Appendix , the location of the oil production of about 237,633 below the 51.24%. Whereas blue indicates production between 237,633 to 472,015, the proportion of about 22:31%. To dark blue, purple, respectively show that oil production over the 1000 and 2000, but relatively smaller proportion than the two colors were. Compared with Figure 3.7, in Figure 3.8 shows that the young blue wider. This is because the production of about 757,262 under the 85.95% . Meanwhile, from 15 to 18 months is not too significant. To contour 19-month to 22-month, the pattern is almost the same.

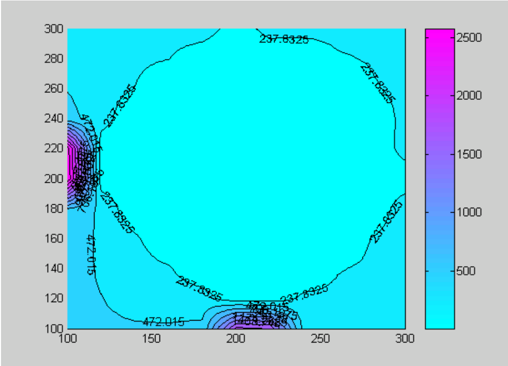


Figure 3.7 Contour of oil production on 121 new wells location at 15-month

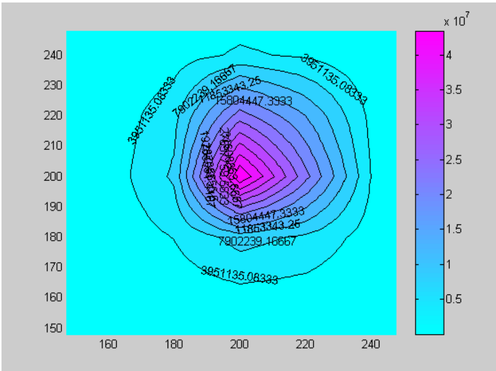


Figure 3.8 Contour of oil production on 121 new wells location at 18-month

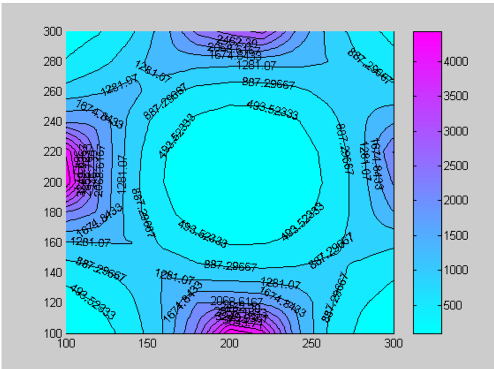


Figure 3.9 Contour of oil production on 121 new wells location at 19-month

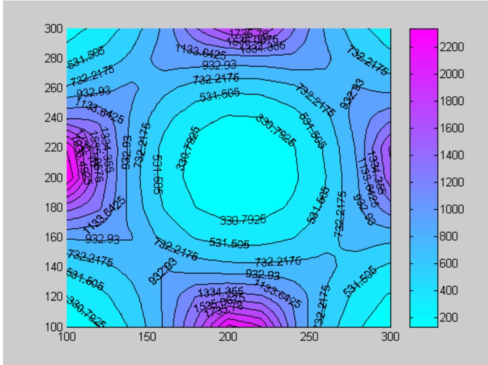


Figure 3.10 Contour of oil production on 121 new wells location at 20-month

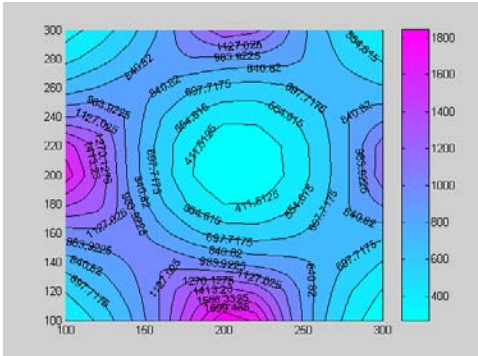


Figure 3.11 Contour of oil production on 121 new wells location at 21-month

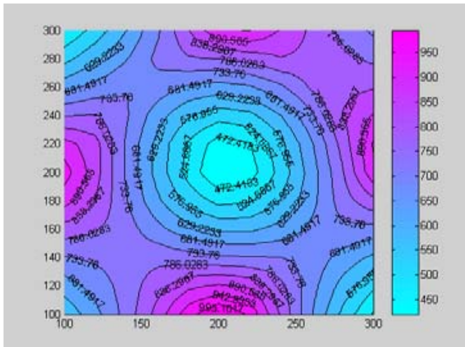


Figure 3.12 Contour of oil production on 121 new wells location at 22-month

From Appendix B, can be seen that oil production from 15 to 16-months tend to fall. Then from 16-month to 18-month , the production tends to rise, and the increase is very significant. This may be caused by: 1) placement of the well location, 2) data are cumulative data of daily production, and 3) the daily sampling time, there was some drilling wells temporarily closed, so this may affect the level of monthly production. Meanwhile, in the next four 19-month to the 22-month, the production tends to fall.

As an illustration, the average of oil production and condesat in Indonesian around 1 million barrels per day (bpd), in which 50% came from the Central Sumatra Basin. As for the oil field as Cepu, between Madura & East Java, is expected to produce approximately 150 thousand bpd. Oil production is not rising so quickly, even the natural tendency falls (EMO & G, 2005). For example, the average oil production and daily condesat in Indonesia in January 2004 1.113.507 bpd, 1,126,634 bpd in February, while March of 1,113,413 bpd (EMO & G, 2004).

4. Conclusion and Suggestion

4.1 Conclusion

1. Product – sum model of oil product to this paper is obtained

$$\gamma(h_s, h_t) = 0.00146\left(1.5\left(\frac{h_t}{10}\right) - 0.5\left(\frac{h_t}{10}\right)^3\right) + \left(\frac{0.00228}{200}\right)h_s - 318.433\left(0.00146\left(1.5\left(\frac{h_t}{10}\right) - 0.5\left(\frac{h_t}{10}\right)^3\right)\right)\left(\left(\frac{0.00228}{200}\right)h_s\right)$$

2. From the contour and estimation of oil production can be shown that oil production from month to month naturally tendency falls.
3. The oil production at some months tends to rise, and the increase is very significant. This may be caused by: 1) placing of the well location, 2) data are cumulative data of daily production, and 3) the daily sampling time, there was some drilling wells temporarily closed.

4.2 Suggestion

1. To estimate the oil production data, we should watch covariate oil reserves, so that estimation accuracy will increase.
2. To improve the accuracy of estimates and further study needs to be semivariogram model with 5 parameters namely k

References

[1] Armstrong, M., (1998). *Basic Linear Geostatistics*, Springer, Berlin.

[2] Cressie, N. A. C., (1985). *Fitting Variogram Models by Weighted Least Squares*, Mathematical Geology 17, 563−586.

[3] Cressie, N. A. C., (1993*). Statistics for Spatial Data*, Resived Edition, John Wiley & Sons, New York.

[4] Christakos, G., (1992). *Random Field Models in Earth Sciences*, Academic Press, Inc, San Diego.

[5] De Cesare L., Posa, D., Myers, D. E., *Estimating and Modeling Space−Time Correlation Structures.*

[6] De Cesare L., Posa, D., Myers, D. E., (2002), *FORTRAN Programs for Space−Time Modeling*, Computers & Geosciences 28, 205−212.

[7] De Iaco, S., Myers, D. E., Posa, D., (2001). *Space−time Analysis Using a General Product−Sum Model*, Statistics & Probability Letters 52, 21−28.

[8] De Iaco, S., Myers, D. E., Posa, D., (2002). *Space−Time Variograms and a Functional from for Total Air Pollution Measurements*, Computational Statistics & Data Analysis 41, 311−328.

[9] Deutsch, C.V., Journel, A.G., (1992), *GSLIB Geostatistical Software Library and User's Guide*. Oxford University Press, New York.

[10] Kitanidis, P.K., (1999). *Introduction To Geostatistics: Application to Hydrogeology*, Cambridge University Press, New York.

[11] Wackernagel, H., (1998), *Multivariate Geostatistics: An Introduction with Applications*, Springer, New York.

CHURN PREDICTION OF CELLULAR TELECOMMUNICATION CUSTOMER WITH COST-SENSITIVE LEARNING APPROACH

Sugihartono[1], Moch. Arif Bijaksana[2], Yuliant Sibaroni[3]

[1,2,3]Institut Teknologi Telkom
[1]mr.sugihartono@yahoo.com, [2]mab@ittelkom.ac.id, [3]yls@ittelkom.ac.id

Abstract

Churn prediction is one of the data mining application to predict the churn customer. Predicting churn is a problem of imbalance class with churn as the minor class. Conventional data mining classification methods does not take loss of benefit due to an error classified a class to another class, this is less appropriate to the problem of churn which is considering benefit as the primary factor in the handling of churn customers. Model evaluations are done through churn prediction model accuracy which are expressed in total benefit, lift top decile, top decile benefit, lift curve, and gini coefficient. A non-cost-sensitive learning algorithm used as the comparison is Boosting with training data which is preprocessed with balancing technique. We investigated two cost-sensitive learning algorithms, Costing and CSRoulette, and our modification of these algorithms, we named them UnderCosting and UnderCSRoulette. Results obtained from the research show that the cost-sensitive learning based sampling method is not always better than non-cost-sensitive learning algorithm in all evaluation parameters. Algorithm based on undersampling: UnderCosting, UnderCSRoulette, and Boosting-UnderSampling are resulted in good performance when cost is low and in high cost resulted in poor performance because low of precision.

Keywords: cost-sensitive learning, sampling, boosting, benefit, cost.

1. Introduction

Churn prediction is one of data mining applications that can support companies in predicting the potential customer churn (stop and / or move to another operator). The number of customer churn each month are relatively small, averaging about 2% compared to the total subscribers (Lu, 2002), but the absolute number of customers who churned and the loss of potential value of money is substantial. The small percentage of the churn group becomes a problem in the model construction because the data is imbalance which means that there are far more records of a class (major class) than the other classes (minor classes) (Yen. et al, 2006). Standard classifiers used on imbalance data tend have low performance, it is because the model tends to give more priority to the major classes or consider minor classes as noise or outlier so they are neglected.

One popular approach in handling imbalance cases is sampling which change the imbalance into balance distribution. This sampling process basically is a balancing process. Balancing technique is generally known in three types:

a. *Oversampling:* doubling minor classes records until their proportion is close to the major classes.
b. *Undersampling:* delete some major class records until the ratio of major and minor classes is almost the same.
c. *Combine over-undersampling:* Combination of Oversampling and Undersampling.

This sampling process can be done with a cost guidance (sampling based cost-sensitive learning). Although cost sensitive learning itself can also be used with other than sampling approach. Cost-sensitive learning considers minor as major record misclassification have a higher cost than major as minor misclassification, resulting a model with the smallest cost (Tan. et al, 2006).

Cost-sensitive learning is one technique in the model development that which take account misclassification cost in its process (Elkan, 2001). Misclassification cost is incurred losses due to data classification error. In general, misclassification costs in the cost matrix described in Table 1.

Table 1. Cost Matrix of Two Class Problem

Cost Matriks	*Predicted Class*		
Actual Class	C(i, j)	Class=Yes	Class=No
	Class=Yes	C(Yes, Yes)	C(Yes, No)
	Class=No	C(No, Yes)	C(No, No)

One approach in cost-sensitive learning is to modify the distribution of training data (Zadrozny. et al, 2003). One of the algorithms that implement this technique is *Costing* (Cost-Proportionate Rejection Sampling with Aggregation) (Zadrozny et. al., 2003) and *CSRoulette* (Cost-Proportionate Sampling Roulette with Aggregation) (Sheng. et. al., 2007).

In this paper churn prediction is implemented with cost-sensitive approach learning by prioritizing algorithm Costing and CSRoulette with accuracy measured by the benefit, lift top decile, top decile benefit, gini coefficient, and the lift curve.

In the case of churn prediction, class misclassification cost of churn group is high, consequently it is very important for corporates to retain their customers since the cost to attract new customers is higher than maintaining existing customers (Mattison, 2002).

In cost-sensitive learning there are two ways to calculate the cost of misclassification, i.e. cost or benefit. An example of cost matrix in Table 2 from Germany credit datasets (Elkan, 2001).

Table 2. Cost Modelling on Credit Card Problem

		Actual Class	
		Bad	Good
Predicted Class	Bad	0	1
	Good	5	0

In the second cost/benefit modeling there is no constant cost matrix, the value of misclassification cost depending on the misclassification cost of each record which may vary considerably between the existing record. The modeling of credit card case above is shown in Table 3.

Table 3. Cost dan Benefit Modeling for Credit Card Case

		Actual Class	
		Fraudulent	Legitimate
Predicted Class	Refuse	$20	-$20
	Approve	- ***x***	0.02***x***

2. Costing and Csroulette Algorithms

2.1 Costing Algorithm (Zadrozny. et al, 2003)

Costing Algorithm (Cost-proportionate rejection sampling with aggregation) is a combination of rejection sampling technique and ensemble. Figure 1 shows the Costing algorithm pseudocode.

```
Costing (Learner A, Sample Set S, count t)
  1. For i=1 to t do
     (a)  S' = rejection sample
          from S with acceptance
          probability c/Z
     (b)  Let h_i ≡ A(S')
  2. Output h(x) = sign (Σ_{i=1}^{t} h_i(x))
```

Figure1. Pseudocode of *Costing* Algorithm

Training data records must be weighted before processed by algorithm Costing. Weighting in the case of churn prediction is as follows:

Record with class *active* = c
Record with class *churn* = R-c
Where:
c : *Cost* customer retention.
R : *Revenue* of customer

Rejection sampling requires random number generator with uniform distribution. This uniform distribution is required in order weightings do not affected the distribution of random numbers used. In rejection sampling, instance is received if random number <= c / Z, where c is the weight of record, and Z is the maximum value of the weight of the data.

2.2 CSRoulette Algorithm (Sheng. et al, 2007)

CSRoulette method have the same scheme with Costing, the difference lies in the samples construction. CSRoulette use Roulette's sampling techniques in selecting the records of training data to be inserted into the sample. CSRoulette algorithm pseudocode is shown in Figure 2.

```
For r = 1 to I do //build I classifiers
   S' = CPRS (T, C)
   Let h_r = B(S')
   Output h(x) = sign (Σ_{r=1}^{I} h_r(x))
```

Figure 2. Pseudocode of Algoritma *CSRoulette*

Roluette sampling as rejection sampling requires a weight for each record. The weigting has the same way as weighting in Costing algorithm. After records has been weighted, the weights of all records mapped in a circle (Sheng. et al, 2007). The width of the circle is proportional to the amount of mapped record's weight. Illustration of mapping record is shown in Figure 3.

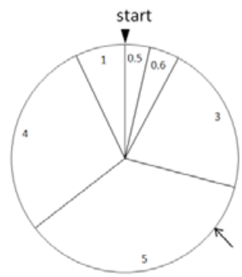


Figure 3. Illustration of *Roulette* Sampling Algorithm

For example if the random number is 7.56 then the selected records are records with a segment wide 5. Factor ***k*** states how large the sample size compared to the size of training data. The value of factor ***k*** must be more than zero. If 0 < ***k*** < 1 then the sample will be smaller than the training data, if ***k***

= 1 then the sample size is equal to the training data, and if ***k*** > 1 then the size is larger than training data.

Roulette sampling algorithm can be developed by increasing the number of needle that will determine where the selected record. The number of needles can be set to *n*, where the needle will be placed on record segment circle with the same distance. This will lead to in a generating random numbers process, the number of selected records is n (Suyanto, 2008).

3. Evaluation Metric

3.1 Imbalance Class Evaluation

Confusion matrix is a tool to express the number of correct and wrong data predicted by a model. A sample of confusion matrix for 2-class classification is illustrated in Table 4.

Table 4. Confusion Matrix on Two-Class Classification

		Predicted Class	
		+	–
Actual Class	+	TP	FN
	–	FP	TP

$$Precision, p = \frac{TP}{TP+FP} \quad (1)$$

$$Recall, r = \frac{TP}{TP+FN} \quad (2)$$

$$F_1 = \frac{2rp}{r+p} = \frac{2 \times TP}{2 \times TP+FP+FN} \quad (3)$$

3.2 Churn Prediction Evaluation

There are some evaluation metrics commonly used in churn prediction, that are *lift curve*, *top decile lift*, *Gini coefficient*, and *top decile benefit*.

Lift curve is obtained by rank and map the results of model predictions into the curve (Neslin, 2003). The lift curve construction can be done by data sorting starting from the highest churn probability until the lowest churn probability. Then proceeded for records that predicted active which sorted, ascending, based on their active probability. Figure 4 is an example of lift curve.

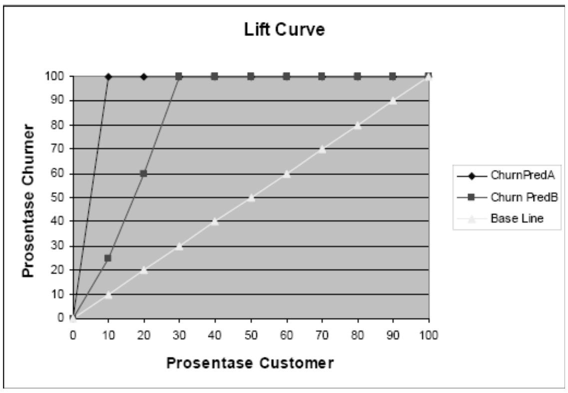


Figure 4. An Example of Lift Curve Comparing Two Prediction Models

Top decile lift is a measure of predictive accuracy to see how many fold increase in churners we found in 10% of the most potential customer churn compared to the percentage of all customers (Neslin, 2003).

$$Top\ Decile = \frac{\hat{\pi}10\%}{\hat{\pi}} \quad (4)$$

- $\hat{\pi}10\%$: churner percentage at the riskiest segment
- $\hat{\pi}$: churner percentage at all customes*r*

Gini coefficient is a churn prediction evaluation to see the accuracy of a classification algorithm at all customers including those that have a low churn probability (Neslin, 2003). Gini coefficient calculation can be done with formula (5).

$$Gini = \left(\frac{2}{n}\right)\sum_{i=1}^{n}(v_i - \hat{v}_i) \quad (5)$$

Where

- o v_i : Churner percentage which their churn probability is the same or more than of *customer* i.
- o $\hat{v}_i$: Customer percentage which their churn probability is the same or more than of *customer* i.
- o n : The number of customer.

The total benefit is the total profit obtained when using the predictive model. In the case of churn prediction the total benefit is calculated by taking into account cost factors and benefit from the prediction result. The calculation of total benefit is the sum of all benefits minus the sum of the calculated cost for each customer.

The calculation of the top decile is the same benefits as the total benefits, but only count for 10% of the total benefit of the most potential customers to churn reduced by the amount of the missing benefits of churn above 10% of the selected customer.

Often the results of several algorithms show almost the same value / differ slightly, but to know the difference is significant or not can be tested by testing the difference between 2 proportions (Walpole, 1993).

4. Experiments

4.1 The Influence of Retention Cost to The Sample Size of Costing Algorithm

We used three kinds of scenarios data with different percentage of retention cost. Figure 5 shows the influence of percentage of retention cost against sample size produced by Costing algorithm.

Costing algorithm sample size increases, following the increase of retention cost as influenced by the boolean value of (**angka_random** <= **c** / **Z**) in the rejection sampling process.

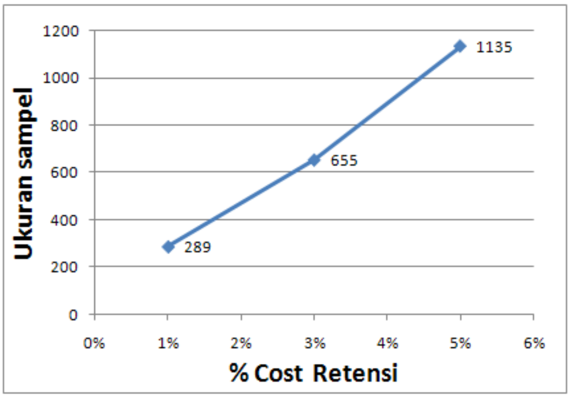


Figure 5. Sample Size of *Costing* Algorithm on Three Retention Cost Scenarios

4.2 The Influence of Retention Cost to The Average Percentage of Sample Imbalance of CSRoulette and Costing Algorithm

The output of the rejection and roulette sampling is a sample with the imbalance percentage is greater than the imbalance percentage of training data. Comparison of the average imbalance sample percentage of the two algorithms are presented in Figure 6.

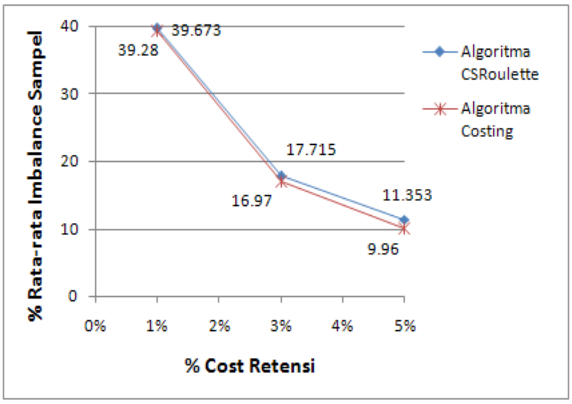


Figure 6. Average Percentage of Sample Percentage on Three Retention Cost Scenarios

In Costing algorithm, the increasing in retention cost percentage is coupled with decreasing in average imbalance percentage of the sample. This is due to increased retention cost causes an increased of the resulting sample size. Because the sample size increases, with the same training data imbalance percentage becomes smaller. It is related to the higher opportunities of boolean value (**angka_random** <= c / Z) with true value.

In CSRoulette algorithm, the increasing value of retention costs will cause the length of churn record smaller because the value of benefit - cost is smaller, while the segment for active record becoming longer. With the increasing size of active record segment, the chance active record to be selected is also greater, so the sample will be increasingly dominated by the active record. Changes in these segments is illustrated in Figure 7.

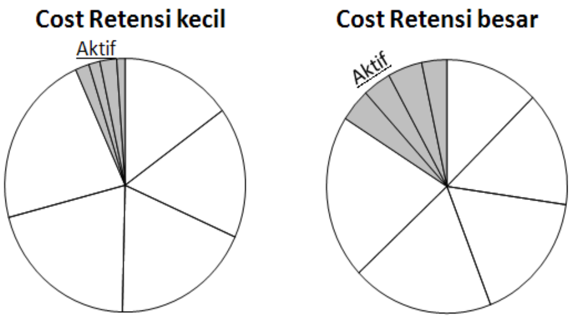


Figure 7. Illustration of Ilustrasi perubahan ukuran segmen record pada algoritma CSRoulette

Table 5 shows the ranking of the best five algorithms on the three scenarios

Table 5. The Ranking of The Best Five Algorithms in 1% Retention Cost Scenario

Total Benefit	*Top Decile Lift*	Luas Area *Lift 10%*	*Top Decile Benefit*
UR1	CS2	BC	BO
UR2	BC	BO	CO
BU	BU	BU	CS2
CO	CO	CS2	BU
UR3	UR1	UR1	CS1

Gini coefficient	*Recall* Minor	*Precision* Minor	*F-Measure* Minor
BC	UR1	BO	BO
BU	BU	CS6	CS6
CO	UR2	CS5	BC
CS2	UR3	BC	CS5
BO	UC	CS4	CS4

Notes:

CO : *Costing*

CS1 : *CSRoulette*, **k**=0.01, needle=1

CS2 : *CSRoulette*, **k**=0.05, needle=1

CS3 : *CSRoulette*, ***k***=0.1, needle=1

CS4 : *CSRoulette*, ***k***=0.5, needle=1

CS5 : *CSRoulette*, ***k***=1, needle=1

CS6 : *CSRoulette*, ***k***=2, needle=1

UC : *UnderCosting*

UR1 : *UnderCSRoulette*, needle=1

UR2 : *UnderCSRoulette*, needle=2

UR3 : *UnderCSRoulette*, needle=3

BC : *Boosting-CobineUnderOverSampling*

BO : *Boosting-OverSampling*

BU : *Boosting-UnderSampling*

5. Conclusion

From the experiments it can be concluded that retention cost influences to the result of cost-sensitive learning algorithms prediction. The higher of costs the more likely model predicts the current record as an active class because the model is built from more imbalance samples. The model more likely failed to predict to churn class as shown from the lower recall. Furthermore Costing algorithm has a low performance in prediction and rank of top 10% customers. In addition, CSRoulette algorithm requires appropriate value ***k*** to give good results. For telecommunications data used in this research the best value ***k*** is 0.05.

Undersampling based algorithms such as UnderCosting, UnderCSRoulette, and Boosting-UnderSampling tend good at a low cost value and bad on the high cost value because they produce low precision. While the number of needles on UnderCSRoulette algorithm is not much effect on prediction results, it will only accelerate the sampling process. The best number of needles from three values tested for UnderCSRoulette algorithm is one.

References

[1] Elkan, C. 2001. The Foundations of Cost-sensitive learning. In *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence (IJCAI'01).* pp 973–978.

[2] Han, Jiawei., Micheline Kamber. 2006. *Data Mining: Concepts and Techniques, 2nd ed*. San Francisco: Morgan Kaufmann Publishers.

[3] Lu, J. 2002. *Predicting Customer Churn in the Telecommunications Industry - An Application of Survival Analysis Modeling Using SAS ®.* Kansas: Sprint Communications Company.

[4] Mattison, Rob. 2002 "*Telco Churn Management: The Golden Opportunity*".

[5] Neslin, Scott. 2003. *Churn Modelling Tournament.* USA: Duke University.

[6] Sheng, Victor S and Charles X Ling. 2007. Roulette Sampling for Cost-Sensitive Learning. In *Proceedings of the 18th European conference on Machine Learning (ECML '07).* pp 724-731.

[7] Suyanto. 2008. Evolutionary Computation - *Komputasi Berbasis "EVOLUSI" dan "GENETIKA"*. Bandung: Informatika.

[8] Tan, P.N., M. Steinbach, and V. Kumar. 2006. *Introduction to Data Mining*. Boston: Pearson Education.

[9] Walpole, Ronald E. 1993. *Pengantar Statistika, 3nd ed*. Jakarta: Gramedia

[10] Yen, S.J., Y.S. Lee, C.H. Lin, and J.C. Ying. 2006. Investigating the Effect of Sampling Methods for Imbalanced Data Distributions. In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics.*

[11] Zadrozny, B., J.C. Langford, and N. Abe. 2003. Cost-sensitive learning by Cost-Proportionate Example Weighting. In *Proceedings of the 2003 IEEE Intl. Conf. on DataMining.* pp 435–442.

THE ANALYSIS AND APPLICATION OF CLASSIFICATION METHODS FOR SOFTWARE DEVELOPMENT *NON-TULIS* UNIVERSITY STUDENT ENROLLMENT SYSTEM (CASE STUDY AT IT TELKOM)

Yuliant Sibaroni[1], Benhard Sitohang[2]

[1]Institut Teknologi Telkom
[2]Bandung, Institute of technology
[1]yls@ittelkom.ac.id, [2]benhard@informatika.org

Abstract

The selection of new students through ″non-tulis ″ route could be predicted based on the GPA value when the students follow the lecture. The problem of this prediction could be carried out by employing the classification method in data mining. The *Classifier* model was developed based on history data of these student's grade during their senior high class year, the major that they choose and score of high school where students is from. The classification methods that were further analyzed was decision tree and *Bayesian Network*. These methods were often used in the problem of the prediction and the level complexity and accuracy of these methods are similar. Based on literature study, method of *C4.5* was chosen from the decision tree because this method is widely used; meanwhile Naïve *Bayesian* was chosen from *Bayesian Network* because theoretically the assumption of this method was regarded to be more precise than the other methods. The experiment result showed, based on the accuracy level, there is no absolute superior method, however generally the accuracy level of these methods was not satisfaction yet. Based on the process time, *C4.5* is better than Naïve *Bayesian*, but their process time are applicable in this data. Based on dispersion of data, there is not relationship between levels of dispersion of data with level of accuracy *C4.5*. From dispersion of data, we can see there are some relationships between attribute so *Naive Bayesian* is not suitable for this data. To be able to be employed, this system still needs improvement to increase the level of their accuracy.

Key Word: prediction, data mining, *Classifier*, *Naive Bayesian*, *C4.5*, accuracy.

1. Introduction

One of the processes of acceptance of the new student in a university was to go through the non-tulis (without test) route. Through this route, the writing test was not carried out, the new registrant will be selected especially being based on the academic achievement during in the Senior High School, score of high school where students is from and the major that they choose.

The determination of the accepted student is from the prediction of the GPA potential that will be achieve the students when follow the lecture in the university. The problem of this prediction could be carried out with the classification method in data mining.

Generally, the classification method consisted of two stages that are learning and classification. The learning stage is the performing process of the classifier model from training-set, whereas the classification stage contains class determination process of the testing -set and the new data.

Two classification methods that will be analyzed further were decision tree and Bayesian. The consideration was the two methods often was used in the problem of the prediction, had the similar complexity and the level of their accuracy is high [1.2.10]

These two methods will be compared to determine the best method. Several criteria to determine the best method are: accuracy, the speed, robustness, scalability and interpretability [1].

2. Basic Theory

2.1 Decision Tree

Decision tree is flow–chart like tree structure, where each internal node denotes a test on an attribute, each a branch represents an outcome of the test, and leaf node represents classes or class distribution [1].

There are some stages in classification process with decision tree, i.e. [1]

- Attribute selection and tree Building
  Attribute selection as root node or internal node was based on impurity measure *from* each attribute. Impurity measures that usually used are i*nformation gain*[1], *gain ratio*[2] and *gini index*[2]. Attribute with highest impurity will be chosen as *attribute test.*
- *Tree Pruning*
  Tree pruning is branch cutting process on the tree to get simplifier model and model with higher accuracy. There are two kinds of pruning, pre pruning and post pruning.
- Classification rules Extraction from *decision tree*
  Classification rules extraction is done after we have the optimal decision tree yang optimal. This process will give output classification rules with IF–THEN rule form.

There are some popular algorithms for tree building, i.e. : *ID3, C4.5* and *CART. C4.5* is successor of *ID3*. *C4.5* and *CART* have similar ability, i.e.: handling continues data, missing value, and has *pruning* process.

2.2 Bayesian

Bayesian classification is based on *Bayes* theorem, i.e. theorem for posterior probability calculation $P(Y \mid X)$. In this classification, *Bayes* theorem is used to calculate class probability of data–set based on given inference data.

Classification rule will be obtained from the highest $P(Y \mid X)$ value from all $P(Y \mid X)$ values that calculated. For the example, Y have 3 class i.e.: Y_1, Y_2 and Y_3, so there are 3 posterior probability that will be calculate i.e. $P(Y_1 \mid X)$, $P(Y_2 \mid X)$ and $P(Y_3 \mid X)$.

Based on relation type between data attribute, there are two Bayesian classification methods, i.e. [1]:

- *Naive Bayesian*
- *Bayesian Belief Network*

2.2.1 Naive *Bayesian*

Naive Bayesian has assumption that relation between an attribute with another attribute is conditional independency for class Y. It is called *naive* because this assumption is difficulty to be happening in reality live. But, this method has high accuracy level for most cases [2].

Naive Bayesian calculates posterior probability for every class in Y to classify a new record. [2]:

$$P(Y \mid X) = \frac{P(Y)\prod_{i=1}^{d} P(X_i \mid Y)}{P(X)}$$

Based on conditional dependency assumption, calculation of $P(Y \mid X)$ can't be done directly. Because P(X) value is fixed for every Y values, so determination class $Y = y_i$ is chosen from class whom has maximum $P(Y = y_i)\prod_{i=1}^{d} P(X_i \mid Y = y_i)$ value.

If *training set has* many attribute, then naïve *Bayesian* formulas above has probability can't predict some records. This problem can be handled using M- estimator. Estimator–M has formula [2] :

$$P(x_i \mid y_j) = \frac{n_c + mp}{n + m}$$

Where

n	:	Sample size in class y_j
n_c	:	Amount of sample in class y_j whom have value x_i
m	:	Sample size (determined by user).
p	:	Parameter whom determined by user, can be supposed as prior probability $P(x_i \mid y_j)$ when n = 0.

2.2.2 Bayesian Believe Network

Bayesian Belief Network or simplifier called *Bayesian Network* use more flexible approach i.e. uses conditional dependency attribute pair. Graph model that descript relation between attributes called *Belief Network* or *Bayesian Network.*

Bayesian Network model building is worked in two step, i.e.: [2]

1. Built network structure
   There are two approaches i.e. [4]
   - *Search and scoring method*

- Dependency analysis method

2. Estimate probability value in the table that is related with every *node*.

   These probability values will produce Bayesian classification rules.

Generally, both of algorithms can be explained as:

S*earch and scoring* method

- Network construction is most suitable structure finding process to data.
- Construction process start from a graft without edge, then finding method is used to add an edge to the graft.
- S*coring* method is used to see weather new structure better than old structure.
- These processes continue until there is no new better structure than old structure.
- Example: K2 algorithm (assumption : node ordering)
- Criteria was used in *search and scoring* method is *Bayesian scoring criterion* ($score_B$)

  $$score_B(d,gp) = score_B(d,G) = P(d \mid G)$$

  Where $P(d \mid G)$ has formula [4]

  $$P(d \mid G) = \prod_{i=1}^{n} \prod_{j=1}^{q_i^{(G)}} \frac{\Gamma\left(N_{ij}^{(G)}\right)}{\Gamma\left(N_{ij}^{(G)} + M_{ij}^{(G)}\right)} \prod_{k=1}^{r_i} \frac{\Gamma\left(a_{ijk}^{(G)} + s_{ijk}^{(G)}\right)}{\Gamma\left(a_{ijk}^{(G)}\right)}$$

  Where $a_{ijk}^{(G)}$ and $s_{ijk}^{(G)}$ are values in $\left(G, F^{(G)}, \rho \mid G\right)$

Dependency Analysis Method

In this method, *Bayesian Network* structure is built with identifies conditional dependency between *node–node*. Determination of dependency is based on statistical test (G^2). This relation then is used as certainty to built *Bayesian Network* structure.

Formula of G^2 statistics [4]

$$G^2 = 2\sum_{a,b} s_{ijk}^{abc} \ln\left(\frac{s_{ijk}^{abc} M}{s_{ik}^{ac} s_{jk}^{bc}}\right)$$

Where

S_{ik}^{ac} is random variable that has same value amount record that contain X_i=a and X_k=c.

s_{ijk}^{abc} is random variable that has same value amount record that contain X_i=a, X_j=b and X_k=c .

3. Analysis

3.1 Preparing of Data

There are some steps in preparing of data i.e.

- *Data cleaning*

  This step is cleaning process to noisy data and missing values. We assume there no noisy data and missing values related with handled data.
- *Data integration and transformation*

  This step is process to integrate some tables to be one table and process to transforms data. Integration process is done manually, while data transformation is not done.
- *Data reduction*

  This step is attribute selection process and discretization process. Attributes that supposed have no relation and redundant is eliminated.

The result of Data preparing is *training set*

Mathematics	Physics	English	Major	Score	GPA
A	B	A	TI	T	T
A	B	B	IF	R	T
B	C	B	TE	S	R
A	B	A	TE	S	R
B	A	C	IF	T	T
.	.	.	.		.

Explanation:

Mathematics, Physics and English are high school grades. We split range of Mathematics, Physics and English into 3 range using equal binning method: A : high, B: Mid, C:Low.

Major is: faculty that student choose, TI: Industrial engineering, IF: Informatics Engineering and TE: Electrical Engineering

Score is high school level; T : high, S: Mid and R: low

GPA : GPA of student, T; high and R :low

We can define T and R manually for example we can define T : GPA ≥ 3,00 and R: GPA < 3,00.

3.2 Evaluation of Performance Classification Method

Based on data, we want a model with highest accuracy so, we will compare performance both of methods based on data prediction accuracy level.

We use *Random Sub sampling method* [2] to measure accuracy level of classifier model. That method is *holdout method* which repeats several time [2]. In *holdout* method, real data will be partition in two separated set: *training set* and *testing set*.

Accuracy for each method can be estimated based on accuracy of *testing set.* Accuracy calculated from percentage of error prediction.

3.3 Determination of *Decision Tree* Algorithm

C4.5 and *CART* algorithms as generally can be applied for similar problem. There is no stronger indication that an algorithm will better than another one. Finally, determination of *decision tree* algorithms based on popularity of those algorithms. *C4.5* algorithm is more popular than *CART*, so *C4.5* is chosen algorithm. [5,14].

3.4 Determination of *Bayesian* Algorithm

Based on *training set*, there are some attribute that supposed influence GPA i.e. mathematics, physics, English, major and score attributes. Prior knowledge that descripts relation of attribute is not yet. The relation is node ordering and there is no dependency relation between attribute.

Based on analyzed of attribute, we get conclusion for three algorithms that we have learned; *Naive Bayesian* is the most suitable algorithm. K2 Algorithm needs node ordering; meanwhile PC algorithm will be more precise if there is relation between attributes.

4. Experiment and Evaluation

4.1 Scenario of Experiment

Generally, scenario of experiment is divided in two parts, i.e.:

a. Experiment to measure result quality
   Experiment was done in 10 GPA value limit i.e. : 2.5, 2.6, 2.7, 2.75,2.8, 2.9, 3.0, 3.1, 3.2 and 3.25 for 4 *testing set* size i.e. : 10%, 15%, 20% and 25%.
b. Experiment to measure process quality especially for response time.
   The scenario is we make fictive data with 2000,3000,4000,5000 and 6000 record size. Training set size is 10%, and GPA limit high =3.00.

4.2 Experiment and Result

We have done experiment and the result is:

Naive Bayesian	*C4.5*			GPA
	Raw Tree	Pre Pruned	Post Pruned	
78.55%	77.40%	78.77%	80.63%	2,50
74.32%	72.70%	74.24%	75.24%	2,60
69.06%	65.08%	66.90%	72.19%	2,70
67.15%	63.89%	65.48%	66.28%	2,75
65.62%	63.03%	64.12%	62.91%	2,80
59.80%	56.36%	58.53%	59.78%	2,90
59.27%	57.76%	57.78%	57.10%	3,00
61.03%	59.97%	61.21%	61.46%	3,10
65.69%	66.93%	67.39%	69.72%	3,20
68.67%	70.03%	71.69%	72.75%	3,25

4.3 Result Analyzed

The level of accuracy both of model will be compared with *difference two proportion statistics test.* [15]

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{pq\left(1/n_1 + 1/n_2\right)}}$$

Where

$\hat{p}_1, \hat{p}_2$: Estimator of proportion first, second method

n_1, n_2 : Total of testing-set size first, second method

p : Joint proportion (estimate by average both proportion)

q : $1 - p$

In [2.5 , 2.75) GPA interval, the result is *C4.5* algorithm has higher accuracy than *Naive Bayesian* method. In [2.75 , 3.00] GPA interval, the result is *Naive Bayesian* algorithm has higher accuracy than *C4.5* method. In (3.00 , 3.25], GPA interval, the result is *C4.5* algorithm has higher accuracy than *Naive Bayesian* method.

In C4.5 method, *post pruning* has higher accuracy level than *pre pruning* or raw *tree* model.

All the result above is done for special conditional GPA mean 2,88 and standard deviation = 0,58. If GPA value increases ‘a’ and standard deviation fixed, then conclusion of GPA interval will increase ‘a’ too.

Observation of data dispersion for several conditions can't explain relation data dispersion and accuracy level. Data dispersion is observed in several conditions:

- Dispersion of all attributes to T GPA
- Dispersion of attribute to score and T GPA
- Dispersion of attribute based on *tree* for 2.5, 2.75, 3.0 and 3.25 GPA limit.
- Dispersion of all attributes based on *tree* for 2.5, 2.75, 3.0 and 3.25 GPA limit.

Observation at data dispersion of mathematics, physics and English to score and major attribute show that data dispersion of mathematics, physics and English is not flat to score and major attribute, it means

- There is dependency relation between attributes.
- Assumption conditional dependency in *Naïve Bayesian* is not filled.

4.4 Analyzed of Process Quality

Process time of *C4.5* is much better than process time of *Naive Bayesian*, especially for large data. It is happen because *C4.5's* model is simplifier than *Naive Bayesian*'s model. But, process qualities both are good enough related with data.

5. Conclusion and Suggestion

5.1 Conclusion

1. In addition to get more efficient model, for most of cases *pruning* method can make increase accuracy of *C4.5*'s model prediction.
2. In *C4.5, post pruning* method has accuracy level better than accuracy level of *pre pruning*.
3. We can use system by combination to get best prediction result

Method	GPA interval
C4.5	[2.5 , 2.75)
Naïve Bayesian	[2.75 , 3.00]
C4.5	(3.00 , 3.25]

This conclusion was taken for GPA mean = 2,88 and standard deviation = 0,58. . If GPA value increases 'a' and standard deviation fixed, then conclusion of GPA interval will increase 'a' too.

4. For case of GPA prediction, there is no absolute superior method between *Naive Bayesian* and *C4.5* methods.
5. Process time of *C4.5* is much better than process time of *Naive Bayesian*, especially for large data. But, process qualities both are good enough related with data.
6. Generally, the result of GPA prediction with *Naive Bayesian* and *C4.5* methods is not satisfactory yet. Observation at data dispersion data can't explain pattern or level *C4.5'* accuracy. In *Naive Bayesian*, the accuracy level is influenced by dependency relation between some attribute so assumption of conditional independency is not filled.

5.2 Conclusion

1. There is a big *range* in data of 'high school grades' (mathematics, physics, English). It indicates discretization process is not good. Some suggestions are we can use standard 'high school grades ' i.e.:

   a. We can use 'high school grades' which have same standard like mathematics, physics, English values which are taken from national test.
   b. We can transform mathematics, physics, English values for each Senior High School to make overall range of 'high school grades' lower.
   c. It is necessary to handle noisy data before we process data.

2. We can use analysis dependency method to replace *Naive Bayesian* method because the experiment shows there are dependency relations between attributes.
3. We should can arrange amount of attribute and amount of attribute's discretization dynamically. Probably, It will make us will get model which have higher accuracy level.

Reference

[1] Han, Jiawei., Kamber, Micheline.(2000), *Data Mining Concepts and Techniques*, Morgan Kaufman Publishers

[2] Tan, Pang-Nin., Steinbach, Michael., Kumar, Vipin. (2006), *introduction to Data Mining,* Person Education,Inc

[3] Witten, Ian H., Frank, Eibe.(2005), *Data Mining, Practical Machine Learning Tool and Techniques, second edition,* Elsevier Inc.

[4] Neapolitan, Richard E.(2004), *Learning Bayesian Networks* ,Prentice Hall.

[5] Duda, R .O, Hart, P.E, Stork D.G(2000) ,*Pattern Classification*, John Wiley&Sons 2nd ed

[6] Quinland, Ross J.(1992) *C4.5 Program for Machine Learning*, Morgan Kaufmann Publishers

[7] Woolf, R.J.(2005) *Data Mining using Matlab.* Faculty of Engineering & Surveying University of Southern Queensland

[8] Pressman, Roger S.(2005), *SOFTWARE ENGINEERING, A Practitioner's Approach* **sixth edition**, Mc Graw Hill

[9] Piatetsky, Gregory(2006), *Data Mining and Knowledge Discovery in Business Databases*(PPT).www.kdnuggets.com/data_mining_course

[10] Santosa, Budi (2007), *Data Mining,* Graha Ilmu

[11] MDReid(2008), *Kullback-Leibler divergence*, en.wikipedia.org

[12] Kaufmann,Andreas(2008), *Entropy*, en.wikipedia.org

[13] Hilliard,Kirk (2008), *Gini coefficient*, en.wikipedia.org

[14] Gurak, Hamilton E., Olive, Findlater(2008), *Decision tree*, http://dms.irb.hr/tutorial/tut_dtrees.php

[15] Walpole, Ronald E.(1997) *Pengantar Statistika edisi ke-3,* PT Gramedia Pustaka Utama

3D RECONSTRUCTION EXTRACTION AND EQUATION OF CALIBRATION PARAMETERS USING SIMPLE SCKEWED CHESSBOARD PATTERN ON STEREO VISION

Agung Toto Wibowo

Institut Teknologi Telkom
atw@ittelkom.ac.id

Abstract

The existence of physical objects represented in the form of three-dimensional. Three-dimensional object projected by camera into a two-dimensional image. This image usage felt less, so it is required projection into a three-dimensional model (3D). The projection process of real world object into 3D models called the 3D reconstruction. The tool called 3D scanners. The process of image projection eliminate depth information (z coordinates of the coordinate system x, y, z). This information needed on the 3D reconstruction process. This information restored with a variety of approaches, one of them is stereo vision. Stereo vision is the reconstruction using two images from two cameras simultaneously. In stereo vision, an important process that must be done is calibration. Calibration process used to find the camera parameters (intrinsic, and extrinsic parameters). Current calibration have some weaknesses that is hard to get calibration parameters, and uncertain to get the parameter itself. In the proposed calibration studies we used 6x8 skewed chessboard patterns. From this chessboard, we obtained parameter and it will develop into a 3D equation for 3D reconstruction process. This proposed research produced the smallest error of 0.678 using RMSE.

Key words : 3D reconstruction, 3D scanner, stereo vision, Skewed chessboard pattern

1. Introduction

The existence of physical objects in the real world represented in the form of three-dimensional. Three-dimensional object projected by camera into a two-dimensional (image). This image usage felt less, so it is required projection into a three-dimensional model (3D). The processes of projection of the real world into 3D models called 3D reconstruction. The projecting tool called 3D scanners.

The process of image projection eliminate depth information (z coordinates of the coordinate system x, y, z). This information needed on the 3D reconstruction process. This information restored with a variety of approaches, one of them is stereo vision. Stereo vision is the reconstruction using two images from two cameras simultaneously.

3D reconstruction in stereo vision performed with several stages: stereo capture, calibration, feature segmentation, feature correspondence, reconstruction of 3D coordinates, 3D object creation. Stereo capture is the process of image retrieval using the camera simultaneously. This process is done using OpenCV library.

In stereo vision, an important process that must be done is calibration. Calibration process used to find the camera parameters (intrinsic, and extrinsic parameters). Current calibration have some weaknesses that is hard to get calibration parameters, and uncertain to get the parameter itself.

In the proposed calibration studies we used 6x8 skewed chessboard patterns. From this chessboard, we obtained parameter and it will develop into a 3D equation for 3D reconstruction process. This proposed research produced the smallest error of 0.678 using RMSE.

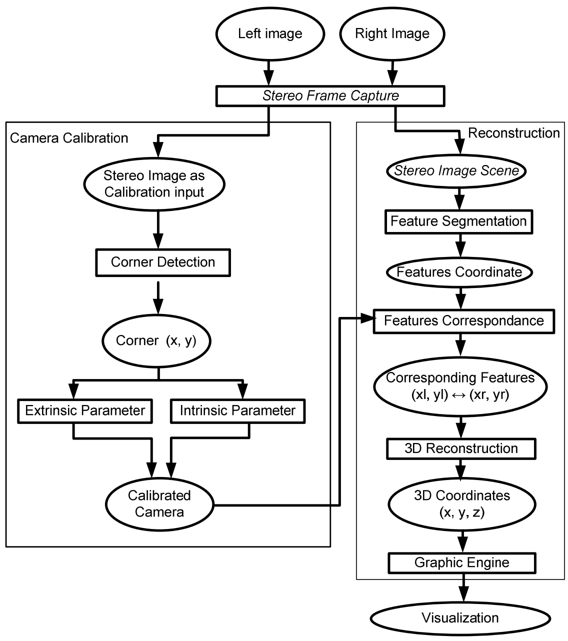


Figure 1. Stereo 3D Reconstruction Process[1][2]

2. Stereo Vision Process

Some of the process will run for 3-dimensional reconstruction with stereo vision. Those processes are [2]:

1. stereo frame capture, to get two frames simultaneously from left and right camera as input system;

2. stereo calibration, is process to obtain the values of intrinsic and extrinsic parameters from both camera;
3. feature correspondence, is process to get the correspondent point on left and right image;
4. 3D reconstruction, is process to restore the depth information (z coordinates) from the disparity information.

Those processes can be described by following block diagram [1][2].

3. Calibration

Real world object captured by camera through the projection process. On stereo vision, both cameras had two values of projection parameters. This value was need on 3D reconstruction process. Those parameters are intrinsic parameters and extrinsic parameters. These parameters extracted from calibration process.

Intrinsic parameter calibration on the camera used to obtain a transformation as the focal length value, and the central projection. Intrinsic parameters provide projections of 3D points in the real world into 2D coordinates of the camera system. Extrinsic parameters provide a transformation from camera coordinate to real world coordinate.

3.1 Intrinsic and Extrinsic Parameters

Key points on the camera model is the focal length (f), *pixel scaling* (s_x, s_y), optical center (o_x, o_y) and skewing factor(γ). Pixel scaling value provides a measure of the width and height of one pixel, and used to calculate the aspect ratio of camera. Skewing factor (γ), defines the angle between the pixel with x and y-axis and allows the camera to be incorrect pixel 900. All values are summarized in the camera parameters.

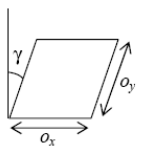


Figure 2. Skewing factor on camera projection[1]

The following values show the intrinsic parameters of camera [1][2].

$$M_{int} = \begin{bmatrix} fs_x & -fs_x \cot\gamma & o_x \\ 0 & \frac{fs_y}{\sin\gamma} & o_y \\ 0 & 0 & 1 \end{bmatrix} \tag{1}$$

If the camera pixel can be assumed to elbow, then the skew factor is ½ π, than intrinsic parameter can define as :

$$M_{int} = \begin{bmatrix} fs_x & 0 & o_x \\ 0 & fs_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

Extrinsic parameters define the influence of the external characteristics of the camera. Extrinsic parameters usually consist of the translation vector and rotation matrix associated with the coordinate frame in the real world coordinates. Extrinsic parameters defined as [1][2]:

$$M_{ext} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \tag{3}$$

3.2 Camera Perspective

Projection process between pixel camera and a point in the real world is known as perspective projection. Linear matrix equation in the camera's perspective projection defined as[1][2]:

$$\begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = M_{int} M_{ext} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix} \tag{4}$$

wrehe M_{int} and M_{ext} represents the intrinsic and extrinsic parameters. Intrinsic matrix is a 3x3 matrix that contains information intrinsic camera. While extrinsic consist 3x4 matrix that stores information the camera rotation and translation. The i subscript is used to denote camera coordinate, w subscript define real world coordinate.

This relation can also written as,

$$P_c = RP_w + T \tag{5}$$

$$P_i = M_{int} P_c \tag{6}$$

where P_i and P_w are *image* coordinate and real world coordinate, while P_c represent point on camera coordinate, R and T is rotation matrix and translation matrix.

4. Common Stereo Vision

4.1 Rectify Image

Rectify image is one part of the calibration process. Results of calibration are the intrinsic and extrinsic matrix of both cameras. Those matrixes then used to calibrate the camera so that the obtained calibrated image. Complete equation of the camera calibration can be found in "*Learning OpenCV, Computer Vision with the OpenCV Library*", and "*Real-time 3D Reconstruction from Stereo*".

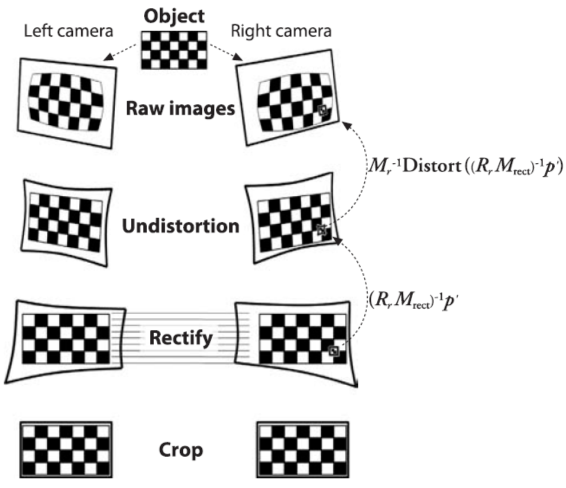


Figure 3. Image rectification process to get calibrated image[1]

The next process is triangulation. Triangulation process is use to reconstruct 3D, using the results of segmentation and feature correspondence. Complete reconstruction equation can be found also in the reference.

4.2 Image Rectification Problem

In stereo vision system, rectification process does not always success. From several experiment using the same algorithm (in the sample OpenCV), we obtained less satisfactory results. Rectification result are sometimes skewed too extreme, so only a few areas that can be used. The following image is example of rectify image using a web camera.

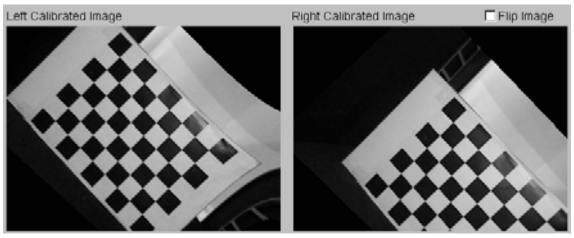


Figure 4. Example of image rectification problem

5. Proposed Calibration

Calibration process obtains some parameters needed to restore the value of z coordinates of objects (information depth). In research that was build, used 8x6 size chessboard printed on A4 paper. The size of each box on the board is 2.9 cm. Chessboard pattern used found in the OpenCV project.

Chessboard calibration placed with a slope angle of θ and captured from both cameras. The value of θ in this research is 40 °, 50 °, and 60 °.

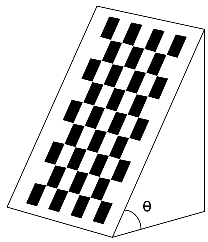


Figure 5. Proposed calibration board

In the calibration process, two cameras captured the chessboard. Both left and right image is then searched features a chessboard-point using *cvFindChessboardCorners* function in OpenCV library.

These points then used to find the calibration parameters. These parameters are $x\,unit,\ y\,unit,\ z\,unit,\ ratio\ x,\ ratio\ y,\ ratio\ z,\ center\ x,$ dan $center\ y$.

$x\,unit$ value represent distance between two cross point in real world. On the printed calibration board, the value $x\,unit$ is 2,9 cm. While $y\,unit$ and $z\,unit$ represent height unit and depth unit of skewed chessboard. All $x\,unit$, $y\,unit$ and $z\,unit$ value counted using the following equation.

$$x\,unit = 2{,}9 \quad (6)$$
$$y\,unit = x\,unit\ \times sin(\theta) \quad (7)$$
$$z\,unit = x\,unit\ \times \cos(\theta) \quad (8)$$

The value of $ratio\ x$ parameter represent ratio between $x\,unit$ with number of horizontal pixel in captured image. To get the value of $ratio\ x$ value, first time we calculated the average projection of cross point in horizontal direction. This value is called $x\ avg$, which is calculated by equation (9):

$$x\ avg = \frac{1}{(nkol-1)\times nbar} \sum_{i=1}^{nbar} \sum_{j=1}^{nkol-1} X_{i,j+1} - X_{i,j} \quad (9)$$

where $nkol$ value represent amount of chess column (6 used in the research), and $nbar$ is the number of rows of chessboard (8 used in the research). The value of $X_{i,j}$ indicates x location of cross point in the calibration image.

The value of $ratio\ x$ then calculated by the following equation :

$$ratio\ x = \frac{x\,unit}{x\ avg} \quad (10)$$

The value of $ratio\ y$ parameter represent ratio between $y\,unit$ with number of vertical pixel in captured image. To get the $ratio\ y$ value, first time we calculated the average projection of cross points in vertical direction. This value is call $y\ average$, which is calculated by following equation :

$$y\ average = \frac{1}{nkol \times (nbar-1)} \sum_{i=1}^{nbar-1} \sum_{j=1}^{nkol} Y_{i+1,j} - Y_{i,j} \quad (11)$$

where $Y_{i,j}$ indicate y location of cross point in calibration image.

The value of $ratio\ y$ then calculated by the following equation :

$$ratio\ y = \frac{y\ unit}{y\ average} \quad (12)$$

Parameter $ratio\ z$ represent ratio between disparity z projection (the difference x left with x right) with $z\ unit$ value. Same as the previous search value ($ratio\ x$, and $ratio\ y$), $ratio\ z$ value is calculated by finding the $z\ average$ score first. The following equation is used to find $z\ average$ and the $ratio\ z$.

$$z\ avg^1 = \frac{1}{nkol \times (nbar-1)} \sum_{i=1}^{nbar-1} \sum_{j=1}^{nkol} (X_{i+1,j}^{left} - X_{i+1,j}^{right}) \quad (13)$$

$$z\ avg^2 = \frac{1}{nbar \times (nbar-1)} \sum_{i=1}^{nbar-1} \sum_{j=1}^{nkol} (X_{i,j}^{left} - X_{i,j}^{right}) \quad (14)$$

$$z\ average = z\ avg^1 - z\ avg^2 \quad (15)$$

$$ratio\ z = \frac{z\ unit}{z\ average} \quad (16)$$

Center x and center y represent value used as the reference coordinate system. The reference coordinates using the smallest value of image coordinates. The value of center x and center y parameters obtained from the equation below:

$$center\ x = \min\left(\bigcup_{i=1}^{nbar} \bigcup_{j=1}^{nkol} X_{i,j}\right) \quad (17)$$

$$center\ y = \min\left(\bigcup_{i=1}^{nbar} \bigcup_{j=1}^{nkol} Y_{i,j}\right) \quad (18)$$

The fifth parameter values above formula will be use in triangulation to restore the coordinate values of feature points at the time of reconstruction of the object.

6. Feature Correspondence

Correspondence conducted by comparing the point in left image and right image at the same index. Difference of height on first cross point represent delta value of Δy.

Δy values obtained from the formula below.

$$\Delta y = center\ y_r - center\ y_l \quad (19)$$

$$y_l = y_r - \Delta y \quad (20)$$

From the corresponding features, will be use to find the disparity value. Disparity value is a value that describes the difference between the values of x in the field of projection. Disparity value obtained by the following equations,

$$disparity = abs((x_r - center\ x_r) - (x_l - center\ x_l)) \quad (21)$$

$disparity$ value will be used to reconstruct the 3D coordinates of the model.

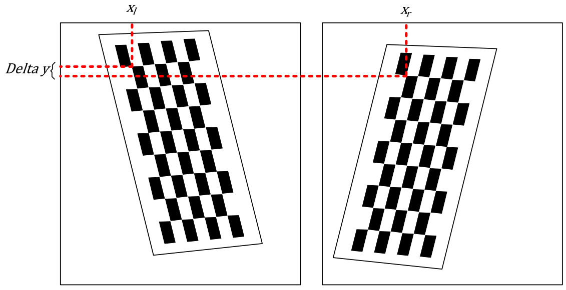


Figure 6. Illustration Δy on Chessboards Image

7. 3D Coordinate Reconstruction

Once known set of points corresponding to, then we will reconstruct the 3D model coordinates (x_m, y_m, z_m). To reconstruct the model coordinates, used the following equation.

$$x_m = (x_l - center\ x_l) \times ratio\ x \quad (22)$$

$$y_m = (y_l - center\ y_l) \times ratio\ y \quad (23)$$

$$z_m = disparity \times ratio\ z \quad (24)$$

To build a 3D object done by connecting the point clouds. One corresponding point become 3D point, we maintained all point in array of point clouds. Point clouds are a collection of 3D points that can be use to build 3D models. The results of the 3D chessboard pattern of development shown in the picture below.

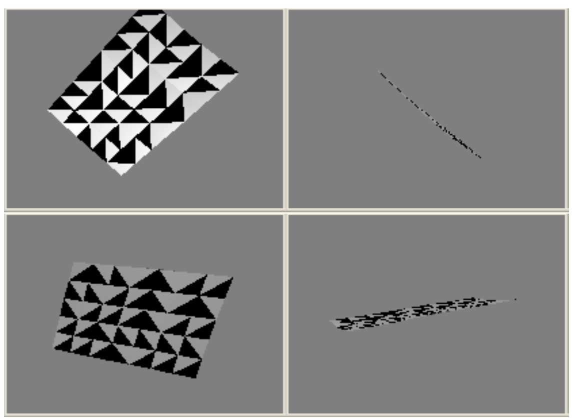


Figure 7. 3D Reconstruction Result of Chessboard Pattern

8. Experiment Result

8.1 Scenario

Proposed calibration testing conducted with the aim to know the best calibration parameters of the proposed system. Parameters that can be change are the degree of calibration board, and the distance between cameras.

At the skew of the board calibration parameters, the parameters used angle 40°, 50°, and 60°. Camera distance parameter testing will be done with variations of value between 7.5 cm, 10 cm, 12.5 cm, 15 cm, and 17.5 cm.

Each combination of these parameters will be test five times to count and compare the results of the angular coordinate reconstruction chessboard with ideal conditions. Comparison value obtained by calculating the 3D coordinates with the ideal conditions (the point of chess pieces) using the proposed reconstruction equation and expressed in the RMSE calculations.

8.2 Result

As mentioned in the previous chapter, the proposed calibration tests carried out with several variations of the calculated parameters in the RMSE. Here are the average resume RMSE results of an experiment in which the parameters of each test performed 5 times.

Table 1. Average RMSE on Experiment

Camera Distance (cm)	Average RMSE		
	40°	50°	60°
7.5 cm	6.106	4.561	2.304
10 cm	6.220	4.603	1.149
12.5 cm	6.873	4.825	0.792
15 cm	6.724	4.834	0.740
17.5 cm	6.769	4.749	0.815

In comparison, the best results for each of the experimental parameters can be seeing in the following table:

Table 2. Best RMSE on experiment

Camera Distance (cm)	Best RMSE		
	40°	50°	60°
7.5 cm	5.911	4.343	2.028
10 cm	5.709	4.055	0.869
12.5 cm	6.429	4.618	0.713
15 cm	6.212	4.716	0.678
17.5 cm	6.200	4.349	0.693

8.3 Analysis of Results

From the above results, the calibration parameters obtained for good angle and distance between cameras is 60°, and 15 cm. This result is the best results during the research process.

When the calibration experiment conducted, the largest errors were generating from cvFindChessboardCorners function. This function was not good at detecting the calibration board with a sharp slope. Calibration slope as much as possible up to the axis of the camera can detect the corner point location. Sample error detection angle shown in the picture below.

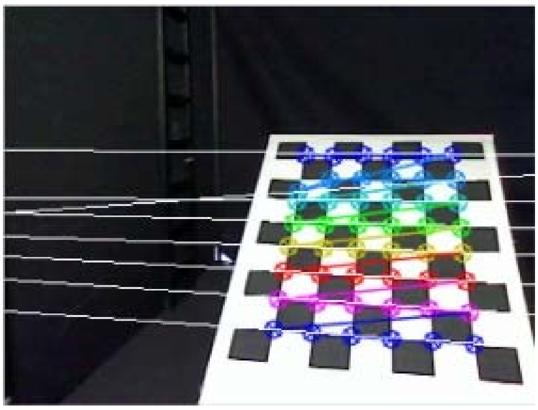


Figure 8. Error Detection in A Small Slope

Experiments were excluded slope angle 70 ° or more because of the calibration formula value of $z\,unit$ (8) involved cos(θ). This is a critical point in the proposed equation. Value of $z\,unit$ in the equation written $z\,unit = x\,unit \times \cos(\theta)$. Value of $z\,unit$ requires high accuracy in 3D reconstruction (remember what is important is to return z information). Value cos(70°) is 0.342 which is small enough as a range object reconstruction.

From the table above (at an angle of 60°) obtained a relationship of information that the further the distance the camera gives a value smaller RMSE. However, on another angle of 50° and 40°, the value of increasing the distance it gives the increase of the RMSE values. From these information drawn a conclusion that there are ideal conditions at the laying of the corner of calibration and a certain distance.

In addition to the above reasons, the influence of small angle (the italics) after the projected differences between segments of the difference x becomes larger. This happens because the results of the straight-line projection, the farther the narrower, so that if the accumulated calibration error will be larger.

Experiment discontinued at the value of the distance between the camera due to greater average RMSE experiencing a turning point at a distance of 15 cm between the camera and the camera distance between 17.5 cm average RMSE becomes larger increases.

9. Conclusion

From the test results and analysis conducted in the previous chapter, obtained some conclusions as follows:

1. Proposed calibration can be use as an alternative to 3D reconstruction.
2. From the model system was built, best calibration parameters obtained with optimum angle is 60°, and the distance between the camera as much as 15 cm.
3. In the experiment this parameter gives the smallest RMSE results with an average value 0,740, and the best RMSE is 0,678.

References

[1] Bradski, G., and Kaehler, A. (2008) *Learning OpenCV, Computer Vision with the OpenCV Library*, 1st ed., O'Reilly Media, California, United States.

[2] Wild, D. (2003), *Real-time 3D Reconstruction from Stereo*, The University of York, Heslington United Kingdom.

[3] Zhu, Z. (2005), *Stereo Vision*, Lecture 5, CSc80000 Section 2, http://www.cs.engr.ccny.cuny.edu/~zhu/GC-Spring2005/CSc80000-2-VisionCourse.html (accessed on 10.00WIB, 4 June 2009).

TRAINING SET WITH ENHANCEMENT TSVQ METHOD FOR MEDICAL IMAGE BASED SUPER-RESOLUTION

Ni Putu Agustini, Muhammad Ashar

National Institute of Malang

Abstract

TSVQ method considering patches as vectors. Vector search is actually the constitution by the low-resolution patches and the patches overlap at high resolution, and dimensions. Quantization regarding the use of factors, the smallest set of search vectors from the set of training vectors. Finding the nearest neighbor of an input vector search is easier and faster. The smallest set of codebook is composed of the codewords. Is calculated for Codebook vectors are closed as a set of training vectors. Codebook has a balanced tree structure (a balanced tree structure), in order to re-encoding (matching process) which only consists of wide reading tree, with decisions at every level to select the left and right leaf nodes. Codewords leaves have been measured to minimize distortions in the distance ready to be determined by training vectors. TSVQ training process with a visual quality better than with the addition of pixels and the interpolation method of image quality biasa.Hasil magnification increases with the PSNR measurement has an average PSNR of 75 - 80% more sharply.

Keyword : TVSQ, PSNR,Super Resolution,Patches

1. Background

Analysis approach has reached such bilinier interpolation or cubic B-spline, on the edge remove fled in detail. Increase in high frequency with less than perfect human beings as the operator in deciding the level of improvement. However, some statistical approach gives good results. Robust and fast to algorihtm no image sharpness, although many are trying to create and study like Kersten and later, Hulbert and paggio using a linear approach, but they still have shortages in some cases. Freeman further propagation using the Bayesian algorithm with a more efficient result, this algorithm uses training data in determining the propagation parameters. In the test try to use one-pass algorithm without using a Markov model previously used network. However, the training sets are constructed from Markov models Network or one pass algorithm is still slow and not efficient and requires further research to refine the process of training based ini.Penggunaan TSVQ selected as one of the methods that can be applied to the training set of images in accelerating and improve the image magnification quality.

2. Concept of Training set using TVSQ

Vector Quantization process are :

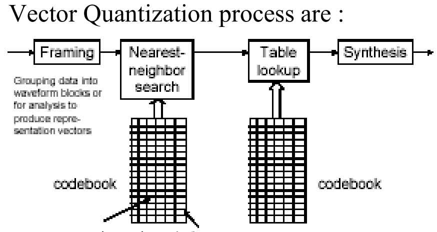


Figure 1. Single-Stage Vector Q

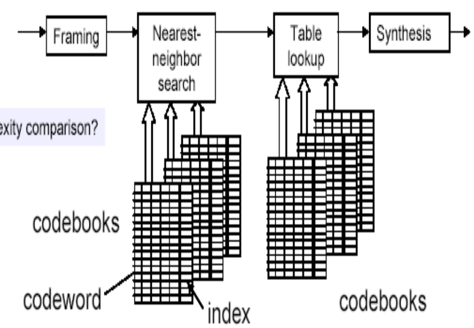


Figure 2. Multi-Stage Vector quantization

Lloyd's algorithm:

Step 1: Initialization codebook. C_m

Step 2: In the *codebook* C_m, for each *training vector* $x \in \Omega = \{x_j, j = 1,2,...,L\}$ *closest codeword* index $i = q(j)$; $i = q(j)$ = index of *the closest codeword* for x_j

Step 3: After the entire training set is stopped, Calculate the center of a new area for each cell of the set correctly; $\Omega_i = \{x_j ; q(j) = i\}, i \in I$

Step 4: Calculate the average distortion of new. If the drop because the amount is not significant during the last iteration, stop;
do repeated steps 2-4.

Codebook initialization:

a. Random

b. Binary split algorithms

1. Global start global expansion, set n = 1;

2. Create a new 2_n of the n codewords (centered)

the codewords by any of the influential (each multiplied by 0. 99 and 1.01) or find codewords in each cell of the most distant.

The area calculation :

Given deviation of the distribution $f_x(x)$

and distortion measure *d (x, y)*,

Limit : $Y = \arg\min \int_x d(x,y) f_x(x)dx$

Euclidean distance:

$$d(x,\hat{x}) = d_2^2 = (x-\hat{x})^t (x-\hat{x})$$

Empirical data: $\overline{x} = \frac{1}{L}\sum_{i=1}^{L} x_i$

With L1 differences and empirical data

$$d(x,\hat{x}) = | x - \hat{x} |$$

the equation is the median.

3. Training Set Model

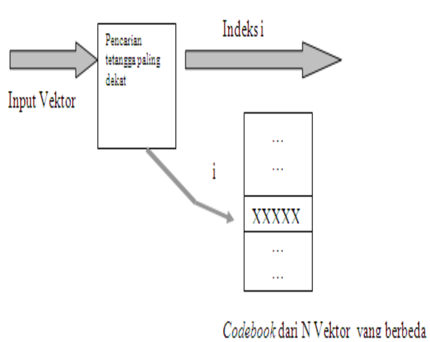


Figure 3. Training process can use the codeword:

3.1 Full Vector Quantization with Lloyd algorithm

Each training signal compared to all the codewords in each iteration.

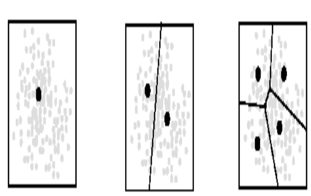


Figure 3. Full Vector Quantization with

3.2 Binary Tree Vector Quantization

Each training signal compared to the two codewords in the area of each during each iteration.

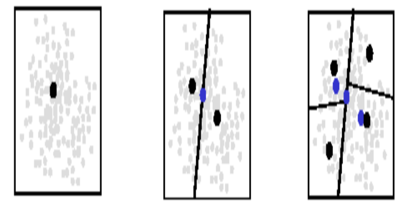


Figure 4. Binary Tree Vector Quantization

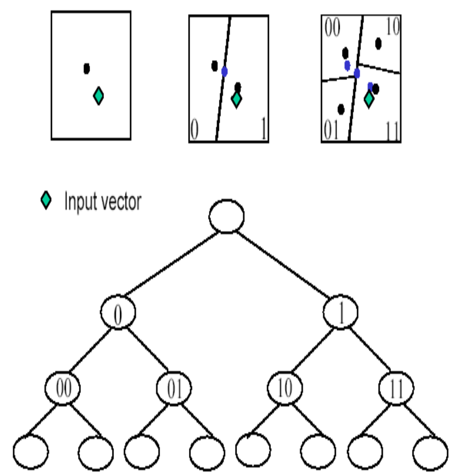


Figure 5. Binary Vector Quantization

Description Tree-Structured Vector Quantization:

a. Overall distortion aberration is sub optimal.
b. A lot faster training in relation to the division of training set (fewer signs and less codewords to compare)
c. Codebook space is greater in the training (nearly two times).
d. When combined with the representation of code and receiving a copy of the entire tree-structured codebook, may offer some additional (minor) useful in robust transmission and get back the signal (compared to codeword get back the best resolution, get the one that received intermediate code which allows reliable in the case of errors or lost bits). Before discussing the application of TSVQ method used in training set, should discuss in advance of the patches are suitable (matching the patches). This matching process using . Search can be writtenαEuclidian normalization, as modified by the parameter as a vector in training set. The equation is:

$$V=(V_L,V_n).$$

VL consists of a vector with a low-resolution data and Vh is the vector with a high-resolution data (overlap) and then obtained:

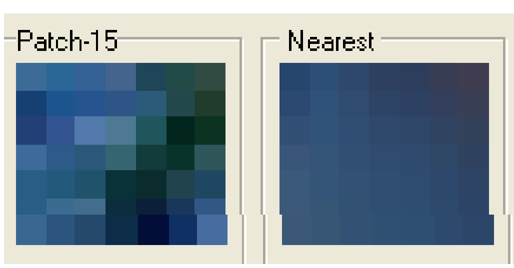
$$d(V,W)^2=|VL-WL|^2+\alpha|Vh-Wh|^2$$

TSVQ method considering patches as constitution by the low-resolution patches and the patches overlap at high resolution, and dimensions. Quantization regarding the use of factors, the smallest set of search vectors of the training vector set. Finding the nearest neighbor of an input vector search is easier and faster. The smallest set of codebook is composed of the codewords. Is calculated for Codebook vectors are closed as a training vector set. Codebook has a balanced tree structure (a balanced tree structure), in order to re-encoding (matching process) which only consists of wide reading tree, with decisions at every level to select the left and right leaf nodes.

4. Results

Examples of simulation results 16 The Best Matching Patches sajadah.bmp image

Figure 6. Image x ray original hand medical image with size 100 x 149 px



Picture 7. Figure Patch hand medical image with the nearest partner

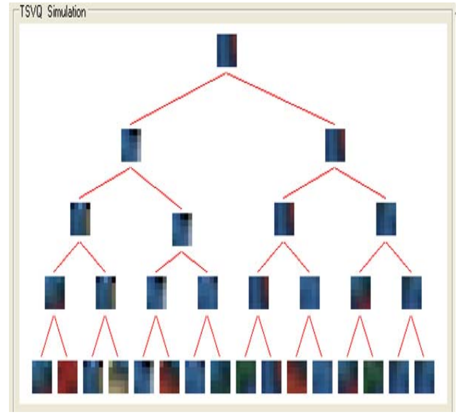
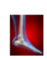
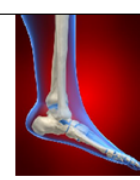


Figure 8. The results simulation 16 The Best Matching Patches.

Table1. Magnification Medical image

File Name	Original Image	Output Image *Super Resolution*	
		Factor 2	Factor 4
Hand			
	100 x 149 px	200 x 298 px	400 x 596 px
Thorax			
	64 x 63	128 x 126	256 x 252
Backbone			
	125 x 84	250 x 168	500 x 336
Foot			
	76 x 75	152 x 150	600 x 592

Enlargement Image Analysis

At the time the image is enlarged factor 2 give PSNR better thanfactor 4 magnified because the pixels are formed closer together, for more details see table 2 .The table shown in the magnification hand.bmp with training sets to produce 100 images with PSNR = 35.83 (factor 2) magnification and PSNR =35.78(factor 4) magnification.

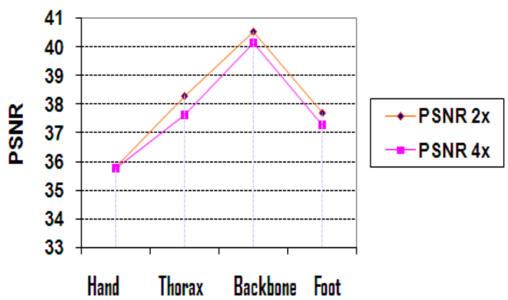


Figure 8.Comparison of PSNR factor 2 and 4

Image enlarged factor 2 with MSE better than factor 4 magnified because of the missing pixels less. For MSE = 16.94 for 2x magnification and PSNR = 17.14 for 4x magnification.

For the time the image is enlarged 2x faster than 4x magnified because the pixels are processed less, for more details see table 2. The table shown in the magnification Sajadah.bmp with training sets to produce 100 images time = 3.30 minutes for 2x magnification and time = 14.50 minutes for 4x magnification.

Table 2. Calculation of MSE for training image variation.

File Name	Number of Image Training	MSE (Factor 2)	Time (m)	MSE (Factor 4)	Time (m)
Hand	100	16,94	3,30	17,14	14,50
Thorax	100	9,58	0,55	11,17	3,55
Backbone	100	5,71	2,30	6,28	10,15
Foot	100	10,98	1,20	12,101	5,35

5. Conclusion

1. TSVQ training process with a visual quality better than with the addition of pixels and the interpolation method of image quality biasa.Hasil magnification with PSNR measurements have an average PSNR of 75 - 80%
2. Visual quality of the image with TSVQ has an average MSE of 12%

References

[1] S.Baker,Limits on Super Resolution and How to Break Them, Proc.IEEE conf.Computer Vision and Pattern recognition, (CVPR) vol.II.IEEE CS press, Los Alamitos,Calif, 2000, pp 3722-3729

[2] Xin Li, New edge directed interpolation, IEEE Trans, Image Processing, 2001.

[3] Freeman ”Exampled Based Super Resolution”, IEEE Computer Graphics and application SN 0272-1716/02

[4] Régis Destobbeleire *Super-resolution,* L. Velho (IMPA, Brazil) and S. Mallat (École polytechnique, France) May - June 2002

[5] A. Herzmann, C. E. Jacobs, N. Olivers, B. Curless, and D. H. Salesin, Image analogies. In ACM SISGGRAPH, 2001. In Computer Graphics Proceedings, Annual Conferences Series.

BUILDING AN INDONESIAN DIGITAL FORENSIC LABORATORY

Charles Lim, Msc. ECSA, ECSP, CEH, CEI[1]

[1]Department of Information and Communication Technology, Faculty of Information Technology,
Swiss German University
[1]**charles.lim@sgu.ac.id**

Abstract

With the rapid rise of Internet usage in Indonesia, the number of cyber crime is also rapidly increasing in Indonesia. The protection of critical national infrastructure becomes national top priority and IDSIRTII (Indonesia Security Response Team on Internet Infrastructure) was formed to control, to prevent and to help the law enforcement to handle the cyber attacks on this infrastructure. In 2008, Indonesian government made a decision to build the first nationwide digital forensic laboratory. Swiss German University, an official partner university of IDSIRTII has been during research and development, in the area of digital forensic and malware analysis. This paper will discuss the proposal of building digital forensic laboratory and necessary professional personnel to support this effort.

Keywords: digital forensics, cybercrime, laboratory

1. Introduction

With more and more Indonesian people using Internet, currently with 30 million active Internet users in Indonesia per day, 10 Gbps total national internet traffic and 1 million attacks daily [1], the risk of cyber attacks to all aspect of live including social, politics, and national security is also on the rise. With the cyber attack incident in 2004 on General Elections Commission web site [2], the government of Indonesia is finally aware of the coming national threats to the vital Indonesian assets including National Financial System, National Electricity, Defense System, National Air Transportation System and many more.

Finally, in year 2007, ID-SIRTII (Indonesia Security Response Team on Internet Infrastructure) was established [3] to fulfill the national interest to monitor, control and prevent cyber crime. The institution also provide critical information service to the public, and include technical consultation to the government as well to the public, so that the relevant institution will be able to anticipate and mitigate any risks of exposing the confidential company assets to the public. Hence, research development into how to perform incident handling and forensic investigation becomes critical to the success of overall ID-SIRTII being. Swiss German University becomes one of the first institution that sign an MOU (Memorandum of Understanding) to collaborate with ID-SIRTII to perform research and development in the area digital forensics and malware analysis. To facilitate research and development, government through ID-SIRTII will invest in several critical laboratories, which include Digital Forensic Laboratory, Malware Laboratory and Honey Net Project Laboratory. In this paper, we will discuss the survey of Digital Forensic Laboratory requirement to support this lab and the required professional certification to maintain the Laboratory.

2. Digital Forensic

Digital Forensics is a discipline that combines elements of law and digital science to collect and analyze data from digital systems, networks, wireless communications, and storage devices in a way that is admissible as evidence in a court of law [4]. Previously, highly skilled and trained personnel perform digital forensics and digital forensics was considered more of an art than of science. This is no surprise, since the field of digital forensics was born largely due to the demand for service to law enforcement community [5]. Today, however, the shift toward the balance of art and science is due to the facts of commonly accepted procedures and more widely accepted digital forensics software [4]. To ensure the digital forensics result report is accepted as a legal proof in the court, digital forensics must abide to certain principles and rules. Following are the 4 principles of Digital Forensics [5]:

1. No action taken by law enforcement agencies or their agents should change data held on a digital device or storage media which may subsequently be relied upon in court.
2. In circumstances where a person finds it necessary to access original data held on a digital device or on storage media, that person must be competent to do so and be able to give evidence explaining the relevance and the implications of their actions.
3. An audit trail or other record of all processes applied to digital device-based electronic evidence should be created and preserved. An independent third party should be able to

examine those processes and achieve the same result.

4. The person in charge of the investigation (the case officer) has overall responsibility for ensuring that the law and these principles are adhered to.

More detail rules of Digital Forensics are explained in [6], which at the core, it emphasizes that digital evidence must be handled with extra care to be qualified as a proof in the court of law. Hence, every actions performed during the investigation of the digital evidence must be properly logged and recorded. This also includes proper handling of the equipments seized, and device or equipment labeling.

Digital Forensics process goes through several phases, which include the following [5]:

1. Evidence Collection
2. Evidence Preservation
3. Evidence Analysis
4. Evidence Presentation

The first phase is commonly performed at the “crime scene,” where items include computers, device and some form of digital media for storage. On the second phase, the evidence must be preserved in a way that is reliable, complete, accurate, and verifiable. On the Evidence Analysis phase, the analysis must be admissible, authentic, complete, reliable, and believable. On the final phase, the evidence must be presented in a way that is understandable to the intended audience and must be supported with proper documentation. During this process, the digital forensic investigator must maintain, so called, chain of custody, which ensures the identity and integrity of the articles from the time collected through the time the results of the analysis reported and subsequently disposed of. Detail of step-by-step procedure preparing each person who is doing the investigation, forensic readiness, can be found in [7].

3. IDSIRTII digital forensics lab requirements

As stated in the official letter released by Indonesia Ministry of Information and Communication that IDSIRTII has seven main roles to play in protecting national critical infrastructure [8], which include:

1. Socialize the enhancement of network security based on internet protocol to all parties involved.
2. Monitoring, early detecting, preventing threats and other intrusions on the internet network in Indonesia.
3. Establish and/or provide, preserve, operate, and develop database system to:
    a. Support the two elements above.
    b. Store the log files, and
    c. Sustain the law enforcement.
4. Give the information services and factual report of internet threats and intrusions to people.
5. Provide a simulation laboratory, workshop and training to enhance the security.
6. Provide the consultation and technical support.
7. Become the contact point, either to related domestic institutions or foreign institutions.

To fulfill point 5 and 7, existence of IDSIRTII Digital Forensics Laboratory is mandatory and the development of such lab is expected to be completed in 2010.

When established the lab will offer services to various institutions, which include law enforcement agencies, government agencies, and public. The services offered are based on the Digital Forensics methodology mentioned above, and the services include:

1. Preservation and authentication of electronic evidence
2. Analysis of computer hard drives and any other electronic media
3. Recovery of deleted files, partitions, and formatted drives
4. Expert witness testimony
5. Internet and network investigations
6. Audio/Video Forensics

Based on the requirement analysis, the laboratory should consist of at least 5 core elements: Modules and Systems, Hardware, Software, Facility and People, as shown in figure 1 below.

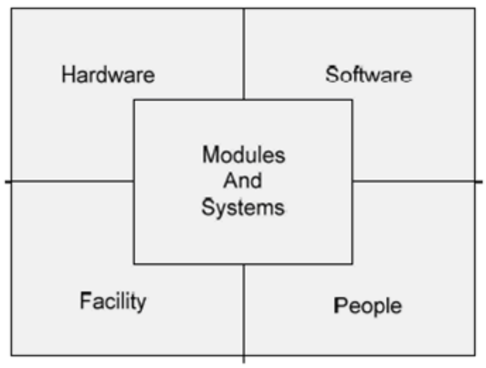


Figure 1. Category of laboratory core elements

To find out detail requirements of the laboratory, IDSIRTII has done benchmarking of various countries’ digital forensics laboratory (NSA, JPCERT, KrCERT, MyCERT, AusCERT, SingCERT, etc.), and the detail requirements are categorized below:

1. Digital Forensics Modules , Software, and Systems
2. Data Acquisition and Imaging Portable Devices
3. Data Sanitizer and Eraser Drives Portable Devices
4. Data Recovery and Encryption Portable Devices

5. Physical Hard Drive Protection Portable Devices
6. Media Duplicator and Media Imaging Portable Devices
7. Mobile Phone Forensics Portable Devices
8. Notebook (Mobile Computer)
9. Forensics Black Bag
10. Forensics Crime Scene Equipment and Tools
11. Anti Static Room Development
12. Evidence Storage Room Development
13. Security Systems
14. Network Systems
15. Training, Workshop, and Certification
16. Laboratory Accreditation and Certification
17. Personnel

For point 1, Two industry standard software for performing forensics are listed, and they include EnCase and AccessData Forensics Toolkit (FTK). For point 2 to point 8, the specifications are very much standardized, except with some yearly update provided by the vendor to support recent hardware technology. For point 9, the tool is enclosed in a "black bag", which includes various toolkit data acquisition devices [9], such as hard drive adapters, adapter converters, etc. Forensics Crime Scene Equipment and Tools are mandatory equipment for field investigator when working on the crime scene and they include such as sterile gloves, flashlight, screwdrivers, bags, tags/labels, communication devices, tapes, camera, mirrors, pens, notebooks (for writing), containers, devices for testing wireless or Bluetooth, and many more. Next, Anti static room development ensures that the laboratory room is designed away from large electrical conduits or magnetic fields since magnetic fields can destroy evidence. Anti static room must include the following: Anti static shield and vacuum room, Anti static tools and equipment, and Anti static personal coat and equipment. Anti static mat is highly recommended for some laboratory room with low in humidity and high elevation location. Storage for evidence is critical to the whole process of digital forensics process and the storage room must meet the following conditions:

1. The facility should use environmental controls and have an ambient temperature of between 15 degree Celsius and 20 degree Celsius. This should be refrigerated air-conditioning, not evaporative, and have 0% of humidity.
2. It should have a dust-free environment.
3. It should be well away from large electrical conduits or magnetic fields
4. Fire control should be by oxygen deprivation, not a sprinkler system.
5. It should not be located near any source of vibration (on the edge of the building near a busy road).
6. The facility should be not near sources of direct ultraviolet light (e.g. the sun). This is because ultraviolet light can rapidly degrade some optical and magnetic media.
7. The number of people authorized to open and/or enter the evidence storage room should be kept to a minimum. Maintain records on who is authorized to access the digital storage room.

In addition, the evidence storage area must be secured from unauthorized entry or tampering. A routine maintenance system for the contents of the evidence must be established. For security systems, standard CCTV system that covers the whole area of the laboratory should be installed. To access the laboratory, personal identification tags must be worn and biometric access control is mandatory for all restricted area. For network systems, the laboratory should be equipped with high speed connection for intranet and internet connection and the systems must be designed with fault tolerance in mind.

To ensure qualified personnel, all the personnel that work in the laboratory, must have the following professional certification:

- Computer Crime Investigator (CCI - EC-Council)
- Computer Hacking Forensic Investigation (CHFI – EC-Council)
- GIAC Certified Forensics Analyst (GCFA)

When completed, the laboratory shall be accredited by the following:

- The American Society of Crime Laboratory Directors/Laboratory Accreditation Board (ASCLD/LAB)
- ISO/IEC 17025:2005

Finally, to run the laboratory, the following personnel are required:

- Laboratory Manager, which is responsible for financial planning, mentoring all staffs, controlling, managing, counseling, and training. The laboratory manager has the highest authority-level which means the laboratory manager can enter all areas inside the laboratory. Therefore, a laboratory manager must be the trusted person.
- Reception Officer, which is responsible for maintaining the guestbook, listing all people who entered the laboratory.
- Digital Forensics Investigators and Response Team (as known as Case Investigator), which are responsible for dealing with clients, finding as much information, investigating the clients, deciding whether the tasks are accepted into our laboratory or not, allocating the priority in which cases are to be dealt with. In case of crime, they will be the first person on site, who will perform the incident response procedures and document the crime scene. Mandatory certifications: CHFI Certification or CCI Certification. Optional certification: GCFA Certification.
- Digital Forensics Examiner and Analyst (as known as Lab Specialist) which are responsible

for creating a replica of the original evidence, analyzing the case and the evidence. Mandatory certifications: ACE Certification and EnCase Certification for whom deal with EnCase and FTK software, GCFA Certification. Optional certification: CCI Certification (digital forensics) and CHFI Certification (internet forensics).

- Digital Evidence Preserver must be special trusted personnel who work with the original evidence and preserve it from any damage. These particular people will have the high authority to enter the evidence

4. Laboratory Design

Since the scope of work of IDSIRTII is the entire country of Indonesia and its requirement to collaborate with other CERT/CISRT in the Indonesia, the size of the laboratory ought to be large enough to cover the necessary requirements. It is recommended that the size of the laboratory at least the size of laboratory as suggested by [12]. The general layout of the General Computer Layout is shown below.

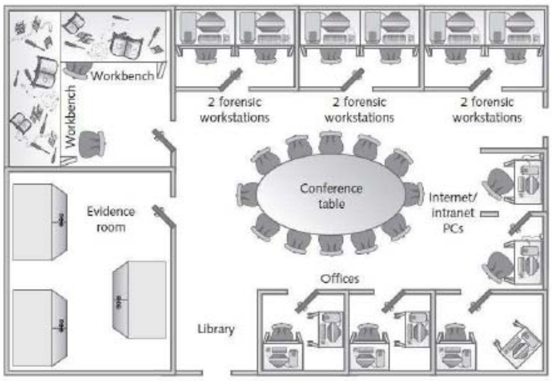


Figure 2. Regional Computer Forensics Lab Layout

In general, the evidence room will require the most secure place with no window on all the walls. In addition, it is common to have a separate room for investigators and supervisors since they might be working on different cases at the different times.

Based on the common practices discussed above, it was proposed for the laboratory to have a single entry door and there is a separation between private and public area. The public area will include conference room, workstations to access Intranet and/or Internet, and library area. The private areas will have 2 different areas: Private workstation areas, which include supervisor area and investigators area and Storage area. Then the most inner area and most secure area will be the evidence room, so called restricted area. The diagram shows the proposed design.

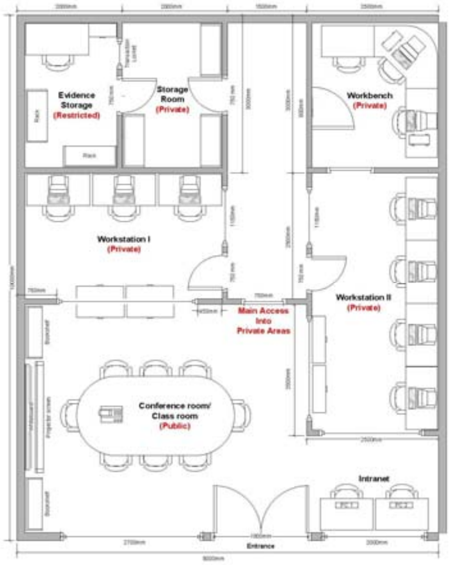


Figure 3 – Proposed Design Layout

Next, the room also needs special installation for Electricity, Security System, Telecommunication, Lighting and Air Conditioning System. The electricity for each of the working areas is best designed with different power line backed with different circuit breaker. One CCTV camera to monitor the public area and two CCTV cameras to monitor highly secure storage room will be used. Telecommunication cable for voice communication and network cable will be pulled into each workstation. The network cabling will be discussed in the next diagram. The diagram below shows the all the required installation including the lighting system. Separate Air conditioning system for restricted storage area is required to provide continuous (7x24 hours, 365 days) cooling for the room.

Figure 4. Proposed Electrical and Security System

The next design is the physical network connection to all the workstation. Since the network connection is critical to the overall work of the supervisors and investigators, the network is designed to be redundant, i.e. primary and secondary network. The secondary network will be used when the primary network is down. The 8 port switches are used to distribute the network connection to all the investigators' workstations. The Conference room in the public area only has single wired connection. If all of the wired connections in the private or public failed, Wireless Access Point is available for use. There are 2 wireless access points: 1 Access Point (AP) for the private area and 1 AP for the public area. The diagram below shows the design the network connections for the lab.

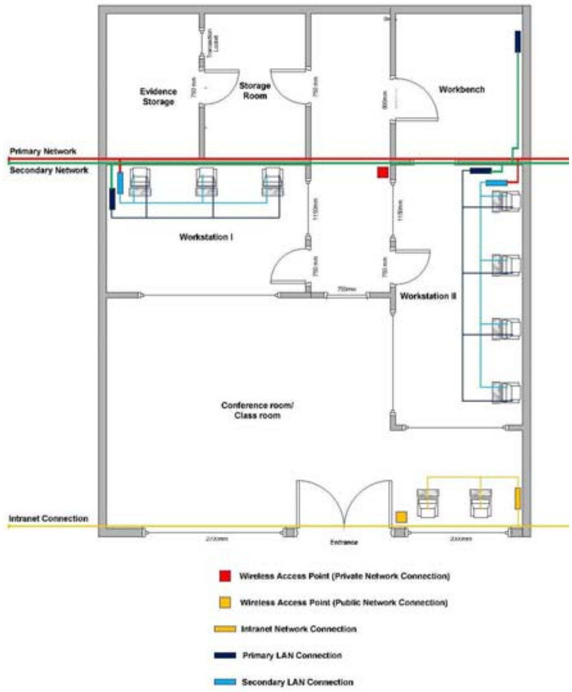


Figure 5. Proposed Network Connections in the Lab

Finally, the private area of the lab is required to be equipped with anti-static material, especially in the storage room and workbench room. This is to ensure, no damage to the evidence is due to the incidental static electric discharge.

In the public area, the following rooms will be found:

1. Conference Room/Class Room
   The conference room can be used for conducting class or training, meeting, and other activities. There are several presentation devices such as LCD projector, screen projector, whiteboard, and conference table in this room.
2. Library
   There is a small library in the public area. All books in the library are reserved (internal use only).
3. Intranet Workstations
   The intranet workstations are used for conducting digital forensics process internally without connecting into the internet, since the internet is insecure. EnCase and FTK software should be installed on the intranet workstations.

In the private area, the following are the equipments and work areas:

1. Biometric based access control and CCTV camera (covering the entrance of public to private area)
2. Workstation I (4m x 2,5m)
   The workstation consists of one Mac PC (Mac OS), and two other PCs (Windows OS). This room can be used for up to three people. EnCase and FTK software should be installed in those PCs. Other additional forensics software (if any) may also be installed. All PCs are connected into the double networks.
3. Workstation II (5m x 2,5m)
   The workstation consists of one Mac PC (Mac OS), and three other PCs (Windows OS). This room can be used for up to four people and has direct access from the workbench. So, the duplicated evidence can be transferred directly and immediately to this workstation for the forensic process. EnCase and FTK software should be installed in those PCs. Other additional forensics software (if any) may also be installed. All PCs are connected into the double networks.
4. Workbench (3m x 2,5m)
   The workbench is the place where the duplication process is being performed. Most of the digital forensics hardware and tools are used in the workbench. All portable devices are usually used in the workbench. This room is anti static.
5. Storage Room
   Storage room is used for storing our tools and equipments. All forensics portable devices are stored here and will be taken when it is needed. There is a door to the evidence storage room which has the highest restriction level in our laboratory, and a transaction locket for transferring evidence in to/out from. The access door is equipped with CCTV camera and thumb scanner/fingerprint scanner.

In the restricted area, one will find the evidence storage room, which is used to store evidence, such as hard drive, PCs, etc.

5. Conclusion

Once the laboratory is completed, attention should be focused on the overall laboratory management, which includes Human Resource Management, Case Management, Facilities Management, Financial Management and

Reaccreditation. The challenge of maintaining digital forensics laboratory is the capability of the laboratory to evolve as new storage, hardware and software processing technology becomes available. The other challenge is to develop the trusted and competent human resource to work with and handle the digital evidence in a professional manner.

6. Future Works

IDSIRTII is still in its infancy and the role of IDSIRTII will continue to expand, including expanding services to private entity. The future works of IDSIRTII also include building malware analysis and honey net laboratories, which are essential to protect Indonesia critical infrastructure against cyber attacks that are on the rise. Finally, the success of these labs depends heavily on the research on how to detect and prevent the malware from attacking our critical infrastructure.

References

[1] Dirjen Pos & Telecommunication – http://www.idsirtii.or.id (accessed 29th August 2009)

[2] General Elections Commission (KPU) web site attack - http://news.zdnet.co.uk/security/0,1000000189, 39152461,00.htm (accessed 29th August 2009)

[3] Establishment of ID-SIRTII - http://www.idsirtii.or.id/index.php/sejarah-id-sirtii/ (accessed on 29th August 2009)

[4] Jones, A. & Valli, C. (2009), "Building a Digital Forensic Laboratory Establishing and Managing a Successful Facility", Butterworth Heinemann and Syngress Publishing.

[5] Mohay, G., Anderson, A. & Collie, O. (2003), "Computer and Intrusion Forensics," Artech House Publishers.

[6] Reyes, A. Wiles, J. (2007), "The Best Damn Cybercrime and Digital Forensics Book Period," Syngress Publishing.

[7] Rowlingson. R. (2004), "A Ten-Step Process for Forensic Readiness," International Journal of Digital Evidence, Winter 2004, Volume 2, Issue 3.

[8] Regulation of Minister of Information and Communication regarding IDSIRTII - http://www.postel.go.id/content/ID/regulasi\telekomunikasi\kepmen\id-sirtii%202007.pdf (accessed 29th August 2009)

[9] Marcella, Albert J. Jr., Menendez, D. (2008), "Cyber Forensics: A Field Manual for Collecting, Examining, and Preserving Evidences of Computer Crimes," 2nd edition, Auerbach Publications

[10] Brown, Christopher L. T., (2006), "Computer Evidence: Collection & Preservation," Charles River Media, Inc.

[11] Electronic Crime Scene Investigation—A Guide for First Responders - www.ncjrs.org/pdffiles1/nij/187736.pdf (accessed 29th August 2009)

THE EFFECTS OF VAK LEARNING STYLE ON LEARNING OUTCOMES IN POWERPOINT-BASED TEACHING

Francisca H. Chandra

Institut Sains Terapan dan Teknologi Surabaya (iSTTS)
fhc@stts.edu, fhc@eduluminari.com

Abstract

The purpose of this study is to explore any differences in learning outcomes among students with different VAK learning styles when learning from PowerPoint modules with varying cognitive load levels. Participants are 59 (fifty-nine) students of iSTTS enrolled in Management of Information Systems class. Participants are divided into two groups. Group A (control) learns from a PowerPoint module taken directly from the course textbook. This original module is mostly text, with minimal graphics. Group B (experiment) learns from an equivalent PowerPoint module that has been modified to include more graphics and animations according to the cognitive load theory. Learning outcomes are measured using retention test. Overall test results indicate that the learning outcomes are better for the experiment group compared to the control group. From the VAK perspective, the aforesaid difference is greater among visually-oriented students from both groups (α=0,005). The auditory-oriented and kinesthetic-oriented students in the experiment group also received better results that their counterparts in the control group, although the difference is not statistically significant. Additionally, in experiment group, visually-oriented students achieved better results compared to the auditory-oriented and kinesthetic-oriented students.

Keyword: cognitive load theory, PowerPoint, learning style

1. Introduction

Information technology has totally changed the way we learn and teach. The modern student is hourly faced with multiple digital channels of dense visual stimuli, such as TV, DVD, and electronic games. The same students therefore possess advanced skills in assimilating and processing visual information. They prefer learning material containing visual as well as verbal information. They are known to dislike learning material that contains only text [20]. Their teachers and lecturers must adapt to this new way of learning. Traditionally lecturers used to employ primarily verbal and textural material, such as verbal lectures, writing on blackboards, and text-only lecture notes [7]. With the advent of computers, LCD projectors, and PowerPoint, lecturers have been racing to make use of these new technologies.

There are two main usages of PowerPoint in teaching. The first is to transfer content from a Word format into PowerPoint using bullet points and templates. The resulting slides are commonly dense with text and bullets. The second usage is to make use of all the features of the software, such as clip arts, graphics, sounds, and animation, without paying attention to the cognitive load these features impose on the learners. The resulting slides often contain high, extraneous cognitive load.

This study will investigate the influence of VAK learning styles when learning from PowerPoint modules with different cognitive load levels. PowerPoint modules will be use as self-paced learning material. The learning outcomes will be measured using retention test.

2. Learning Styles

In receiving and processing new information, every individual has their own preference. There are those who prefer graphical and visual information, those who prefer to study in the morning, those who prefer to study in groups, and those who prefer to study while listening to music. All of these preferences are called learning styles. There is no commonly accepted definition of learning styles, even the term itself varies, alternative called thinking styles, cognitive styles, learning modalities, or learning preferences [6]. All of them refer to the ways a person receives and processes information.

Studying the various definitions, theories, and models, [6] identified 71 learning style models. Further, Coffield, et al selected 13 most influential models. One of them is the model developed by Dunn & Dunn. This model is based on 5 stimuli strands, a) environmental, b) emotional, c) sociological, d) psychological, and e) physiological, all of which affects who an individual learns [9] in [6]. Physiological stimulus determines the perception (visual, auditory, kinesthetic, or tactile), time-of-day energy levels, the need for intake (food and drink) and mobility while learning. This study only covers the physical preference modalities, the VAK (Visual, Auditory, and Kinesthetic).

The Dunn & Dunn model tends towards measuring preferences than strengths. One of the positives of this model is that it works on reinforcing existing strengths instead of overcoming weaknesses.

The visual modality accesses images, created or memorized. Color, mental portrait, and pictures dominate this modality. Visual learners exhibit a number of common characteristics, they are organized, attentive, and well-groomed, they remember pictures well, prefer to read than being read to, and close their eyes when memorizing things.

The auditory modality accesses sounds and words, created or memorized. Music, tone, melody, rhythm, dialogue, and sounds dominate this modality. Auditory learners exhibit a number of common characteristics, they easily lose focus, speak with melodious cadence, prefer to listen than to read, read text outloud, and like to discuss matters.

The last modality is the kinesthetic modality. This modality accesses all motions and emotions, created or memorized. Movement, hand-eye coordination, cadence, emotional response, and physical comfort dominate this modality. Kinesthetic learners exhibit a number of common characteristics, they move a lot, touch their conversation partners, stand close to others, learn by doing, and memorize while walking [8].

Overall, a learner uses all three modalities to receive and process information. However, the VAK theory submits that there will be one or two dominant modalities for a given learner. The dominant modality determines what is the best learning style for a given individual. It is worth noting, however, even for the same individual, not all learning tasks are the same. Different learning styles are more effective for different learning tasks.

Learning styles can be formed and shaped by the environment. For example, the traditional learning process prescribe a certain sequence of learning. Up to the third grade, teaching is done primarily through the kinesthetic modality. The next five grades often learn visually, and gradually the teaching focuses more on auditorial delivery. University education is traditionally auditorial, with lecturers talking in front of the class, writing in the blackboard, handing out text-heavy lecture notes, all of them full of auditorial stimuli.

This traditional learning method persists despite the fact that most people are visual learners. [14], state that 40% of American students are visual learners, 20%-30% are auditory learners, and the rest are kinesthetic learners. According to [10], the learning style of most technical students are visual, sensing, inductive, and active.

For the generation born between 1982 and 1991, often called the Net Generation (Net Gen), dense visual stimuli is everyday staple. In school and out of school, every day and every hour, they are faced with all kinds of digital channels delivering rich visual information. TV, DVD, electronic games, and computers are part of the Net Gen daily life. The Net Gen is therefore exceptionally skilled in assimilating and processing visual information. They expect education to be delivered visually, often refusing to read text-heavy, graphics-poor material [10].

Therefore the most appropriate learning method for the modern student is multimedia learning, which can present a combination of verbal and pictorial representations, fully utilizing the visual and auditory modality.

Multimedia learning

Learning methods can be analyzed from three perspectives, 1) delivery medium, 2) representation format, and 3) sensory modality [15]. The format perspective is concerned with how the material is represented. Material can be represented verbally with on-screen text or narration, and pictorially with static graphics or animation. The sensory modality perspective is concerned with the learner's sensory reception, with the visual and auditory channels [15].

There are two elements in a multimedia learning presentation. The first is the visual element, delivered in the forms of static visuals (photographs, graphics, illustration) and dynamic visuals (animation, video). The second is the auditory element, such as on-screen text, narration, and background music. In multimedia learning, it is often the case that the two elements are simultaneously used to deliver information. In this situation, a problem arises where the processing demands of a learner become greater than the learner's processing capacity. This situation is called cognitive overload [4]. When designing PowerPoint-based teaching, a popular method of multimedia learning, the educator must be careful to minimize or altogether avoid cognitive overload.

Cognitive Load Theory

Cognitive Load Theory (CLT) refers to the way humans process information in their brains. Three types of memory are used in processing information, 1) sensory memory, 2) working memory (short-term memory), and 3) long-term memory. The cognitive load theory is mainly concerned with the working memory, or short-term memory. As the name implies, this type of memory can only retain a limited amount of information, and only for limited duration. In certain conditions, these limitations can hinder the learning process. The main principle of CLT is that the quality of learning will increase if sufficient attention is given to the function and limitations of the working memory. According to CLT there are three cognitive load types that affect the working memory, 1) Intrinsic Cognitive Load, 2) Germane Cognitive Load, and 3) Extraneous Cognitive Load [5],[16].

Intrinsic Load is the cognitive load that arises naturally from the information being processed. The level of the load is dependent on the level of interactivity between the elements of the information. For example, to memorize new words from a foreign language, the intrinsic load is low. However, when the same learner is asked to combine the same words into a meaningful sentence, the added interactivity between the information elements increases the intrinsic load. The level of of load depends on the content. Intrinsic load, being a natural part of the information, cannot be avoided or reduced. They must instead be managed, and there are several ways to do that [5], [16].

Germane Load is the cognitive load created by learning activities intended to improve learning outcomes. This is an effective load created by the use of schemas, constructs, and automations. An example of germane load is providing the learners with working examples relevant to the material, which will increase cognitive load, but in a positive context.

Extraneous Load is the cognitive load created by the way the learning is presented. This load can by fully controlled by the educator. Learning design that does not take this load into account tend to inhibit the learning process. This load is prevalent in multimedia learning, where the problems are characterized by the issues of *redundancy*, *split attention*, and *modality* [5], [16].

Cognitive Load Theory in Multimedia Learning

Cognitive Load theory in multimedia learning assumes three things, 1) *Dual Channel*, 2) *Limited Capacity*, 3) *Active Processing*.

Dual Channel is the concept that humans have two independent channels by which they process information, the verbal channel and the visual channel. The visual channel processes information captured through the eyes, such as pictures, illustrations, video, and screen text. The verbal channel process information captured through the ears, such as sound, music, and voice. The eye and the ear are where sensory memory is captured, which is then transmitted via the dual channels into Working Memory, which is then passed to Long-Term Memory. Educators must be mindful of the various limitations of the channels and the memory spaces, which must be managed effectively to improve learning quality.

Limited Capacity. Humans can only receive a certain amount of information within a given period [19]. When a learner receives information presented as a picture and accompanied by a spoken narrative about the picture, the student will need to simultaneously process the visual information presented in the picture and the verbal information presented in the narrative. This situation is called *split attention*, where the learner needs to focus on both the visual and the verbal information. Considering that both are presenting the same material, the overall effect is redundant, and the load is extraneous [4].

Active Processing is a concept where the learner can learn better if they fully attend to the material presentation. The learner will select relevant information, organize them into a coherent mental structure, and integrate them into their prior knowledge.

PowerPoint

PowerPoint is a Microsoft software product that provides a graphical interface to design multimedia slides for display on computer or projector screen. This software combines graphics, video, diagrams, animations, sounds, and text to create a multimedia presentation. PowerPoint interacts well with other products from the Microsoft Office Suite such as Microsoft Word and Microsoft Excel.

History of PowerPoint

The first PowerPoint version was developed by Bob Gaskin and Dennis Austen from the software house Forethought. The software was called “Presenter”. In April 1987, Forethought released PowerPoint 1.0 for Apple Macintosh. At that time it was still in black and white, and its main function was to create simple transparency slides for use with overhead projectors. The color version was released a year after the release of color Macintosh. In July 1987, Microsoft bought the software for 14 million US dollars. In 1990, PowerPoint for Windows 3.0 was released as part of the Microsoft Office Suite.

PowerPoint and education

Initially PowerPoint was targeted at salesmen who could use it to create attractive sales presentations. The business world in general also appreciated PowerPoint’s easy-to-use features. Educators immediately saw that the same features that were useful to the business world was also useful to them. Unfortunately, at that time most PowerPoint books were user manuals focused on how to use the various software features. There were very few literature on how to utilize and integrate those features into effective teaching presentations. The lack of such literature and PowerPoint’s ease of use proved to be a trap for bad teaching [13].

PowerPoint features

Like most other presentation software, PowerPoint revolves around the *slide* object. Each slide is the digital equivalent of single physical slide used in slide projectors, which were ironically made obsolete by PowerPoint. A slide can contain text, graphics, video, and other objects. Compared to other software in the Microsoft Office Suite, PowerPoint is simpler to use. This advantage can sometimes become a disadvantage for educators, who often use only the standard, more common features.

Such standard features include template, AutoContent Wizard, and bullet points. Also included is the basic word editing features used in Microsoft Word. Bullet points and AutoContent Wizard are two features whose pros and cons are often debated in the education world [1], [12], [23], [24].

PowerPoint also provides three animation types to accompany various transitions between slides, 1) *entrance*, 2) *emphasis* and 3) *exit.* These animations can be further controlled with the *custom animations* feature.

PowerPoint version 2007 brought about major feature changes. Responding to criticism of its AutoContent Wizard, PowerPoint 2007 no longer provides the feature. Responding to criticism of the bullet point feature, PowerPoint 2007 provides the SmartArt feature, which this author considers a useful feature to encourage more meaningful presentation than the standard bullet points.

SmartArt provides an easy way to present complex information in graphical format. It is especially powerful in depicting the various connections and processes in a given subject matter. There are eight SmartArt categories, List, Process, Cycle, Hierarchy, Relationship, Matrix dan Pyramid, where each category accommodates a certain process and relation concept.

Managing Cognitive Load in PowerPoint Presentation

The management of cognitive load in PowerPoint presentations should be based on a) the seven principles of multimedia design [15], b) nine ways to reduced cognitive load in multimedia learning [17] and c) five ways to reduce cognitive overload in PowerPoint [2].

The seven principles of multimedia design are:

Multimedia Principle. Learners learn better through graphics and text compared to text-only.

Spatial Contiguity Principle. Learners learn better when the related text and graphics are presented close together.

Temporal Contiguity Principle. Learners learn better when the text and graphics are presented simultaneously rather than sequentially.

Coherence Principle. Learners learn better when irrelevant words and graphics are removed. Irrelevant words and graphics are considered to be extraneous load.

Modality Principle. Learners learn better when utilizing both sensory modalities, visual and verbal. For example, a presentation should contain both graphics and narration.

Redundancy Principle. Learners learn better when there are no redundant modalities. An example of redundancy is an animation accompanied by both spoken narration and written on-screen text.

Individual Difference Principle. The effect of multimedia design is more pronounced on low knowledge and high spatial learners compared to high knowledge and low spatial learners.

Further to the seven principles, [17] promote nine ways to reduce cognitive load in multimedia learning. The nine ways are 1*) off-loading, 2) segmenting, 3) pretraining, 4) weeding, 5) signaling, 6) aligning, 7) eliminating redundancy, 8) synchronizing,* and 9*) individualizing.*

Five principles specifically for designing effective PowerPoint teaching presentations are: *1) The Signaling Principle, 2) The Segmenting Principle, 3) The Modality Principle, 4) The Multimedia Principle,* and *5) The Coherence Principle* [2].

This study uses on the aforementioned principles to modify existing PowerPoint modules to reduce their cognitive load levels. In this study, the presentation of verbal information is limited to on-screen text. No sounds, narration, or music is used.

When designing PowerPoint-based teaching, educators must create presentation that functions as *cognitive guidance*:

1) The presentation needs to use both visual and verbal methods, utilising relevant ClipArt and SmartArt features. Avoid using pictures only as decoration. If there is no relevant pictures that can be used, consider using *signalling* by color-coding or other methods.
2) Populating slides with excess information will create extraneous cognitive load. However, if the material content cannot be meaningfully reduced, use *segmenting* and *signalling* to better structure the delivery. Use the *slide sorter view* to review the overall flow of the presentation.
3) The presentation needs to be designed in such a way that assists the learner to select, organize, and integrate the material with prior knowledge (Atkinson, 2004).

3. Methodology

3.1 Subject and Design

The subjects are 59 students of Institut Sains Terapan dan Teknologi Surabaya who are enrolled in the Management of Information Systems class. The students are divided into two groups, A (28 students) and B (31 students). At the beginning of the experiment a test is used to determine the student's learning styles. The instrument is a questionnaire from [8,166-167]. The module is taken from textbook "Management Information

System" by Raymond McLeod Jr. and George P. Schell, 10th edition.

There are four PowerPoint modules used in the subject, which is a downloadable student resource packaged with the aforesaid textbook. These original modules contain mostly text, with minimal graphics. These modules are then modified with the goal of reducing extraneous load by 1) adding context-appropriate graphics in accordance with the multimedia principles of dual channels and sensory modality, 2) weeding the content, meaning placing text appropriately so it is integrated with the graphics, 3) signaling by using PowerPoint 2007's SmartArtGraphics feature, and 4) adding animation and color to emphasize key learning points.

The study measures the learning outcome using retention test. Each module has a test consisting of 10 multiple-choice and 10 true-false questions. The test for all modules is derived from the subject's textbook.

3.2 Procedure

At the beginning of the experiment, the students are tested to determine their learning styles. The result of that test is as in table 1.

Table 1. Subject by Learning Style

	Subject by learning styles						Total	
Group	V		A		K			
		%		%		%		%
Experiment	17	55	8	26	6	19	31	100
Control	11	39	8	29	9	32	28	100
Total	28	47	16	27	15	26	59	100

Table 1 show that the visual learning style is the most dominant in the group, at 40%-55% across the groups. The rest are balanced between the auditory and kinesthetic learners.

The experiment is carried out in the computer laboratory. Students in the control group studies from the original module with High Cognitive Load (HCL), while those in the experiment group studies from the modified module with Low Cognitive Load (LCL). There are 4 modules that are explored in this study, modules A, B, C, and D. Overall, there are 8 modules and module variants: AHCL, BHCL, CHCL, DHCL, ALCL, BLCL, CLCL, DLCL. The experiment is carried out during the MIS lecture and each student is given two modules for every session. When learning a module, the students are given 15 minutes to review the material and are asked not to take any notes. Following the review, students are given the module's retention test, which is administered electronically on the laboratory's computers. The scoring awards 1 point for a correct answer, zero point for an incorrect answer, and zero point if the question is not answered. The test is sequential, i.e. it does not allow a student to go back to a previous question. The test time limit is 8 minutes for the whole test. The test result is immediately displayed onscreen to the participant after the test is completed, and it is also stored on the computer.

After finishing the first module, students are given a rest period of 10 minutes, after which they proceed to the next module.

4. Results and Discussion

The test results are validated statistically using independent *t*-test. Table 2 shows that for all the modules, there is statistically significant difference between test scores for the original, High Cognitive Load version and the modified, Low Cognitive Load version.

Table 2. Retention Test Score Results by Module and Cognitive Load

Module	N	Mean	SD	F	Sig.
AHCL	28	12.71	1.761	4.497	0.021
ALCL	31	14.10	2.663		
BHCL	31	13.50	2.333	0.010	0.000
BLCL	28	15.87	2.078		
CHCL	28	12.36	2.512	2.669	0.000
CLCL	31	15.52	1.630		
DHCL	31	12.68	2.195	0.186	0.030
DLCL	28	14.45	2.234		

The figures in tables 3, 4, and 5 show that among the visual learners, there is significant difference between the experiment and control group. Test results show that the experiment group, who learned from the LCL modules, scored higher than the control group, who learned from the HCL modules. For all modules (A, B, C, and D) the test results are consistently higher for LCL modules compared to HCL modules.

Table 3. Retention Test for Visual Learners Across All Modules

Module	N	Mean	SD	F	Sig.
AHCL	11	11.91	1.514	4.168	0.030
ALCL	17	14.88	2.690		
BHCL	11	11.91	1.758	3.764	0.000
BLCL	17	15.82	2.430		
CHCL	11	12.55	2.659	0.774	0.001
CLCL	17	15.71	1.611		
DHCL	11	12.64	2.014	0.111	0.004
DLCL	17	15.06	1.983		

Table 4. Retention Test for Auditory Learners Across All Modules

Module	N	Mean	SD	F	Sig.
AHCL	8	12.63	1.506	1.136	0.183
ALCL	8	14.00	2.330		
BHCL	8	14.13	1.458	1.776	0.083
BLCL	8	15.75	1.982		
CHCL	8	13.13	2.949	2.056	0.089
CLCL	8	15.38	1.847		
DHCL	8	13.13	2.800	0.185	0.385
DLCL	8	14.38	2.774		

Table 5. Retention Test for Kinesthetic Learners Across All Modules

Module	N	Mean	SD	F	Sig.
AHCL	9	13.78	1.856	0.174	0.108
ALCL	6	12.00	2.098		
BHCL	9	14.89	2.571	2.318	0.278
BLCL	6	16.17	1.169		
CHCL	9	11.44	1.810	0.094	0.001
CLCL	6	15.17	1.602		
DHCL	9	12.33	2.000	1.400	0.610
DLCL	6	12.83	1.472		

For students with auditory learning sytles, although for all modules test results are higher for LCL modules compared to HCL modules, however the LCL test results are not significantly different than the HCL test results. Kinesthetic learners, however, showed inconsistent results across all modules. Module A showed slightly higher results for HCL modules, while modules B, C, and showed the reverse. Only module C show significant difference between LCL and HCL modules ($F(1,13) = 0.094$, $p = 0.001$).

5. Conclusion

Experiment results showed that there are more visual learners compared to other styles, this is consistent with [14], [10] who noted in their studies that most technical students are visual learners. Also the subjects are born around the year 1990, often called the Net Gen [20]. The visual learners score worse in modules with no or minimal graphics. The difference is even more maker when compared to the auditory learners. This is consistent with [20] study which found that visual learners disliked material with too much text. For auditory students, their test results did not vary among the module variants, because for them the words are more important that the graphics. Also interesting to note is the inconsistent results among the kinesthetic learners, which is most likely because the study material is ill-designed for their learning styles. Overall, the experiment group scored higher than the control group.

The author's hope is that educators would pay more attention to the principles of cognitive load, like weeding and signaling, when designing PowerPoint presentations. This study is limited to the use of PowerPoint modules are self-paced learning and the test is limited to retention test. Further studies should explore the use of PowerPoint as material presentation and the use of transfer test to measure the learning outcomes.

References

[1] Atkinson, Cliff. 2004. *The Cognitive Load of PowerPoint: Q&A with Richard E. Mayer*. http://www.sociablemedia.com. Retrieved on 1 May 2008.

[2] Atkinson, Cliff., & Mayer, Richard E. 2008. *Five ways to reduce PowerPoint overload*. http://www.sociablemedia.com/pdf/atkinson_mayer_powerpoint_4_23_04.pdf. Retrieved on 1 May 2008.

[3] Barbe, WB. and M.N. Milone,"What We Know About Modality Strengths," *Educational Leadership,* Feb. 1981, pp. 378-380. In Felder R. M. 2002. Learning and Teaching Styles in Engineering Education. *Engineering Education, 78*(7), 674-681 (1988).

[4] Clark, R, Mayer, Richard E., & Feldon, David, F. 2005. Five Common but Questionable Principles of Multimedia Learning. In Mayer, R. (Ed.) *Cambridge Handbook of Multimedia Learning*. Cambridge: Cambridge University Press.

[5] Clark, R., Nguyen, F., & Sweller, J. 2006. *Efficiency in Learning: Applying Cognitive Load Theory*. Pfeiffer.

[6] Coffield, F., Moseley, D., Hall, E., & Ecclestone, K. 2004. *Learning styles and pedagogy in post-16 learning. A systematic and critical review.* London: Learning & Skills Research Centre.

[7] DePorter, B., & Hernacki, M. 1992. *Quantum Learning: Unleashing the Genius in You.* Translated by: Alwiyah Abdurrahman. 2007. Bandung: Kaifa.

[8] DePorter, B., Reardon, M., & Singer-Nourie, S. 1999. *Quantum Teaching: Orchestrating Student Success.* Translated by: Ary Nilandari. 2009. Bandung: Kaifa.

[9] Dunn, R. 2003. *The Dunn and Dunn learning style model and its theoritical Cornerstone*. In Coffield, F., Moseley, D., Hall, E., & Ecclestone, K. 2004. Learning styles and pedagogy in post-16 learning. A systematic and critical review. Learning & Skills Research Centre.

[10] Felder R. M. 2002. Learning and Teaching Styles in Engineering Education. *Engineering Education, 78*(7), 674-681 (1988).

[11] Felder R. M., & Soloman, B. A. Learning Styles and Srategies. http://www4.ncsu.edu/unity/lockers/users/f/felder/public/ILSdir/styles.htm. Retrieved on 8 June 2009.

[12] Jones, A.M.2003. *The use and abuse of PowerPoint in teaching and learning in the Life Sciences: A Personal Overview*. http://bio.itsn.ac.uk/Journal/vol2/beej-2-3.pdf. Retrieved on 1 May 2008

[13] Klemm, W.R. 2003. *College Teaching*. Computer Slide Shows: A trap for bad teaching. 55(3), 121-124.

[14] Liu, Y., & Ginther, D. (1999). Cognitive Styles and Distance Education. *Online Journal of Distance Learning Administration, 2*(3), 18. In Sankey, M.D.2007. Multimodal Design for Hybrid Course Materials: Developing and Evaluating a New Paradigm for Course Delivery. A thesis submintted in fulfilment of the requirements for the degree of Doctor of Education. Faculty of Education Queensland University of Technology.

[15] Mayer, R.E. 2001. *Multimedia Learning*. Cambridge: University Press.

[16] Mayer, Richard E. 2005. *The Cambridge Handbook of Multimedia Learning*. Cambridge University Press.

[17] Mayer, Richard E., & Moreno, Roxana. 2003. Nine Ways to Reduce Cognitive Load in MultimediaLearning. *Journal Educational Psychologist.* 38(1),43-52. http://homepages.nyu.edu/~jpd247/2251/readings/mayer_moreno_2003.pdf. Retrieved on 27 May 2008.

[18] Mayer, Richard E., & Moreno, Roxana. 2008. *A Cognitive Theory of Multimedia Learning:Implications for Design Principles*. http://www.unm.edu/~moreno/pdfs/chi.pdf. Retieved on 27 May 2008.

[19] Miller, George A. 1956. The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information. *Journal Classic in the History of Psychology*. 1-16. http://psychclassic.yorku.ca/Miller/. Retrieved on 5 January 2008.

[20] Oblinger, D. G., & Oblinger, J. L. 2005. Is It Age or IT: First Steps Toward Understanding the NET Generation. In Oblinger, D. G., & Oblinger, J. L.(Eds.) *Educating the Net Generation*. An Educause e book.

[21] Richardson, J., *Working With People,* Associate Management Inst., San Francisco, Calif., 1984.

[22] Sankey, M.D.2007. *Multimodal Design for Hybrid Course Materials: Developing and Evaluating a New Paradigm for Course Delivery*. A thesis submintted in fulfilment of the requirements for the degree of Doctor of Education. Faculty of Education Queensland University of Technology.

[23] Tufte, Edward R. 2003. *The Cognitive Style of PowerPoint*. Graphic Press LLC.

[24] Tufte, Edward R. 2006. *The Cognitive Style of PowerPoint: Pitching Out Corrupts Within (Second Edition)*. Graphich Press LLC

CYBERCRIME FROM ISLAMIC SHARIAH LAW POINT OF VIEW

Irene Salomo[1], Charles Lim, M.Sc., ECSA, ECSP, CEH, CEI[2]

[1,2]Binus International University
[1]irene_math2003@yahoo.com.sg, [2]charles.lims@gmail.com

Abstract

The significant growth of cyber world recently has aroused many countries to regulate it by cybercrime law. Similarly, in 2008 Indonesia government passed Information and Electronic Transaction Law. However, most of those laws are enacted only according to common laws. Authors will explain what Islamic Shariah Laws in general, what kind of Shariah Laws that are related with cybercrime, and how Quran Islamic Shariah Laws can respond and be implemented towards cybercrime. Authors will base this paper on literature research related with this topic. Authors found out that some verses in Al-Quran and Hadith can judge cybercrime. In fact, there are more than one billions Moslems in the world, while 90% of 250 millions Indonesia populations are Moslem. Besides of that, most of the Moslems still consider Islamic Law as the highest law and principles, instead of common laws, this situation may be significant for the prevention and awareness campaign of cybercrime especially for Moslem. Therefore, it's worthwhile to comprehend and analyze the cyber crime from Islamic Shariah point of view.

Keywords: cybercrime law, Islamic Shariah Law

1. Introduction

There are some definitions of cybercrime. First, according to Forester and Morrison, it is any crime that uses computer as main weapon [1]. Second, cybercrime is any crime that exploits computer technology as the main component [2]. Finally, according to Tavani (2000), cybercrime is any crime that can be committed only through cyber technology and in cyber environment i.e. Internet [3]. Based on the third definition, the examples of cybercrime are the offences towards computer systems' confidentiality, identity and availability, like identity theft, hacking into others' systems, and distributing malicious code, also the infringement of digital intellectual properties' copyright, etc. On the other hand, cyber pornography and sexual abuse can be classified as cyber-related crimes since both can be committed with and without cyber technology. In this paper, authors will focus on the third definition of cybercrime which is stated by Tavani.

From study conducted by McAfee, cost of cybercrime may reach $1 trillion globally, this includes lost of intellectual property and the cost for fixing the problems [4]. Obviously, this is not a trivial number and it shows how serious cybercrime has threatened the real world, like business.

One of the methods to control cybercrime is by regulating it through cybercrime law. In Indonesia, cybercrime is specifically regulated in Information and Electronic Transaction Bill (*Undang-Undang Informasi & Transaksi Elektronik*) article 30-37, with fine ranging from 600 million to 12 trillion Rupiahs and 6 to 12 years jail.

The purpose of this paper is to discuss how Shariah Law can respond and be related to cyber crime.

2. What is Shariah Islam

One of Islamic distinctive contribution towards human life is Shariah Law. Different with other law, this law covers the entire aspect of Moslems' life [5], consequently cyberspace cannot be resistant to Shariah Law as well.

Shariah in an Arabic word means 'way' or 'path'. According to Oxford Islamic Studies Online, Shariah is "God's eternal and immutable will for humanity, as expressed in the Quran and Muhammad's example (Sunnah) [6]. It aligns with the fact that Islam is not only religion but also the way of life.

The sources of Shariah Law can be classified into following:

a. **The Qur'an**
It's the most fundamental and main resources of Moslem. According to Moslem, the Qur'an is word of Allah revealed to the Prophet Mohammad.

b. **The Sunnah**
It is the written tradition and practices of Qur'an by Prophet Mohammad, including his judgments, opinions, and decisions regarding certain issues. Therefore, Moslems are expected to follow his examples.

c. **Ijma (consensus)**
In case, there is no specific law available in either the Qur'an or the Sunnah, Moslem will

find the consensus from community. This derives from what Mohammad said that it is impossible for Moslem community to have an agreement on flaw.

d. Qiyas (analogy)

Qiyas means analogy, logic, thinking, and legal precedent that can be used as the last source of Shariah Law for Moslem, under situation of they could not find any guidance from the three previous sources [7].

The main objective of Shariah Law is to protect five important values which are religion, human life, intellect, linage, and property.

In Shariah Law, there are 3 types of crimes exist.

a. Hadd Crimes

This is the most severe crime with exact punishment according to the Qur'an, no diminishing or augmenting is allowed. These crimes include murder, apostasy from Islam, theft, adultery, defamation, robbery, and alcohol-drinking. Hadd crimes are considered as physical violence that are against Allah.

The punishments for had crimes are corporal punishments, such as amputation, hanging, stoning to death.

b. Tazir Crimes

It is least severe crimes compared with Hadd crimes, and considered against society. These kinds of crimes are not found in Qur'an so judges are given freedom to decide appropriate punishment for the guilty, as long as they are responsible to Allah and Moslem community. The examples of punishments for tazir crimes are canning, fines, counseling, public condemnation, confiscation of property, and so forth.

c. Qesas Crimes

This is retaliatory crimes, means that the victims have right to claim for retribution and retaliation. These crimes consist of physical violence, such as murder, and committing injure [8].

3. How Shariah Islam Relates with Cyber Crime

In September 2008, more than 300 Shia websites, including Grand Ayatollah Ali al-Sistani's website became the victims of Sunni hackers. While Shia hackers retaliated by attacking more than 77 Sunni websites [9].

Both examples may illustrate how cybercrime issues may be used as weapon to attack opponents, even among religious parties.

Although computer and Internet had not existed during the revelation and documentation processes of Allah's Words, we can still analyze how Shariah Islam relates and deals with cybercrime.

However we need to determine first in Shariah Law what kind of crime cybercrime is. Obviously it is not either Hadd or Tazir crimes since it is not physical attack or violence. Actually all crimes that cannot be classified under both crimes and are incomplete Hadd crimes, can be considered as Tazir crimes [10].

According to Moslem, Shariah comes from God who is perfect, so the Shariah Law must be infallible as well, unlike human law which is imperfect. Therefore, many Moslems still consider that they need only to comply fully with Shariah Islam, instead of common human-made law [11].

Computer system, software, identity card, passwords are all human's properties. Consequently we need consent before accessing them and are not allowed to obstruct it.

There are some verses in Al-Qur'an can be subjected to cybercrime:

a. ALLAH (God) said: "O believers! Do not enter houses other than your own until you have sought permission and greeted their inmates "(Ayah, An-nur, 27). It means that according to Islam, you can not intrude to someone's property and privacy, including computer system without their permission. The intrusions itself can be in form of virus, DOS (Denial of Service), worm, Trojan horse, and other attacks as what crackers did usually.

b. ALLAH said: "O believers! Avoid immoderate suspicion, for in some cases suspicion is a sin, Do not spy on one another" (Al-Hujurat, 12). It means that Islam does not allow you to spy others' properties and environments, like what crackers do in fingerprinting and social engineering actions.

c. Prophet Mohammad (peace be upon him) said: "It is better for a Muslim to mind his own business" (Al-Muatta, 1980). Therefore, as Moslems they are asked not to let their curiosity to make them intrude on others' areas.

d. Prophet Mohammad (peace be upon him) said: "Permission is for having a look" (Al-Bukhari, 1987). This means that having a look inside others' computer system only can be done when they give us permission, and it must not be used for other purposes, like hacking [9].

One of the biggest factors why cybercrime may happen is the existence of trust. We give trust for others for the good reason and purpose; unfortunately they misuse it for negative purpose. Islamic teachings also respond towards this issue. In Al-Bukhari, it is written that Prophet Mohammad (peace be upon him) said: "The signs of a hypocrite are three: Whenever he speaks, he tells a lie. Whenever he promises, he always breaks his promise. If you trust him, he proves to be dishonest. If you keep something as a trust with him, he will not return it." (Al-Bukhari, 1987). In Al-Qur'an it is stated that ALLAH said: "O believers! Do not consume one another's wealth through unlawful means; instead, do business with mutual consent." (An-nessa, 24) [10]. Regarding cybercrime issues, the crackers have violated these laws by abusing others' trust.

In Hadith, Prophet said "No harm shall be inflicted [on anyone] nor reciprocated [against anyone]' to criminalise emerging crimes". Therefore, it clearly implied that cybercrime is illegal since it causes harms towards computer system and human as well.

We can use those understandings to campaign and promote cybercrime awareness. The advantages of using religion approach are:

a. Shariah Islam is believed to come from God; hence it exerts more influences and has stronger appeal than other human-made law.
b. Most Moslems have already been familiar with these Shariah principles, so they will be more susceptible to aware about cybercrime if we move on from something familiar towards something less familiar.
c. Shariah Islam can reach broader element of society rather than cybercrime law itself that may be understood by certain elements only.

Some real actions that can be done are:

a. Give adequate understanding about cybercrime to Islamic clergies, like Imam and preachers. Ask them to relate their preaching sometimes with cybercrime. Because of their highly-respected positions and charisma, those clergies may use their influence to campaign security awareness towards our society that still consider their speaking as worthy cause.
b. Promotion through other media, like leaflet, Internet, poster, etc. and put some familiar religious terminologies there and emphasize that committing cybercrime is basically not different with violating Shariah Islam

4. Conclusions

Understanding how Shariah Law can be related with cybercrime may help us to campaign and promote security awareness, especially among Moslems, since Shariah Law is the highest law comes from perfect God according to Moslems, and recalled that 90% of Indonesia population consists of Moslems. Besides of that, any Moslems are expected to have fear of God so obeying His rules is uncompromised matter. Ultimately, we expect to reduce and prevent cybercrime cases in the future.

We also prove that Shariah Law can be responsive and used in the context of modern issues, such as cybercrime.

5. References

[1] Forester, Tom and Morrison, Perry, 1994, *Computer Ethics: Cautionary Tales and Ethical Dilemmas in Computing*, The MIT Press, America.

[2] Girasa, Roy.J., 2001, *Cyberlaw: National and International Perspectives*, Prentice Hall, America.

[3] Spinello, Richard.A and Tavani, Herman.T., 2001, *Readings in Cyberethics*, San Val, America.

[4] Mills. Elinor. *Cybercrime cost $1 trillion last year, study*. Located at http://news.zdnet.com/2100-9595_22-264762.html, January 2009.

[5] Casciani. Dominic. *Q&A : Sharia Law Explained* .Located at http://news.bbc.co.uk/2/hi/uk_news/7234870.stm, July 2008.

[6] Oxford Islamic Studies Online. *Shariah Definition*. Located at http://www.oxfordislamicstudies.com/article/opr/t125/e2168?_hi=2&_pos=1, n.d.

[7] Huda. *Sources of Islamic Law*. Located at http://islam.about.com/od/law/a/sources.htm, n.d.

[8] Madkoar. Mohammed. Salam. *Human Rights from an Islamic Worldview: An outline of Hudud, Ta'zir & Qisas*. Located at http://www.islamawareness.net/Shariah/sh_article002.html, n.d.

[9] Maghaireh. Alaeldin. *Shariah Law and Cyber-Sectarian Conflict: How can Islamic Criminal Law respond to cyber crime?*. Located at http://ccrimejournal.brinkster.net/alaeldinijccdec2008.htm, n.d.

[10] Rudolph. P., 2005 *Crime and Punishment in Islamic Law*, Cambridge University Press, England.

[11] Al-A'ali.Mansoor. *Cybercrime and the Law: An Islamic View*. Located at http://www.webology.ir/2007/v4n3/a46.html, September 2007.

IMPLEMENTATION OF PARTICLE SYSTEM USING SMOOTHED PARTICLE HYDRODYNAMICS (SPH) FOR SIMULATING LAVA FLOW

Khairul Hamdi

Telkom Institute of Technology
khairul.hamdi@gmail.com

Abstract

Lava flow simulation has a complex phenomenon because of many parameters involved. The method used to simulate lava flow is Lagrangian approach by considering fluid lava as compiled by the particle system. Smoothed method Particle Hydrodynamics (SPH) implemented for the interaction between particles that is element of the fluid. Physical parameters particle based fluid lava is density, viscosity, pressure, surface tension, temperature and external forces. The SPH method allows the use of particle systems can interact with other objects such as terrain. Determination of nearest neighbor particles using the staggered grid algorithm, function of smoothing kernel used the kernel Spiky kernel and Poly6 kernel that more in stability and accuracy. Collision detection between particles and the terrain using the methods Plane Sphere Sweep Test and rendering methods used quadtree algorithm. Produced lava flows more realistic to simulate some physical parameters.

Keywords: SPH, smoothing kernel, staggered grid, terrain

1. Introduction

One of the objects that quite interesting study is the modeling of volcanic phenomena. Area of Indonesia which lies a ring of fire is characterized by high volcanic activity and there are lots of volcanoes, there is a challenge for researchers to study the various volcanoes and make the modeling of the volcano as a learning tool for the public and the parties other stakeholders. One of the important modeling is the modeling of lava flow.

Regarded as a lava flow of fluid formed by the particle system. The behavior in the fluid is a result of the interaction between the particles in the system. Interaction between particles is modeled using the method of smoothed Particle Hydrodynamics (SPH). The SPH method is a method commonly used in Computational Fluid Dynamics (CFD)-based gridless Lagrangian approach. SPH method has the following advantages:

1. Suitable to be used to model fluid with high deformation and displacement interface;
2. has a level of accuracy, adaptive and high stability;
3. a very broad application of the modeling of microscopic (atomic components) to macroscopic scale (modeling clusters of galaxies);
4. SPH application has to include a variety of physical aspects such as the viscosity of the fluid, external forces, internal forces, density, heat transfer and so forth.

Lava flow is a complex event with a variety of physical and chemical parameter. The purpose of this research are as follows:

1. Simulation of lava flows that have physical parameters of density, pressure, viscosity, temperature and gravity;
2. simulation of lava flows that can interact with other objects;
3. visualize of objects and more realistic lava flows.

Previous research conducted by [Cani, 1999] using the SPH method to model the flow of lava by calculating the density, viscosity and heat transfer. Further development to add other parameters such as particle size, smoothing length and rendering carpet that surrounds the top of the fluid for a more realistic visualization.

2. Terrain and Particle Systems

Terrain is generally implemented for landscape objects. The use of terrain as a model of this landscape found in rally games and real-time strategy. Terrain modeling can be build using data heightmap dummies and the processed data from the digital map.

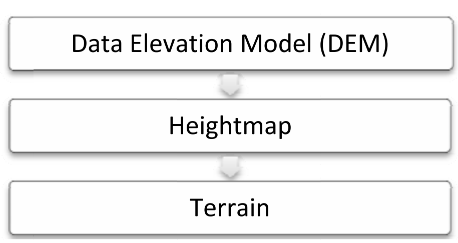


Figure 1. ***Heightmap Data Processing.***

Heightmap data containing elevation information used as input data to generate terrain. Terrain is used as a model volcano, volcano model used is the actual model for heightmap input data obtained from digital maps satellite captured. Stages in the heightmap data processing are as follows:

1. DEM data derived from satellite imagery, each point of the DEM map contains information about location and height;
2. DEM data using a heightmap converted using 3DEM software, the output obtained is the image grayscale, white color indicate maximum height and black for the lowest areas;
3. adjust color depth using the heightmap image editor, with this application is the color depth of 8 bits, so has the range 0 to 255, meaning a 256-scale terrain elevation.

Heightmap data can be either RAW or BMP file, grayscale image is then used as input data to generate terrain. In order to generate terrain is terrain engine.

The particles are entities that do not have space dimension, an object that has the parameters position, speed and style of interaction between each other and can be influenced by external forces. Particle system is a collection of point masses in three-dimensional space associated with a particular force and influenced by external forces such as gravity and friction between particles.

3. Physical Parameter of Lava Flow

Lava flow as a natural phenomenon is a complex system, complex and involves many physical and chemical parameters. To simulate the lava flow, the various parameters that have simplified and the parameters a little contribution can be ignored. Physical parameters which are owned by the lava flows are as follows.

1. Viscosity dependent of the material composition of lava issued from the magma chamber. For the same chemical content, viscosity values will decrease exponentially as the temperature decreases;
2. density of lava influenced by the it's material. Containing molten lava flow and the aluminum silicate melt due to high temperatures. This material makes a high - density lava. Mass density value of basaltic lava is 2500 kg.m^3;
3. the temperature of lava is highest when from the crater. The temperature of lava in this state between 1200 ° - 1400 ° C [ESDM, --]. This high temperature decreases exponentially affected by the surrounding environment. Lava thermal energy absorbed by the surrounding environment such as soil and mountain air. Lava cooling rate proportional to the temperature gradient between the lava with the surrounding environment;
4. color depends on the temperature, according the Wien's displacement law when the temperature is reduced, the black body radiation curve moves to the direction of lower intensity and longer wavelength [Wikipedia, 2008] according to equation (1):

$$T\lambda_{\max} = 2.898*10^6\, nmK \tag{1}$$

Temperature of 3000 K, black body radiation which has a high intensity with 700 nm wavelength which produces a red color. For the lava, the initial temperature of 1200 ° C will make incandescent lava and reddish orange, the color of lava low temperature increasingly faded into blackish red.

4. Design and Implementation of Lava Flow

Design and implementation of lava flows using SPH method is done by making a model volcano using the terrain, the particle moves along the terrain. The particles have collision detection with the terrain and is reflected by the reflecting factor 0.25. Interaction between particles using the SPH method.

4.1 Generating Terrain

The method of *generating terrain* is vertex *terrain* elevation map in accordance with the scale of points contained in the pixels of heightmap. Processing into terrain grayscale bitmap, *grayscale* bitmap data converted first into polygonal *mesh.* This method, considered as a *heightmap* image lattice has height values. This grid will be changed into series composed vertex. The value of x and y coordinates of each pixel is converted into vertex position. The value of each color pixel location will be transformed into z position in each vertex. When done reading the *grayscale* values in the range of 0 to 255, then made projections altitude, with the highest point is 255 and the lowest point is 0, so the *heightmap* which has formed the scale 256.

Four square pixels that have a dimension of 2 x 2. Vertex Storage as a list containing information on the position coordinates x, y and z. Data stored triangle has 3 index vertex list, which describes vertex-vertex used for each triangle. Thus there are two types of data that is *vertex* and *vertex* index. For vertex positioning, also calculated the surface normal for each triangle which is the result of the cross product between 2 vectors normalized.

4.2 Determination of Particle Position in The Terrain

Heightmap size used is 128 x 128 pixel to make the program run faster because of the large heightmap need greater computing. Each point of one pixel formed on the terrain vertex. Three vertex forming a triangle, so that four vertex forming two triangles as in Figure 2.

Figure 2. ***Determination Distance per-Pixel on Each Vertex.***

Triangle [0] has a vertex corresponding to vertex [0], vertex [1] and vertex [128]. Triangle [1] corresponds to vertex [1], vertex [128] and vertex [129]. Number of triangle is formed which has 32,258 from triangle[0] to triangle [32,257].
Two of the hypotenuse triangle coincide to form a box. The number of boxes in one row or one column is 127 pieces. Each box has the index of the box [1] until the box [16,129].

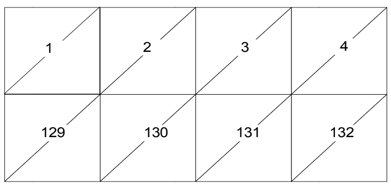


Figure 3. ***Indexing Box for Determining The Collision Area.***

A particle that has a position (x, y, z) will experience a collision with the terrain when the same position in z-axis in accordance with projected terrain elevation of the terrain heightmap. For the determination of the collision in the coordinate position (x, z), determined the value image with equation (2):

$$x_{local} = ceil\left(\frac{x}{4}\right) \qquad z_{local} = ceil\left(\frac{z}{4}\right) \tag{2}$$

Factor 4 used because the value has a value per pixel distance 4. Pixel distance value declared that there are many grids in one pixel. After value image dan and image acquired, determined in the index box is how the particle using equation (3):

$$indexbox = (x_{local} - 1) + ((z_{local} - 1) * (width - 1)) + (z_{local} - 1) \tag{3}$$

Figure 4. ***Top*** **Triangle** ***and Bottom Triangles on a Local Box.***

Once the particle's position is determined, then determining whether the particle's position in the triangle left or right triangle, the value determined in image using equation (4):

$$x_{temp} = x - (jarakperpiksel * (x_{local} - 1)) \qquad z_{temp} = z - (jarakperpiksel * (z_{local} - 1)) \tag{4}$$

Triangle top and bottom of the triangle bounded by the diagonal line defined by equation (5):

$$z = mx + c \tag{5}$$

If

$$x_{temp} > z_{temp}$$

then the particle's position in the left triangle, or

$$x_{temp} < z_{temp}$$

then the particle's position in the right triangle.

4.3 Lava Flow Simulation Using SPH

Smoothed Particle Hydrodynamics method used for modeling the interaction between particles. SPH model used is the model proposed by [Muller, 2005]. SPH modeling presented include determining the nearest neighbor, *kernel smoothing,* determining the force due to pressure and viscosity determination.

In Figure 5, to determine the value of *the kernel Smoothing* first determined the nearest neighbor and *kernel smoothing* function applies to the nearest neighbor particles. After *Smoothing kernel* function is determined, the next step is to calculate the value of density, viscosity, pressure and surface tension.

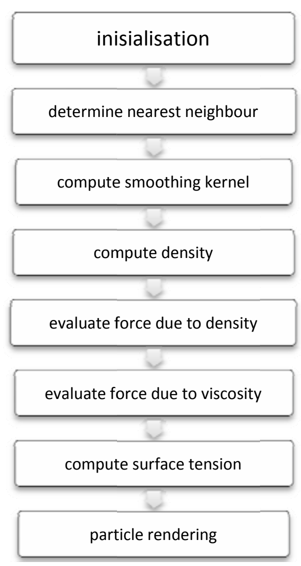


Figure 5. ***System Design of Smoothed Particle Hydrodynamics***

4.4 Determination of Nearest Neighbors

Data structure used for modeling of lava is staggered grids method developed by [Kipfer, 2006]. This data structure was developed for the system of particles with low density particles. Data structures using staggered grids more efficient in use memory and computing.

Linear list used to store the particles. Grid used to mapping virtual particles. Index bin (i_x, i_y, i_z) to be combined to make 64-bit identifier for each bin id = $|0|i_z|i_y|i_x|$, based particle sorting this identifier list. To determine the nearest neighbor particles, have been undertaken prior to that contained in the y-axis, suppose the radius of the kernel used is r and the coordinates of particle reference points to be determined neighbors are i (x, y, z) particle sorting and then performed in positions (x, y, z-r) to (x, y, z+r). Having determined the particle sorting in the area, then do the sorting of particles in the area (x, y-r, z) to (x, y+r, z). The last step sorting particles in the area (x-r, y, z) to (x+r, y, z). The final result obtained is a list of neighboring particles i (x, y, z) with the radius r. At last known position of any particle which includes the nearest neighbor of the reference particle i (x, y, z). These three steps are performed for every particle every time step simulation. The use of staggered grid approach, there are three simple online lists. In determining the nearest neighbors of each particle is checked neighboring particles and thus require large computing.

4.5 Smoothing Kernel

Stability, accuracy and speed of the SPH is used depends on the *kernel smoothing* is used. *Kernels* used in the modeling of this lava designed by [Muller, 2005]. The advantages possessed by this kernel is *r* squared value can be evaluated without calculating the roots when calculating the distance. If the kernel is used to calculate the pressure force, the particles will agglomerate with high pressure. A very close distance between particles create a repelling force is lost due to the gradient of the kernel to zero at the center. To overcome this problem is *spiky kernel* introduced by Desrun [Muller, 2005]. Spiky kernel will be to *generate* the repulsive force between particles, so particles do not accumulate in a cluster with a high pressure.

Viscosity is a phenomenon caused by the friction which reduces the fluid kinetic energy and convert it into heat. To calculate the viscosity force, used a third type of kernel was introduced by [Muller, 2005].

4.6 Force due to pressure and viscosity

Has a parameter of fluid pressure and viscosity. The pressure experienced by a particle is due to an accumulation of pressure neighboring particles. Particle amount of pressure effect depends on the distance of neighboring particles to the particle reference. Particles with the closest distance the most significant influence while the particles are not included in the list of nearest neighbor particles have no influence at all [Sariel, 2008]. The force experienced by a particle due to pressure expressed in equation (6).

$$f_i^{tekanan} = -\sum_j m_j \frac{p_i + p_j}{2\rho_j} \nabla W(r_i - r_j, h) \quad (6)$$

Viscosity depending on the force and the difference of speed. Force due to viscosity is defined by [Muller, 2005] as follows.

$$f_i^{viskositas} = \mu \sum_j m_j \frac{v_i + v_j}{\rho_j} \nabla^2 W(r_i - r_j, h) \quad (7)$$

4.7 Rendering Carpet

Construction of the carpet can be used as a regular *mesh* surrounding *terrain.* The approach is implemented in the construction of this carpet is the carpet virtual method. Carpets constructed and rendered only when the particle is in a *terrain. Quadtree* data structure used to identify the particle position efficiently. The method used for the construction of this rug is as follows [Kipfer, 2006] :

1. Particles stored in the leaf / child quadtree*;*
2. for each node, the maximum height values drawn *tree,* this value defines the virtual carpet surface;
3. visible part of the carpet experience through *tree rendering* depends fluid height above the *terrain.*

Carpet consists of a virtual square n x n, which is the resolution of the carpet during the *rendering* process. *Tree* node storing an absolute minimum *terrain* height of this region, namely the carpet and high-speed value. Then the kids do *tree* storage by storing the corresponding absolute height using the operating room max. Then the carpet is pushed to the right height above the particle. The next node in the *quadtree inner* propagated *updated* with the value of the height for each step down.

Rendering is done efficiently using a *quadtree* method, if the value of the particle height is found to be below the *terrain* elevation values, then the recursion is canceled and does not occur *rendering* carpet. When it comes to child nodes, can be part of the section to be experiencing *rendering* Carpet accelerated by the acceleration of gravity and curved position under the carpet. If the particle is displaced position, then the carpet will come up to the surface of the *terrain.* This stage is not experienced because the position of the particle rendering is no longer there. If the particles are still there, then the second stage of *quadtree* algorithm will restore the correct height value. Carpets can be update incremently, spatial data structures used in the SPH model has information that can be used to *update* this carpet.

Smooth surface can be made to minimize the local curvature so that the carpet will look rendered a separate or isolated. The spatial surface should not be connected because there is a separate particles from a collection of other particles.

In the development of terrain using a quadtree method, each node has 4 children, the node has no children is called a leaf.. The terrain is flat, which does not have a lot of detail, so that part quadtree node has no children. While the terrain that has a lot of detail, the section has a node in a quadtree data structure has a child in accordance with the depth terrain detail in that area.

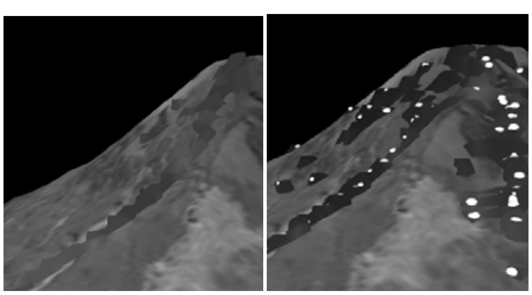
5. Result and Discussion

Specification compiler software is Microsoft. NET, OpenGL graphics utility for graphics applications. SPH computation initialization by setting the variable values as listed in Table 1.

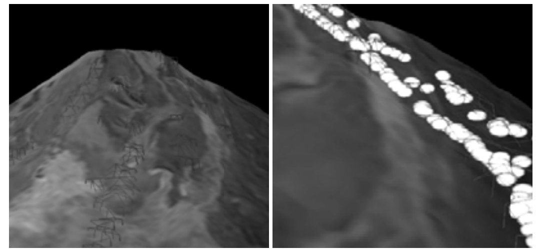
Table 1. *Initial Variable Values SPH Applications.*

Variabels	Value
The number of particles	1,000
Gravitational acceleration	-9.8
Sampling distance	0.5
Smoothing length	2.5
Initial density	2500
Particle radius	0.5
Particle volume	5

The simulation is run using an Intel Core 2 Duo 2.0 GH and graphics card NVidia GeForce 8600 grade of frame persecond (fps) is 100 when the number of particles is 39 fps. When the number of particles rendered 500, resulting fps 8 and drops to 3 fps at 1000 particles.



(a) **(b)**



(c) **(d)**

Figure 6. Simulation of Lava Flows Using SPH,
(a) Rendering of Lava Using Carpet,
(b) Rendering Carpet and Particles,
(c) Rendering Wireframe,
(d) Rendering Particles.

Figure 6 visualize lava flow with the number of particles 1000. Fps value decreased sharply as the number of particles increases because of high computing on the selection of neighbor particles. Lava flow simulation to produce visually realistic enough to describe the flow of lava on the slopes of the mountain.

References

[1] B. C. Vemuri, Y. Cao, Chen, L. (1998) Fast collision detection algorithms with applications to particle flow. Computer Graphics Forum, hal. 121–134.

[2] Brid,R., Muller-Fischer, M. (2007), Fluid Simulation, SIGGRAPH, Course Notes, Zürich.

[3] Crawford, Chris (2003), Chris Crawford on Game Design, New Riders, Indianapolis.

[4] Gomez , Miguel (1999), Simple Intersection Test for Games, Gamasutra, http://www.gamasutra.com/features/19991018/Gomez_1.htm, 26 Mei 2008, 9.35 WIB

[5] Har-Peled, Sariel (2008), Quadtrees – Hierarchical Grids, Creative Commons, San Francisco, California.

[6] Kipfer, P., Westermann, R. (2006), Realistic and Interactive Simulation of Rivers, Graphics Interface, Quebec.

[7] Liu, G.R., Liu, M.B. (2003), Smoothed Particle Hydrodynamics – a meshfree particle method, World Scientific, Singapore.

[8] Muller, Matthias, et.al. (2003), Particle-Based Fluid for Interactive Applications. Dept. Comp. Science, Federal Institute of Tech. Zurich (ETHZ), Zürich.

[9] Muller , M., Solenthaler, B., Keiser, R., Gross, M. (2005), Particle-Based Fluid-Fluid Interaction, Eurographics/ACM SIGGRAPH Symposium on Computer Animation, Zürich.

[10] Reiterer, Florian (2006), Particle Based Real Time Fluid, Institute of Computer Graphic and Algorithms, TU Vienna, Vienna.

[11] Roy, Trina, M. (1995), Physically Based Fluid Modelling using Smoothed Particle Hydrodynamics. Master Thesis, University Illinois at Chicago, Chicago.

[12] Samet,H. (1989), Spatial Data Structures: Quadtrees, Octrees, and other Hierarchical Methods, Addison-Wesley, Massachusetts.Gkj

[13] Sempe, L. (2002), Terrain Rendering using Heightmaps, http://www.spheregames.com, 20 Mei 2008, 09.30 WIB.Fj

[14] Snook, Greg (1996), Real-Time 3D Terrain Engines Using C++ and DirectX 9, Charles River Media, Hingham. Massachusetts, 119-152.

[15] Stahler ,Wendy (2004), Beginning Math and Physics for Game Programmers. New Riders, Indianapolis.Guk

[16] Stora, Dan et.al., Animating lava flows. In Graphics Interface, pages 203–210, 1999, Paris.Gfjk

[17] ________, Black Body Radiation, WIKIPEDIA, http://en.wikipedia.org/wiki/Talk:Black_body, 12 juli 2008, 10.10 WIB.

[18] ________, Differential Equations, Linear Equation Application. Lecturer Notes. http://www.math.stcc.edu/DiffEq, 2 Juli 2008, 15.35 WIB

[19] ________, Center of Vulcanology and Mitigation of Geology Disaster, Introduce to Vulcanology, Department of Energy and Mineral Resources Indonesia, http://merapi.vsi.esdm.go.id/vsi/, 6 Januari 2008, 8.10 WIB.

ANALYSIS AND DESIGN of E-PROCUREMENT PT. ADHI JAYA

Suparto Darudiato

Department of Information System, Faculty of Computer Science, Binus University
supartod@binus.edu

Abstract

PT Adhi Jaya is an enterprise in the paper industry. To provide raw materials, it relies excessively on the suppliers which already signed mutual agreement with; therefore, it would not compare them with others to find the best ones. In order to solve this matter, an analysis and design of a web-based procurement information system, to request, to order, up to receive the ordered material, must be conducted. E-procurement information system is analyzed and designed using object-oriented approach and UML notation.

Keywords: Analysis, Design, Information, E-Procurement

1. Introduction

Lately, the development of an enterprise relies on information technology which can support business activities. By employing an exact IT, the company can achieve competitive advantages to compete within industry.

Applying IT brings hope to stand company's existence in business relationship with other companies, including with suppliers and customers. This encourages companies to compete to be the most successful company in terms of the relationship quality betterment with suppliers and customers by fulfilling their demands. The company that already applied it, has earned enormous benefit, such as accelerating all manual activities and lessening the expenses. Therefore, the company can do its business processes effectively and efficiently.

E-procurement is an example of IT that administers materials supply. To meet materials demand, the company will contact suppliers. In manual process, these activities take a long chain and need many expenses. This can affect to production process. Therefore it is better to make this process faster and more efficient. E-procurement is the exact solution to ease, accelerate, and make more efficient material supply process. E-procurement brings additional value to the company, is improve the relationship quality with other companies, in this case suppliers, to agree a better mutual consent which matched with characteristic and capabilities of each other

2. Electronic Business (E-Business)

E-Business is the changing from manual operation to electronic devices-assisted one. According to [5], there are many forms of electronic media that have been used, such as EDI (Electronic Data Interchange), e-mail, EFT (Electronic Funds Transfer), electronic publishing, image processing, electronic bulletin boards, shared databases, bar coding, fax, automated voice mail, CD-ROM catalogues, the internet, and website.

The benefit of using e-business:

a. To diminish even to wipe out the role of intermediaries, retailers, and service provider, in order to cut off the expense
b. More choices and information to customers
c. To reduce time
d. To improve service
e. To collect and analyze customer data and demand
f. To construct a virtual company that requires less resource to sell more
g. To access the global market, suppliers, and distribution channels

3. E-Procurement

E-procurement is the web-based integrated purchasing activities between a buyer and a supplier, covering from requesting, purchasing, delivering, until paying, by using internet. E-procurement is a form of e-commerce to provide goods and services to companies. According to [3], e-procurement is an activity using internet to create business opportunities to widen the connected network. The company and its suppliers are not just partners which connected via traditional EDI network, but using web to co-operate with other business entities which connects to internet.

There are many positive values that are obtained by applying e-procurement in the purchasing process. They are:

a. The administrative process can be done faster, more accurate, and lesser cost.

b. In a auction system, the bidding process can be done openly, so to be able to choose more competent suppliers.
c. The company and its suppliers can trace all transaction faster.

All of them happen in a life cycle and sequence chain of e-procurement as said by [2] and shown figure 1 above.

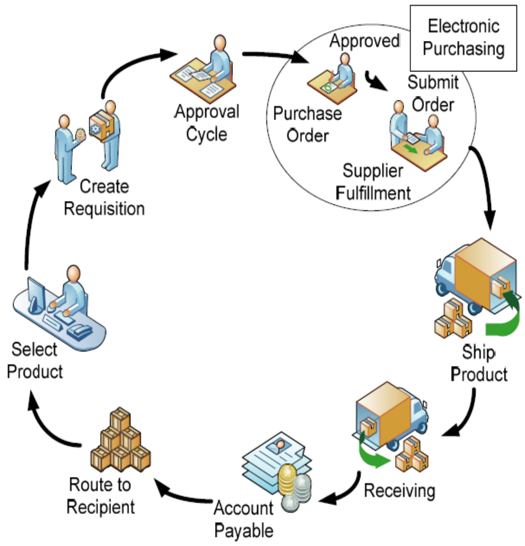


Figure 1. life cycle and sequence chain of e-procurement [2]

The procurement process emphasizes three separated work flow, ordering, fulfilling, and paying workflow as said [2] and shown Figure 2.

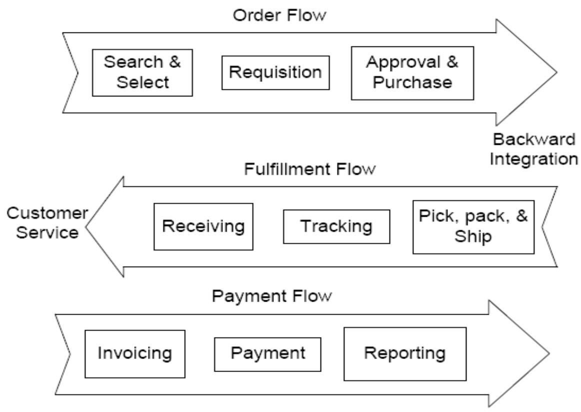


Figure 2. Three critical process flows [2]

4. The Proposed Business Processed

The new system (shown Figure 3) can do so many automations to save more resources for instance time, expense, and energy. It will use web-based technology that can run online. By employing it, production staff and manager, purchasing staff and manager, also warehouse staff can check materials data, material requests, purchasing request, purchase order, price proposal, returned goods, receiving receipt, employee (except username and password of other employees) and suppliers data at PT Adhi Jaya e-procurement websites, but need to login first.

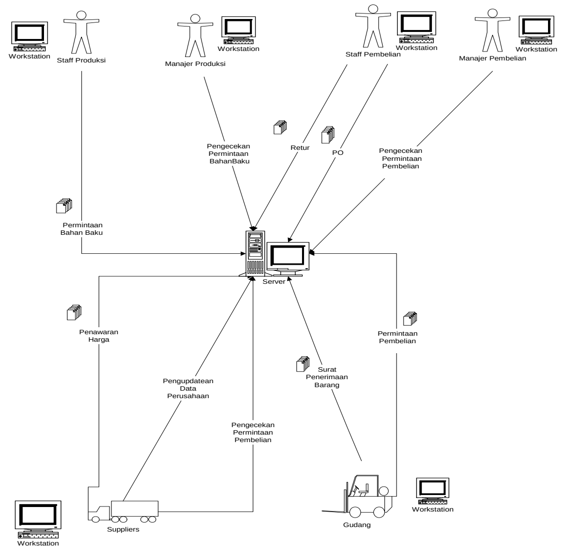


Figure 3. Rich picture of E-Procurement PT. Adhi Jaya

Nevertheless, the authorized party to do update material requests is production manager, purchasing request status and price proposal is purchasing manager, PO and return status is purchasing staff, receiving receipt status is warehouse staff. Staff of each department can update data that are not in his department boundary.

The enterprises which will be one of PT Adhi Jaya suppliers, can register on the websites by filling the required data thoroughly. Purchasing staff are to validate and clarify the data. If it is true, purchasing manager will decide which companies can supply. Then, the system is to send username and password login as supplier to the chosen company.

Suppliers can notice the demanded materials list (purchasing request) on e-procurement websites. It can send a proposal by clicking "*Buat Penawaran Harga*" button on "Permintaan Pembelian" form and filling the form. Price proposal is done within tender process. A supplier is not allowed to offer more than one price proposal for the same purchasing request, but is allowed to do some changes to the submitted data before closing date. By clicking the button, the suppliers is considered to consent all the requirements, chapters, and clauses of the agreement which aforementioned on the websites.

Additional price proposal can be submitted during the given time period. When it comes to end, No proposal is received and purchasing staff will compare the proposals one to another based on certain criteria, then choose some of the bests. Manager purchasing will decide the best among them. The system will generate and send PO to the chosen supplier.

When the supplier deliver materials to the warehouse, warehouse staff validate the data on the delivery document against the data on PO, by logging in to websites. If it's correct, warehouse staff will compose receiving receipt and confirm checked status of receiving receipt. Otherwise, the status will be "pending" and there will be a note, recorded by purchasing staff. The note will send to the supplier to ask for replacement. If it's fulfilled, warehouse staff will compose a new receipt and confirm it to purchasing staff. Purchasing staff is to update the "pending" status to "success" status of the return goods. Automatically, the status of receiving receipt will change to be "checked".

.

5. Usage

Usage is utilized to determine how an actor interacts with the system. It is visualized in a use case diagram. The result of this process is a description of all use cases and actors. According to [1] use case is the communication between an actor and the processes within the system boundary.

Actor is an abstraction of user or other system that interacts directly with the system; meanwhile use case is an interaction pattern between the system and an actor.

Use case diagram is a diagram that shows a collection of use cases and actors, also the existing interaction. So, use case diagram is used by developers to gain users expectations of the designated system. Use case diagram e-procurement PT Adhi Jaya is shown figure 4.

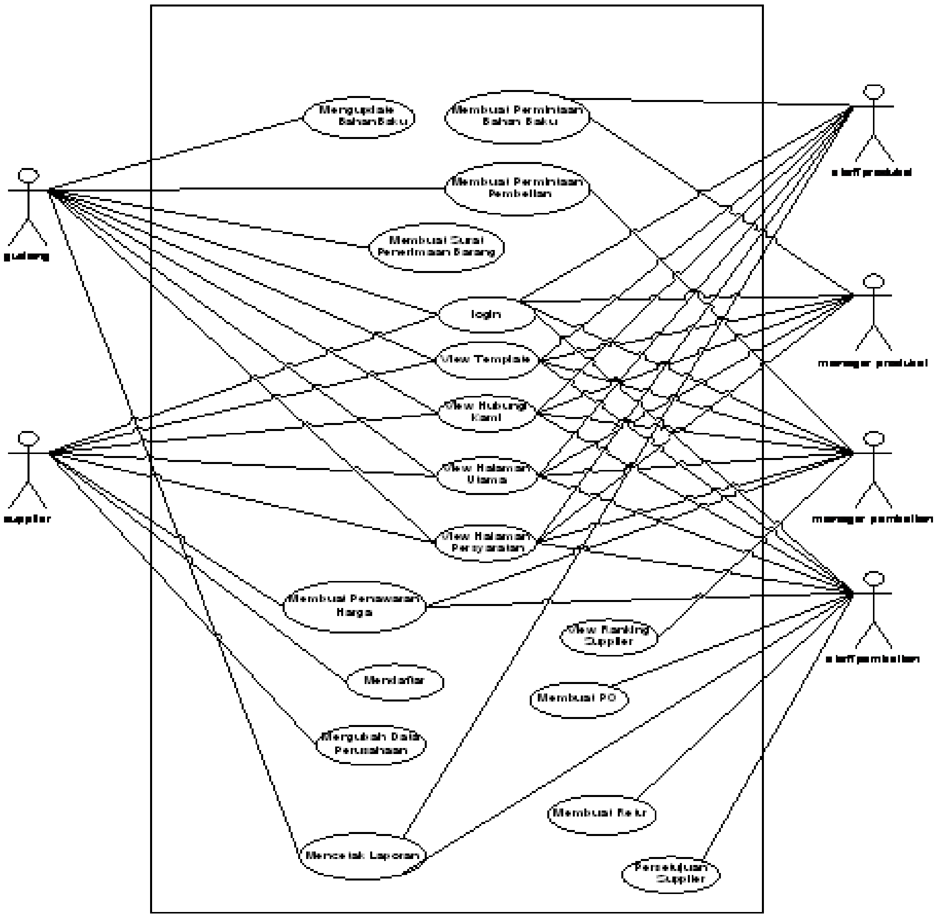


Figure 4. Use Case Diagram of E-Procurement PT. Adhi Jaya

6. Classes

Class is a group of objects that have the same structural relationship, behavior pattern, and attributes; meanwhile object is an entity instance which has its own identity, state, and behavior. Class Diagram is a diagram that conceptualizes the associated problem domain, by showing structural relationship among classes and objects in model.

The relationships in class diagram are labeled structure. The structures In class diagram are classified association, generalization, and aggregation. Class diagram of e-procurement system PT Adhi Jaya is shown Figure 5.

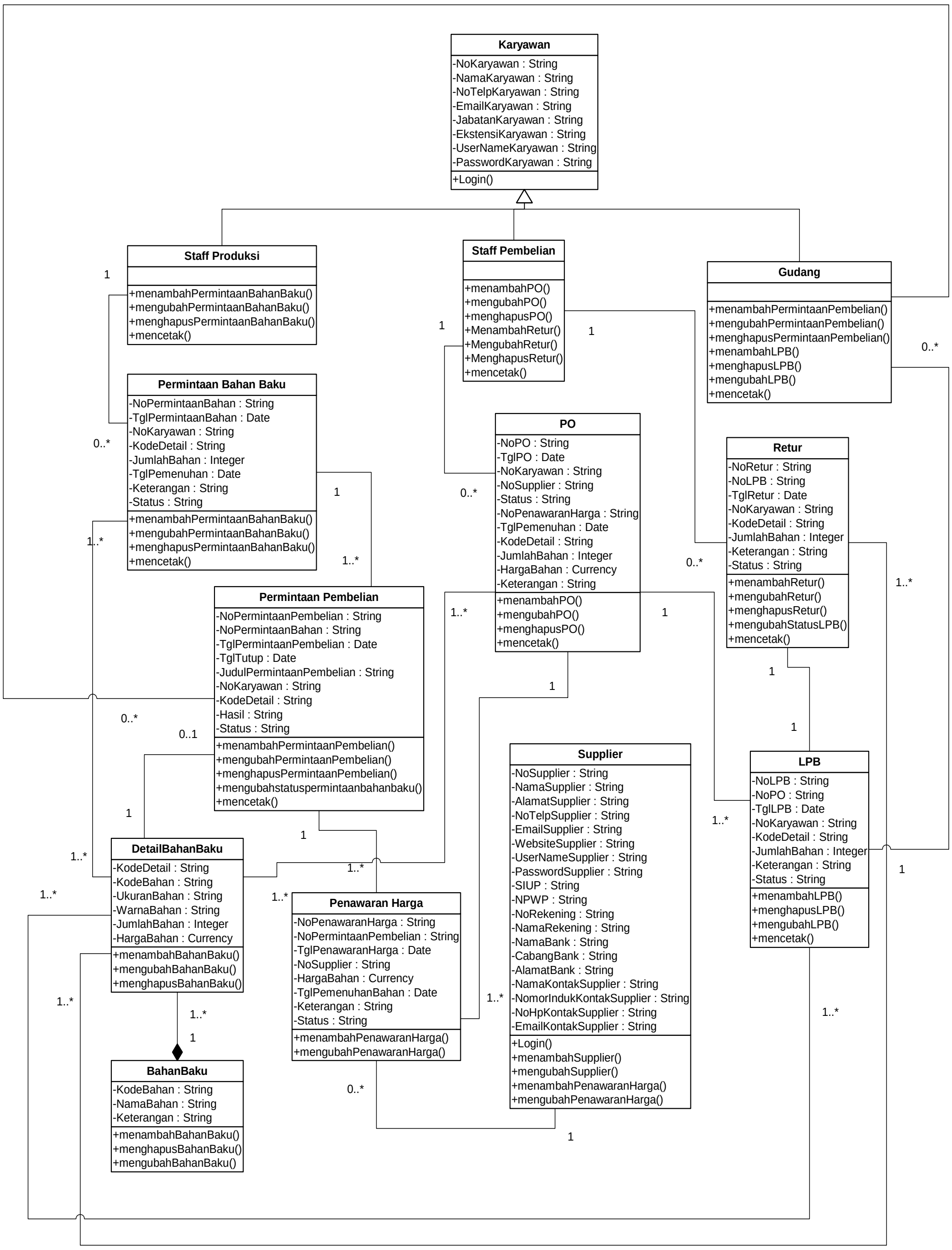


Figure 5. Class Diagram of E-Procurement PT. Adhi Jaya

7. User Interface

When designing the user interface, the issues that have been considered is dialog features such as feedback, escape, help, error, wizards, suggestions, and shortcuts because they are important in order to assist users interact with the system. The quality of general user interface is named usability (reference to [4]). Figure 6 to 9 are examples of user interface of e-procurement system PT Adhi Jaya.

Figure 6. Home Page.

Figure 7. Purchase Request Page

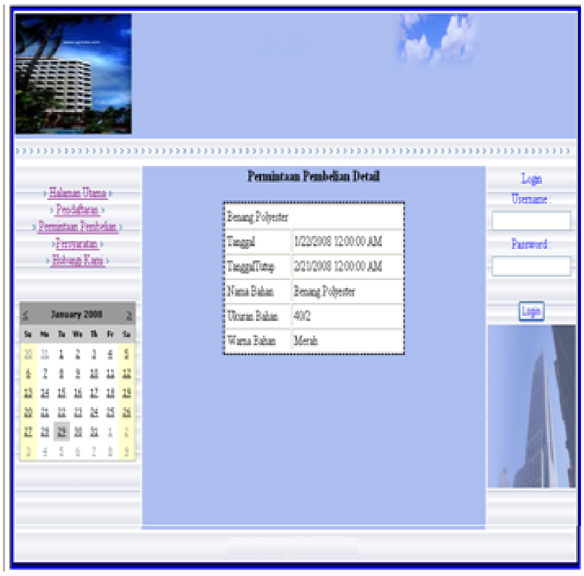


Figure 8. Purchase Request Details page

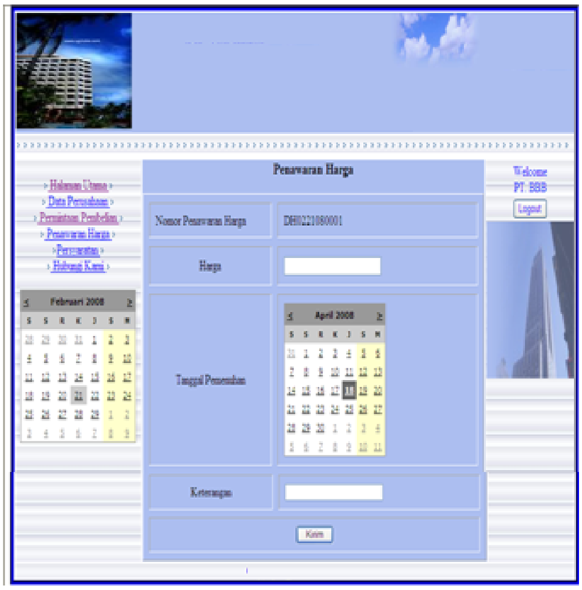


Figure 9. Makes Offer Price Page

8. Architecture Process

Architecture process envisages the physical form of the system. There are two abstraction levels in these process activities. Level one is describing all the available program components in the system processors. Level two is the collaboration process of the object structure. Architecture process e-procurement system is shown picture 10.

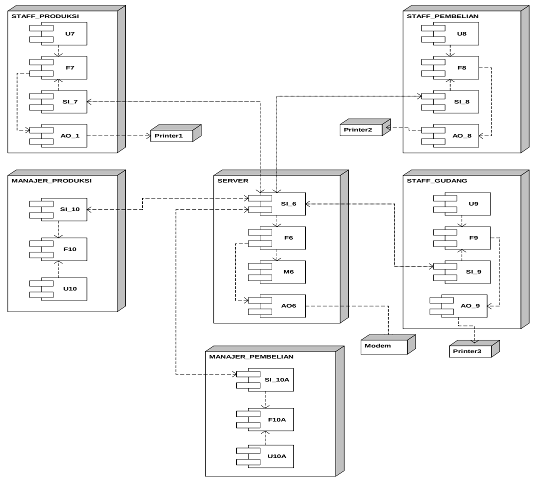


Figure 10. Deployment Diagram of E-Procurement

6. Computer Networking

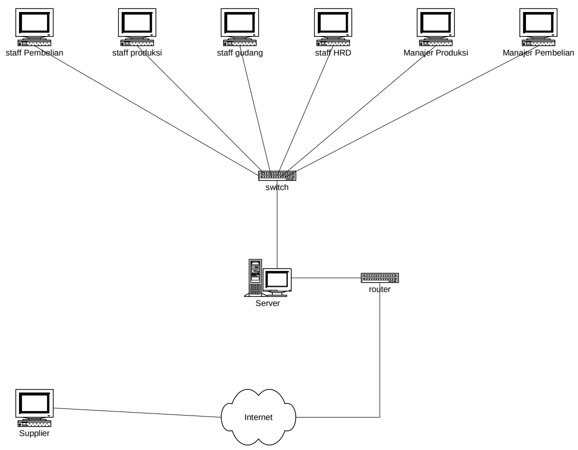


Figure 11. Computer Networking

7. Conclusion

There are some conclusions about the analysis and design e-procurement PT Adhi Jaya:

a. To ease purchasing staff job desk to inform suppliers about demanded materials.
b. To facilitate suppliers, in this case the ones those have been registered and connected to the internet, including the interested companies to be the supplier candidates, in order to be able to submit a proposal of PT Adhi Jaya' purchasing request.
c. To assist in choosing the best suppliers based on their submitted price proposal and the given criteria.
d. To speed up purchasing staff task to send PO as a confirmation to the chosen suppliers to finish administrative activities, such as to specify the agreement, purchasing approval, etc.
e. To aid purchasing staff to return goods to suppliers as a confirmation there is inappropriate delivery.
f. To assist purchasing staff to send username and password to the chosen suppliers.
g. To facilitate warehouse staff to control material usage by registering materials, checking the quantity, observing in-and-out goods, etc which still in warehousing boundary.

8. Recommendation

To reach a good sustainability, some enhancements and functions addition are required:

a. Supplier selection can be done automatically based on the given criteria.
b. The feasibility analysis to suppliers can be done automatically and the result can be served to the interested parties.
c. The system needs to be complemented with supply methods, such as EOQ, ROP, and minimum stock.
d. The system is required to provide information about fast-moving and slow-moving goods.

References:

[1] Bahrami, Ali.(2001). *Object-Oriented Systems Development: using the unified modeling language*. McGraw-Hill, Singapore.

[2] Kalakota, Ravi and Robinson, Marcia.(2004). *E-Business 2.0:Roadmap for Success*. Pearson Education, New Jersey.

[3] Laudon, Kenneth C, Jane P. (2008). *Management information systems : managing the digital firm. 9th Edition*. Penerbit : Sine Nomine. Publisher : Mc Graw-Hill Companies.

[4] Mathiassen, Lars.(2000). *Object Oriented Analysis & Design. Marko Publishing ApS*, Denmark.

[5] Russell, Roberta S. and Taylor III, Bernard W.(2003). *Operations Management, 4th edition*. Pearson Education, New Jersey.

ANALYSIS OF PROPAGATION CHANNEL USING WALFISCH-IKEGAMI MODEL ON MOBILE WiMAX SYSTEM

Andri Mulyaman[1], Uke Kurniawan Usman[2], Saleh Dwi Mardiyanto[3]

[1,2,3] Institut Teknologi Telkom

[1]andrimulyaman@yahoo.co.id, [2]uku@ittelkom.ac.id, [3]sdm@ittelkom.ac.id

Abstract

The IEEE 802.16e standard or known as Mobile WiMAX is an improvement of previous WiMAX standard to support user mobility. Mobile WiMAX conforms the capability of portable and mobile applications which optimize the previous standard performance to resolve dynamics of mobile wireless channel. To resolve dynamic mobile wireless channel, an accurate channel model is required so that Mobile WiMAX's network could be well planned, specifically for analyzing influence of problems because of multipath, on the performance of Mobile WiMAX. That problem is multipath fading such as Doppler shift. In this paper, the writers learn about propagation channel models for Mobile WiMAX system using one of empiric propagation models, the COST 231 Walfisch-Ikegami model. This model is believed to be accurate to describe mobile wireless in urban area because it calculates parameters such as building height, distance between transmitter and receiver, and width of street in that area. The simulation results show that the best performance for Mobile WiMAX system is achieved when LOS distance is one kilometer and the velocity is 0 km/h. On that condition the pathloss is 110.56 dB. To reach BER 10^{-3} it needs SNR 8.86 dB. In NLOS condition at 50 km/h and 120 km/h, mobile WiMAX has very bad performance, where BER 10^{-3} cannot be achieved.

Keyword: Mobile WiMAX, Walfisch Ikegami, Path loss, Propagation Channel, Doppler Shift, Bit Error Rate

1. Introduction

The formation of IEEE 802.16-2004 that supports fixed and nomadic access service shows the standard's development to mobility orientation until the formation of IEEE 802.16e standard which is also known as Mobile WiMAX. IEEE 802.16e standard conforms the capability for portable and mobile application which is an optimalization of performance from the previous standard so it can resolve the dynamics of mobile channel.

In mobile wireless propagation, fading is the main characteristic. There are 2 kinds of fading, which are small scale fading and large scale fading. Therefore, a propagation model must be designed to describe and analyze the influence of the characteristic mentioned.

Walfisch-Ikegami model is the most accurate empiric propagation channel model to describe mobile wireless propagation in the urban area because it considers some parameters such as : the building height between transmitter and receiver, the distance between transmitter and receiver, the roads width in the urban area, etc.

2. Related Theory

2.1. WiMAX

In December 2005, IEEE released IEEE 802.16e or usually known as mobile WiMAX which is an amendement of the previous standard to support mobile application. IEEE 802.16e standard is not fully standardized. For some research this standard still adopts the previous standard which is 802.16d. Generally, mobile WiMAX will be developed to support both fixed and mobile application on licensed band between 2-6 GHz for various condition, whether LOS or NLOS. Mobile WiMAX also support portable application.

The 802.16e amendment includes Physical Layer and MAC Layer to combine fixed and mobile services in a licensed band. It is to upgrade the IEEE 802.16d standard so it can support subscriber station mobility on the vehicular velocity and to combine fixed and mobile broadband wireless access. The functional higher layer is the handoff mechanism between base stations.

2.2. OFDM and OFDMA

OFDM (Orthogonal Frequency Division Multiplexing) is a transmission technique which uses several orthogonal frequencies.

In the OFDM system, input data flow is divided into some paralel data flow with lower data rate (increasing symbol duration) and each paralel data flow mentioned is modulated and transmitted through separated orthogonal subcarriers.

Cyclic Prefic (CP) holds an important role to maintain the orthogonality of the OFDM subcarrier in the channel condition with selective frequency. CP is sequence of bits formed by copying some bits of OFDM symbol, and putting the bits in front of the mentioned symbol. With CP addition, the OFDM signal will not experience ISI as long as the channel's delay spread is shorter than the CP duration. ISI just gives influence to the CP part of the symbol, meanwhile the OFDM payload data will not experience distortion because of ISI.

2.3. OFDMA Symbol and Subchannelization

OFDMA symbol structure is divided into three types of subcarriers, which are :

- Data subcarrier, for data transmission.
- Pilot subcarrier, for estimation and syncronization.
- Null subcarrier (is not used for transmission), is used as guard band and DC carrier.

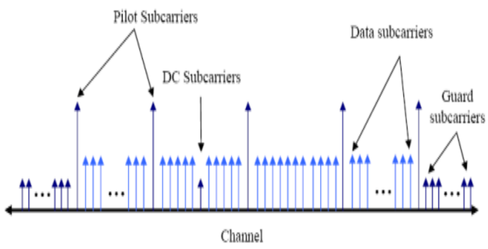


Figure 1. OFDM Subcarrier structure

There are 2 types of subcarrier permutations in subchannelization, which are: diversity and contiguous. Diversity permutation fetchs the subcarrier in pseudo-random way to form a subchannel. That permutation gives the diversity in frequency and averages the inter-cell interference. Several kinds of diversity permutations include: DL FUSC (Fully Used Subcarrier), DL PUSC (Partially Used Subcarrier), UL PUSC, and several additional permutation. On the other hand contiguous permutation groups the neighbourhood subcarrier into a subchannel. The examples of this permutation are: DL and UL AMC (Adaptive Modulation and Coding).

Table 1. DL FUSC Mobile WiMAX Parameters

Parameters	Values				
System Bandwidth	1,25	2,5	5	10	20
FFT Size	128	256	512	1024	2048
Gurad Subcarriers	22	N/A	86	173	345
Subcarriers	106	N/A	426	851	1703
Data Subcarriers	96	N/A	384	768	1536
Pilot Subcarriers	9	N/A	42	83	166
Subchannel	2	4	8	16	32

2.4. Scalable OFDMA

Scalability is conducted by adjusting the amount of FFT point by maintaining the distance between subcarrier in the amount of 10.94 kHz. Because the bandwidth subcarrier and symbol duration are permanent, the effect to the upper layer becomes minimum when the bandwidth is being scaled. There are 2 bandwidth profiles being developed for scalable OFDMA. They are 5 and 10 MHz.

Table 2. Scalable OFDMA Parameters[16]

Parameters	Values	
System Channel Bandwidth (MHz)	5	10
Sampling Frequency (F_p in MHz)	5,6	11,2
FFT Size (N_{FFT})	512	1024
Number of Sub-Channels	8	16
Subcarrier Frequency Spacing	10,94 kHz	
Useful Symbol Time (T_b=1/f)	91,4 microsecond	
Guard Time (T_g=T_b/8)	11.4 microsecond	
OFDMA Symbol Duration (T_s=T_b+T_g)	102.9 microsecond	
Number of OFDMA Symbols	48	

2.5. Doppler Shift

A frequency shifting can be calculated using the following equation :

$$\Delta f = (f\, v \cos\theta)/c = v/\lambda \qquad (1)$$

where,
f = carrier frequency
v = user velocity
θ = signal arrival angle relatively towards the direction of user velocity

The maximum frequency shifting occurs when the moment of signal arrival angle is 0^0, which value could be positive or negative so that the maximum value of spectral extent is twice as the maximum frequency shifting (f_m).

$$f_m = \frac{f\, v}{c} \qquad (2)$$

Doppler spread determines the channel coherence time which means the statistical measurement from time duration where the response impuls from the channel could still be categorized as invariant and quantifying the channel response similiarity at different time. The value of channel coherence time could be estimated using the equation below.

$$T_c = \frac{9}{16\pi f_m}, \text{ for 50\% time correlation} \qquad (3a)$$

$$T_c = \sqrt{\frac{9}{16\pi f_m^2}}, \text{ for geometrical approach} \qquad (3b)$$

2.6. Propagation Channel Model

Propagation channel model is a channel characteristic estimation model. Generally, propagation channel model can be separated into 3 major categories which are :

- Empirical model, a model which is gained from the field measurement in the locations considered as the representation of the wireless environment sample.
- Deterministic model, a model which is developed from the electromagnetic pulse propagation theory and is used to calculate the transmission in the observed location.
- Sthocastic model, a model which includes random variable as the representation of the enviroment condition which always changes from time to time and from one location another.

Amongst all three propagation channel models mentioned above, empiric propagation channel model is the most preferable model for researcher or industry because it is easy to use and its tolerance toward the inavailability of terrain information. The empiric propagation channel model includes : Okumura-Hatta model, Walfisch-Ikegami model, Lee model and Longley-Rice model. Amongst four models mentioned above Walfisch-Ikegami model is the fittest model for the urban area.

2.7. Walfisch-Ikegami Model

This empirical model is a combination of J.Walfisch and F.Ikegami model which has been developed by EURO-COST on COST 231 project.

In the calculation, this model only considers the transmission strip linearly on the vertical field between transmitter and receiver. So the the focus points only to the subjects that cross the transmission strip. In turban area wich has many buildings, the buildings which are calculated are those which are passed by the vertical field of transmission strip.

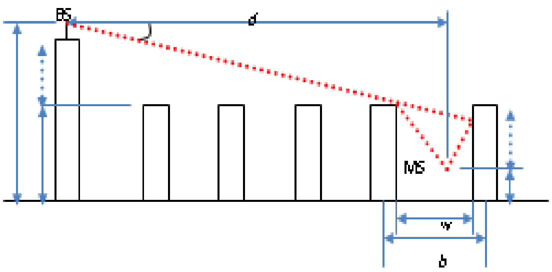


Figure 2. COST 231 Walfisch-Ikegami Parameter[9]

A transmitter on the building roof with h_b height transmits waves with f frequency so that the receiver across the building can receive the signal. The parameters which are obtained from the building are :

- The average value of road width (w)
- The average value of the distance beetwen building (b)
- The average value of building height (h_r)

The main parameters of the Walfisch-Ikegami COST 231 model are :

- Frequency (f) = 800 – 2000 MHz
- The transmitter height (h_b) = 4 – 50 m
- The receiver height (h_m) = 1 – 3 m
- The distance between transmitter and receiver (d) = 0.02–5 kms

Walfisch-Ikegami also makes the comparison and differentiates the two differ situations, which are : LOS (Line Of Sight) and NLOS (Non Line Of Sight). The LOS spreading is a direct spreading between transmitter (BS) and receiver (MS). When the LOS situation happens, the function which is used on the prediction using this model is very simple. Matemathicaly the function is represented as :

$$L_p = 42.6 + 26\log d + 20\log f \qquad (4)$$

The equation on the NLOS situation is more complicated. The total losses from this NLOS case is the summation of free space loss (L_0), multiple diffraction loss (L_{msd}), and rooftop-to-street diffraction loss/losses because of diffraction from the roof - road(L_{rts}).

$$L_p = \begin{cases} L_0 + L_{rts} + L_{msd} & ; L_{rts} + L_{msd} > 0 \\ L_0 & ; L_{rts} + L_{msd} \le 0 \end{cases} \qquad (5)$$

$$L_0 = 32.4 + 20\log d + 20\log f \qquad (6)$$

$$L_{rts} = -16.9 - 10\log w + 10\log f_c + 20\log(h_r - h_m) + L_{ori} \; ; h_r > h_m \qquad (7)$$

$$L_{ori} = \begin{cases} -10 + 0.354\varphi & ; 0^\circ \le \varphi < 35^\circ \\ 2.5 + 0.075(\varphi - 35) & ; 35^\circ \le \varphi < 55^\circ \\ 4.0 - 0.114(\varphi - 55) & ; 55^\circ \le \varphi < 90^\circ \end{cases} \qquad (8)$$

$$L_{msd} = L_{bsh} + k_a + k_d \log d + k_f \log f - 9\log b \qquad (9)$$

$$L_{bsh} = \begin{cases} -10\log(1 + h_b - h_r) & ; h_b > h_r \\ 0 & ; h_b \le h_r \end{cases} \qquad (10)$$

$$k_a = \begin{cases} 54 & ; h_b > h_r \\ 54 - 0.8(h_b - h_r) & ; d \ge 0.5km, h_b \le h_r \\ 54 - 0.8(h_b - h_r)\dfrac{d}{0.5} & ; d < 0.5km, h_b \le h_r \end{cases} \qquad (11)$$

$$k_d = \begin{cases} 18 & ; h_b > h_r \\ 18 - 15\left(\frac{h_b - h_r}{h_r}\right) & ; h_b \le h_r \end{cases} \qquad (12)$$

$$k_f = -4 + \begin{cases} 0.7\left(\frac{f}{925} - 1\right) & \text{; Medium city and suburban centre} \\ 1.5\left(\frac{f}{925} - 1\right) & \text{; Metropolis centre} \end{cases} \qquad (13)$$

3. System Modeling and Simulation

3.1. Simulation Parameters

Mobile WiMAX OFDM parameters that are used in this research have been adjusted with OFDMA scalable parameters.

Table 3 OFDM Parameters on the simulation[16]

Carrier Frequency (f_c)	2,5 GHz
Modulation	QPSK CC code rate 2/3
Channel Bandwidth	5 MHz
Sampling frequency	5,6 MHz
FFT Size (N_{FFT})	512
Amount of symbol per Subcarriers	48
Symbol period	91,4 µs ($T_b = 1/f$)
Guard time	11,4 µs ($T_g = T_b/8$)
Symbol duration	102,9 µs ($T_s = T_b + T_g$)

The parameters of Walfisch-Ikegami model which are used on the simulation are made based on the scenario below :

Table 4. Walfisch-Ikegami Parameters

Carrier Frequency (f_c)	$2.5 \cdot 10^3$ MHz
Transmitter height(h_b)	50 m
Building height(h_r)	30 m
Receiverheight (h_m)	2 m
Road width(w)	12 m
Width between building(b)	20 m
Signal arrival angle (φ)	45°

3.2. Simulation

The simulation is designed to be implemented by Matlab programming language. The simulation is to obtain and compare the BER (Bit Error Rate) curves towards various SNR (Signal to Noise Ratio) values.

Simulation can be divided into 3 main parts, which are :

- Channel coding and modulation or mapping.
- OFDM
- Propagation channel model COST 231 Walfisch-Ikegami which is the core of this research.

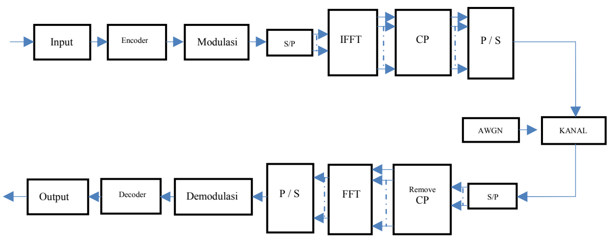


Figure 3. Simulation Model

a. Channel Encoder and Modulation

Before entering channel encoder block, data with 2304 bits is generated by using *randint* Matlab function. Then it is randomized using randomization block to prevent long zero-bit.

Mobile WiMAX system uses concatenated FEC (Forward Error Controller) type which consists of Reed-Solomon encoder and Convolutional Code encoder.

Table 5. FEC QPSK Parameters on Mobile WiMAX

Modulation	Uncoded Block Size	Coded Block Size	Overall Coding Rate	RS code rate (N, K, T)	CC code rate
QPSK ½	24 byte	48 byte	½	(32, 24, 4)	2/3

Before the modulation or mapping process, interleaver is held beforehand with the following parameters :

- N_{cpc} = 2 (for QPSK modulation)
- N_{cbps} = 384 (using 16 channel)

b. OFDM Transmitter

After being modulated, the data is converted into parallel. IFFT block in the OFDM system is functioned to generate orthonogonal carrier frequencies. Then, the data is inserted with cyclic prefic or addition of guard interval to reduce the ISI (Inter Symbol Interference) effect.

c. Channel

Channel is made based on COST 231 Walfisch-Ikegami model parameters scenario. In this channel block calculation of pathloss, Doppler shift, and other Walfisch-Ikegami parameters are conducted for analysis purpose.

d. AWGN

This block gives noise on system that is Gaussian distributed, then output of AWGN block will be mixed into transmitter output signal, which is already distorted by Walfisch Ikegami's channel.

e. OFDM Receiver

Guard interval added on the carrier side will be taken out so that the original sent symbol is obtained. FFT blok is functioned as the local oscillator on the receiver which separates the carrier frequency with different OFDM symbols in that frequency. The amount of FFT point is equal to IFFT point. OFDM symbols which are still on parallel data flow form are converted into serial data flow form.

f. Channel Decoder and Demodulation

In this block the process is the reverse of the channel encoder process. The data experiences demapping or demodulation process to be converted into binary data. Then, channel decoder process is conducted.

The binary datas from the decoder are compared with the input data to obtain the BER value.

3.3. Simulation Scheme

The schemes on this simulation are :

- Simulation for LOS and NLOS condition
- Simulation for various user velocity
- Simulation for various transmitter and receiver distance

The calculation for pathloss value on the LOS and NLOS is conducted with the following equations :

- **LOS Condition (d=1000 m)**

$$L_p = 42.6 + 26 \log d + 20 \log f$$
$$L_p = 42.6 + 0 + 67.96$$
$$L_p = 110.56\ dB$$

- **NLOS Condition (d=1000 m)**

$$L_p = 32.4 + 20 \log d + 20 \log f$$
$$L_p = 100.4\ dB$$

$$L_{ori} = 2.5 + 0.075(\varphi - 35)$$
$$L_{ori} = 3.25\ dB$$
$$L_{rts} = 38.48\ dB$$
$$k_a=54\ ;k_d=18\ ;\ k_f = -1.45$$
$$L_{bsh} = -18\log(1 + h_b - h_r)$$
$$L_{bsh} = -23.8\ dB$$
$$L_{msd} = -23.8 + 54 + 0 - 4.93 - 11.71$$
$$L_{msd} = 13.58\ dB$$
$$L_p = L_0 + L_{rts} + L_{msd}$$
$$L_p = 100.4 + 38.48 + 13.58$$
$$L_p = 152.46\ dB$$

Then the same calculation is conducted for 2000m and 5000m distance, so that the pathloss value are obtained as shown on table 5 and 6 below.

Table 5. Pathloss value for LOS condition

Distance (d)	Pathloss (L_p)
1000 m	110,56 dB
2000 m	118,38 dB
5000 m	128,73 dB

Table 6. Pathloss value for NLOS condition

Parameter	Value		
Transmitter-Receiver Distance (d)	1000 m	2000 m	5000 m
Free space loss (L_0)	100,40 dB	106,42 dB	114,38 dB
Multiple diffraction loss (L_{msd})	13,58 dB	18,99 dB	26,16 dB
Rooftop-to-street diffraction loss (L_{rts})	38,48 dB	38,48 dB	38,48 dB
Pathloss total (L_p)	152,46 dB	163,89 dB	179,02 dB

Pathloss calculation on COST 231 Walfisch-Ikegami can be represented as path gain value, where path gain(dB) is defined as the negative of pathloss(dB): P_G=-P_L[5].

As for the calculation to obtain dopplers frequency value is shown below:

- for v = 120 kms/hour

$$f_d = \frac{2.5\cdot 10^9 \cdot \left(\frac{(120\cdot 1000)}{3600}\right)}{3\cdot 10^8} = 277.78\ Hz$$

Then the same calculation is conducted for several velocity (0 km/hour, 3 kms/hour, 15 kms/hour, and 50 kms/hour). So the Dopplers frequency are obtained as shown in Table 7.

Table 7. Dopplers Frequency

Status	Velocity (kms/hour)	Dopplers frequency (Hz)
Static	zero	zero
Pedestrian	3	6,95
Low speed moving user	15	34,73
Medium speed moving user	50	115,74
High speed moving user	120	277,78

4. Simulation Result and Analysis

4.1. The Comparison of Various Distance

This subchapter observes the effect of distance of user to transmitter towards the mobile WiMAX system performance for several distances, which are : 1000 m, 2000 m, and 5000 m, where the user is on static condition (v = 0 km/jam).

- **LOS condition**

On LOS (Line Of Sight) condition it is found that the building has no effect, so the effect only comes from the distance of user to transmitter.

The pathloss value on LOS condition is only affected by user distance parameter, not by loss as a result of diffraction, because in the calculation rooftop-to-street diffraction loss (L_{rts}) and Multiple diffraction loss (L_{msd}) are unaccounted.

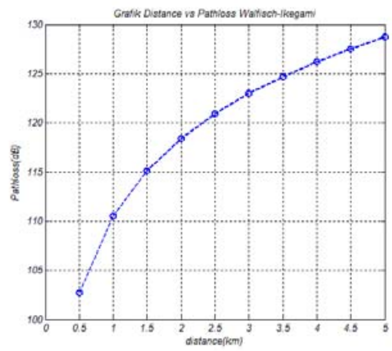


Figure 4. Pathloss graphic on LOS condition

Figure 5 below is simulation result of mobile WiMAX system performance on the LOS condition for various user distance :

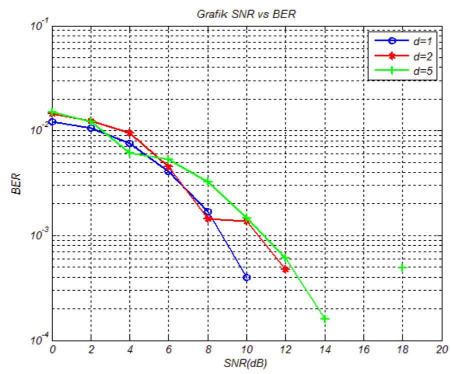


Figure 5. SNR-BER graphic for various distance on LOS condition

Figure above shows the performance of mobile WiMAX system on LOS condition for static (idle) user in the distance of 1000 m, 2000 m, and 5000 m.

Table 8. System Performance on LOS Condition for Various Distance

Distance (m)	1000	2000	5000
Pathloss (dB)	110,56	118,38	128,73
SNR (dB) with BER 10^{-3}	8,86	10,89	11,12

The best mobile WiMAX performance is achieved at the distance 1000m with pathloss 110.56 dB. BER target (10^{-3}) is reached at SNR value 8.86 dB, meanwhile on the distance of 2000m with pathloss 118.38dB, BER target is reached at SNR value 10.89 dB and for the distance of 5000m with pathloss 128.73 dB, BER target is reached at SNR value 11.12 dB.

- **NLOS Condition**

At the NLOS condition, mobile WiMAX system performance is not only affected by user distance, but also by the buildings between user and transmitter so the pathloss value is gained as the result of diffraction, especially the diffraction from the building rooftops.

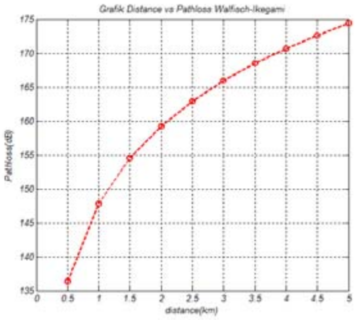


Figure 6. Pathloss Graphic at NLOS Condition

Pathloss value on NLOS condition is greater than LOS condition because on NLOS condition the loss as the result of diffraction is also counted.

On NLOS condition, pathloss value is a summation from various attenuations that occur, which are : free space loss (L_0), multiple diffraction loss (L_{msd}), and rooftop-to-street diffraction loss (L_{rts}). So in NLOS condition mobile WiMAX system performance is worse than LOS condition.

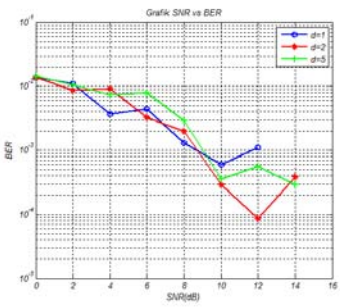


Figure 7. SNR-BER Graphic for Various Distance on NLOS Condition

Figure 7 shows mobile WiMAX system performance on NLOS condition for idle user on the distance 1000 m, 2000 m, and 5000 m. On NLOS condition the best performance is achieved at 100m, with pathloss value 152.46 dB. BER target (10^{-3}) is reached on SNR 9.07 dB. As for the distance of 2000m with pathloss 163.89 dB, the BER target is reached on SNR 9.18 dB and at the distance of 5000 m with pathloss 179.02 dB , BER target is reached at SNR value of 9.52 dB.

Table 9. System Performance on NLOS Condition for Various Distance

Distance (m)	1000	2000	5000
Pathloss (dB)	152,46	163,89	179,02
SNR (dB) on BER 10^{-3}	9,07	9,18	9,52

At SNR greater than 10 for every distance the BER value is increasing. This is because the response of the channel rapid change as the result of randomized channel. This condition shows that the mobile WiMAX performance on NLOS condition is worse than LOS condition.

4.2. Comparison of Various Velocity

This simulation is conducted to compare mobile WiMAX performance on various user velocity level whether in LOS or NLOS condition.

- **LOS Condition**

In this phase the simulation is conducted for various user velocity levels, which are : 0 km/hour, 3 km/hour, 15 km/hour, 50 km/hour, and 120 km/hour for LOS condition on the user velocity 1000 m towards transmitter with pathloss 110.56 dB.

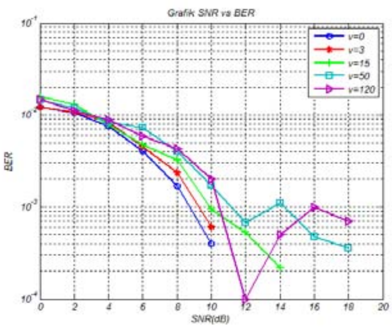


Figure 8. SNR-BER Graphic for Various Velocity on LOS Condition untuk Berbagai Kecepatan Pada Kondisi LOS

Figure 8 shows that the best mobile WiMAX performance is achieved at 0 km/hour of user velocity. And the worst occurs at 50 km/hour and 120 km/hour, where the system performance degrades. Table 10 represents system performance with BER and SNR parameters that based on Figure 8.

Table 10. System Performance for Various Velocity on LOS Condition

	distance (d) = 1000 m Pathloss (L_p) = 110,56 dB				
Velocity (km/hour)	zero	3	15	50	120
SNR (dB) on BER 10^{-3}	8,86	9,52	9,87	11,4	11,06

As shown in table 10, to reach the BER target 10^{-3} SNR of 8.86 dB is needed at 0 km/hour velocity, meanwhile at 50 km/hour it is reached with 11.4 dB SNR and at 120 km/hour with 11.06 dB. So, it is clear that the higher user velocity is then the higher BER value is, which also means system performance degradation.

At 120 km/hour of velocity different thing occurs where in SNR > 12dB the BER value rises significantly. This is because at that moment channel response changes rapidly so that the random channel causes it. This condition is caused by the shifting phenomenon of working frequency received by user along with the rising user velocity, in this case the higher user velocity is, then the higher the doppler shift is.

- **NLOS Condition**

In this phase the simulation for effect of various user velocity level on NLOS condition towads mobile WiMAX system performance is conducted. Simulation is held in 1 km distance from user to transmitter with total pathloss 152.46 dB. In this NLOS condition mobile WiMAX system performance is affected not only by user velocity but also from various occured attentuations such as : 100,40 dB free space loss (L_0), 13,58 dB multiple diffraction loss (L_{msd}), and 38,48 dB rooftop-to-street diffraction loss (L_{rts}) .

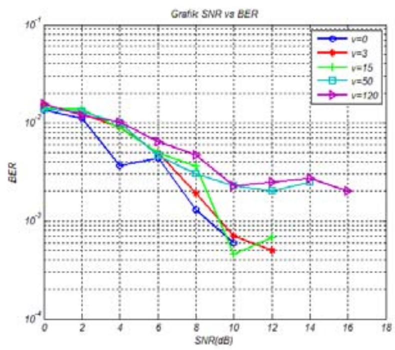


Figure 9. SNR-BER graphic for various velocity on NLOS condition

Figure 9 shows that the best mobile WiMAX performance is achieved at 0 km/hour user velocity. Meanwhile the worst performance is obatained at the 50 km/hour and 120 km/hour user velocity, where the BER target is never achieved even if the SNR value is raised, as the result of shifting shifting working frequency which is received by user and the occurence of various attentuation that happened as the result of diffraction from building rooftops beetwen receiver and transmitter.

Table 11. System performance for various velocity on NLOS condition

	Distance (d) = 1000 m Pathloss total = 152,46 dB				
Velocity (km/hour)	zero	3	15	50	120
SNR (dB) on BER 10^{-3}	9,07	9,5	9,7	-	-

From the table above it is seen that at 0 km/hour user velocity to reached 10^{-3} BER target, 9.07 dB SNR is needed. Meanwhile, at 50 km/hour and 120 km/hour, 10^{-3} BER target is never reached.10^{-3}.

At the high velocity such as 50 km/hour and 120 km/hour the performance of mobile WiMAX is very poor. This happens because of occurence of frequency shifting which is the effect of doppler shift that causes channel coherence time. This condition can be handled by using equalization technique, diversity technique, error control and power control.

4.3. The Comparison of Pathloss Value on LOS and NLOS Condition

In this phase the comparison of pathloss value on LOS and NLOS condition which is the result of various user distance simulation on Walfisch-Ikegami channel is conducted.

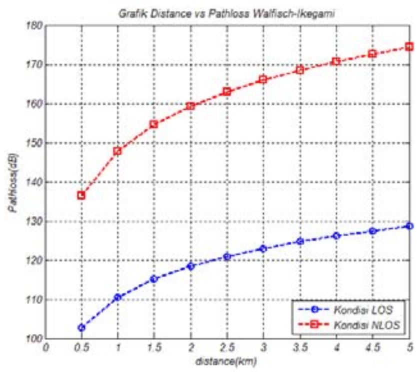


Figure 10. The comparison of Pathloss value on Walfisch-Ikegami channel

Pathloss value is hugely different between LOS and NLOS condition. Pathloss value on NLOS condition is higher than in LOS condition.

Table 12. Comparison of Pathloss value on Walfisch-Ikegami channel

distance (m)	500	1000	2000	3000	5000
Pathloss on LOS condition (dB)	102,7	110,6	118,4	122,9	128,7
Pathloss on NLOS condition (dB)	141,1	152,5	163,9	170,6	179,1

Table 12 proves that pathloss on COST231 Walfisch-Ikegami model on NLOS condition is higher than in LOS condition. This condition happens because pathloss value on NLOS condition is the total of various attentuation that occured such as : free space loss (L_0), multiple diffraction loss (L_{msd}), and rooftop-to-street diffraction loss (L_{rts}).

Meanwhile on LOS condition pathloss value is obtained just from free space loss calculation.

5. Conclusion

Based on simulation and analysis results, several conclusion are obtained :

1. On LOS condition with static (idle) user, the best performance of mobile WiMAX is gained at 1000 m of user distance and to reach the BER target, SNR 8.86 dB is needed. For user distance 2000 m and 5000 m, the performance of mobile WiMAX is relatively good.
2. On LOS condition with static (idle) user for each level of user distance, mobile WiMAX experiences the performance degradation. The BER value increases along with the increasing SNR value. The best performance is obtained at 1000 m distance where the BER target is gained with 9,07 dB as the SNR value.
3. The faster user moves then the worse the system performance is, because the increasing velocity causes higher Doppler Shift. On LOS condition the best performance is achieved in 0 km/hour, and to gain the BER target 8.86 dB of SNR is needed. The worst performance occurs at 120 km/hour where 11.06 dB of SNR is needed to reach the BER target.
4. On NLOS condition for the velocity of 50 km/hour and 120 km/hour, mobile WiMAX has very poor performance where the BER target is never achieved.
5. Pathloss value on NLOS condition is higher than the LOS condition because of the diffraction by the building rooftop which causes rooftop to street diffraction loss.

6. Future Work

1. To handle the doppler shift effect which is channel coherence time, equalization technique and power control could be implemented for better mobile WiMAX performance.
2. The research for various conditions in different building height levels and various levels on other Walfisch-Ikegami parameter is necessary to predict better mobile WiMAX propagation channel.
3. For more accurate research, case study in Indonesia urban area is advised.

References

[1] Bariah, Khairul. 2007. Performance Of Worlwide Interoperability For Microwave Access (WiMAX) In Non Line Of Sight And Multipath Environment. Universiti Teknologi Malaysia.

[2] Chandra, Praphul., M. Dobkin, Daniel., Bensky, Alan., Olexa, Ron., A. Lide, David., Dowla, Farid. 2008. Wireless Networking. Newnes.

[3] Ervan Surahmat, Andi. 2008. Perancangan Algoritma Subkanalisasi Pada Mobile WiMAX IEEE 802.16e. Bandung: Institut Teknologi Telkom.

[4] G. Andrews, Jeffrey., Ghosh, Arunabha., Muhamed, Rias. 2007. Fundamentals of WiMAX. Prentice Hall.

[5] Goldsmith, Andrea. 2005. Wireless Communications. Stanford University.

[6] Jain, Raj. 2007. Channel Models A Tutorial.

[7] Kurniawan, Uke. 2005. Propagasi Model, Slide mata kuliah Sistem Komunikasi Bergerak. Bandung: Institut Teknologi Telkom.

[8] P. Das and D. Koch. 1991. IEEE Procedings Of Southeast Con. On The Use Of Visual Programming Languages For Communication System Simulation.

[9] Rinne, Jukka. 2003. COST-231 Path-loss models. Tampere University of Technology and Institute of Communications Engineering.

[10] Roca, Amalia. 2007. Implementation of a WiMAX simulator in Simulink. Spain: Institut fur Nachrichtentechnik und Hochfrequenztechnik.

[11] Triggs, Alan. 2001. Lecture 7: Propagation, Fading, Link Budgets & Coverage. Southern Methodist University.

[12] Tse, David., Viswanath, Pramod. 2005. Fundamentals of Wireless Communication. Cambridge University Press.

[13] Waskita, Henry. 2007. Pemodelan Kanal Mobile WiMAX Dengan Menggunakan Model Kanal SUI. Bandung: Institut Teknologi Bandung.

[14] WiMAX Forum. 2004. WiMAX's Technology for LOS and NLOS Environments.

[15] WiMAX Forum. 2005. Fixed, nomadic, portable and mobile applications for 802.16-2004 and 802.16e WiMAX networks.

[16] WiMAX Forum. 2006. Mobile WiMAX – part 1: A Technical Overview and Performance Evaluation.

PERFORMANCE ANALYSIS OF ZIGBEE PROTOCOL ON WIRELESS PERSONAL AREA NETWORK (WPAN)

Vera Suryani[1], Yudi Satria Gondokaryono[2]

[1]Institut Teknologi Telkom
[2]STEI, Bandung Institut of Technology
vra@ittelkom.ac.id[1], yudi@stei.itb.ac.id[2]

Abstracts

In this research, Zigbee/IEEE 802.15.4 protocol was studied by its energy, delay, packet loss, and throughput measurement as performance parameters of the protocol. Measurement conducted by simulation using *Network Simulator 2* (NS-2). Measurement result is analyzed further using *Analysis of Variations* (ANOVA) method to observe what kind of parameters that effect to Zigbee/IEEE 802.15.4 protocol. Experiment result shows that AA alkaline type battery that used as power supply of node (as device) can last for about 17 month. Based on ANOVA calculation, time interval effects energy consumption, neither to payload size. Otherwise, energy parameter is also affected by packet size and transmission distance.

Keywords : Zigbee/IEEE 802.15.4, WSN, WPAN, , *delay, packetloss, throughput.*

1. Introduction

Wireless Sensor Network (WSN) is a distributed system and consist of sensor devices, data packet transmitter, and operated on wireless network, has limitation in energy. WSN was implemented for monitoring and controlling processes in industrial field, for patient monitoring in medical field, house automation application, and traffic controlling. Every sensor devices was completed by power supply, usually as battery.

Zigbee protocol/IEEE 802.15.4 is wireless communication protocol which has communication range about 100 metres and low power consumption. This characteristic makes Zigbee/IEEE 802.15.4 protocol frequent used with WSN to save sensor device's power. In wireless communication, power saving becomes important issue.

In this research, power consumption of sensor devices and WSN using Zigbee/IEEE 802.15.4 protocol was studied. We use total energy consumption for 1 hour observation in star and mesh topology as performance parameter. Other parameters that measured are *delay, packetloss,* and *throughput*.

2. Theory

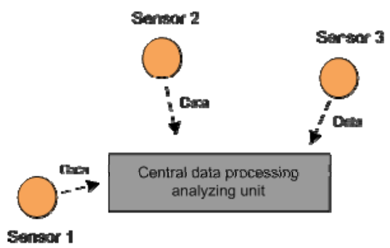
A. WPAN (Wireless Personal Area Network)

WPAN is a network that connecting wireless devices in short range, for few meters only. WPAN is plug in technology, in ideal condition PAN devices can connect automatically within the range. Devices can be locked by other devices selectively to avoid interference or unwanted information access from other devices [13]. WPAN technology developed by IEEE for low power consumption is 802.15.4 WSN.

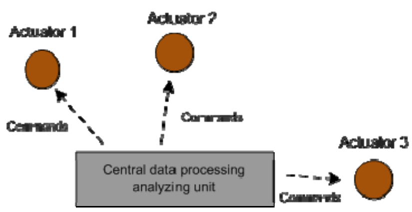
B. *WSN (Wireless Sensor Network)*

WSN is build for monitoring, controlling and combination both of them [1]. For monitoring, WSN is usually used for temperature telemetry, pressure telemetry, etc. In controlling, WSN used for command transmission for actuator such as switch, robot, etc.

In industrial field, developers or vendors need standardization to ensure devices from different vendors can communicate. Standard for WSN is IEEE 802.15.4 by IEEE and Zigbee that was developed by *Zigbee Alliance* [1].



(a)



(b)

Figure 1. (a) WSN for M*onitoring*
(b) WSN for *Controlling* [1]

C. *IEEE 802.15.4 and Zigbee*

Zigbee come from words *zig* and *bee*, *zig* means zig-zag movement. The name noticed that the protocol imitates bees behavior in spreading information about source of honey to other bees [4]. Zigbee was developed by Zigbee Alliance as alternative protocol which has low power consumption, low data rate, low cost, and targeted for wireless network in automation and remote controlled application. IEEE Committee 802.15.4 worked for low data rate protocol decided to merge with *Zigbee Alliance* and use Zigbee as trade mark.

Fundamental differences between Zigbee and IEEE 802.15.4 lies on layer coverage handles by each Committee. IEEE 802.15.4 focuses on 2 lowest layer, physical layer and MAC layer. Whereas Zigbee Alliance focus on 2 other protocols, layer *network* and layer *application* [2]. Zigbee/IEEE 802.15.4 architecture can be seen at Figure 2 below.

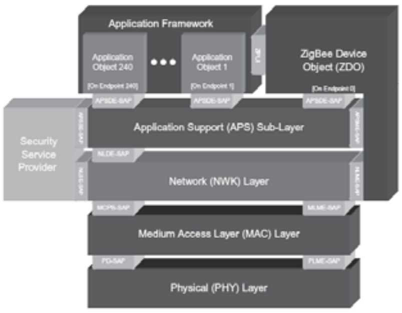


Figure 2. Zigbee Architecture

D. *Adhoc On-demand Distance Vector (AODV) Protocol*

AODV routing protocol works on necessity to find path, path will available only if needed by source node to transmit data packet. AODV uses destination sequence number to identify newest path. Major differences between AODV and *Dynamic Source Routing* (DSR) are DSR uses source route on data packet and brings complete path to destination. In AODV, source node and intermediate node saves information for next hop about each data flow transmission. In routing on-demands protocol, source node floods route request packet in network when a route does not available for wanted destination. The method could result some routes from a single route request. Major different between AODV and other routing on-demand protocol is the usage of *destination sequence number* (DestSeqNum) to determine up-to-date path to destination. Node updates its path information only if DestSeqNum from received packet bigger than last DestSeqNum (which is saved in last packet).

A RouteRequest bring source identifier (SrcID), *destination identifier* (DestID), *source sequence number* (SrcSeqNum), *destination sequence number* (DesSeqNum), *broadcast identifier* (BcastID), and *time to live* (TTL). DestSeqNum indicates novelty of path which received by source. When intermediate node received a RouteRequest, it will prepare a RouteReply if valid route to destination is exist. Validity of route is detemined in intermediate node by comparing sequence number in intermediate node with destination sequence number in RouteRequest.

If a RouteRequest received repeatedly, the pair indicated by the BcastID-SrcID is removed. All intermediate nodes have a valid route to the destination, or destination node itself is allowed to send packets to RouteReply source. Each intermediate node, while a forward RouteRequest, enter the address of the node before and BcastID. It use a timer to remove this entry in the case of a RouteReply not received before the time runs out. This helps in saving active paths. When a node receives a packet RouteReply, information about the previous node where the packet is received is replaced by the new information.

3. Modeling And Simulation

The modeling of energy usage and AODV routing protocol in WSN network using IEEE 802.15.4 was conducted by *Network Simulator 2* (NS2). Modelling and simulation will describe below.

A. *Modeling*

Energy

Energy on the Network Simulator 2 modelled as an attribute of the node that represents the energy level on the node. Modeling is made using a class `EnergyModel` [14] and defined in the following:

```
class EnergyModel : public TclObject
public:
EnergyModel(double energy) energy_ =
energy;
inline double energy() return energy_;
inline void setenergy(double e) energy_
= e;
virtual void DecrTxEnergy(double
txtime, double P_tx)
energy_ -= (P_tx * txtime);
virtual void DecrRcvEnergy(double
rcvtime, double P_rcv)
energy_ -= (P_rcv * rcvtime);
protected:
double energy_;
```

As shown in the code above, there is only 1 class variable, namely `energy_` which represents the energy level of a node at a time. `EnergyModel (energy)` requires the energy-intial as a parameter. Class method used to reduce the energy level of nodes at the time when sending the packet is `(DecrTxEnergy (txtime, P_tx))` and at the time of receiving the packet is `(DecrRcvEnergy (rcvtime, P_rcv))`. `P_tx` is the amount of energy sent by the PHY layer of nodes and `P_rcv` is the amount of energy received by the PHY layer of nodes. At the beginning of the simulation, the parameters `initialEnergy_` in use as `energy_`, where the value will be reduced each time the process occurs sending and receiving of packets on the node. When the energy level of nodes to be continuously reduced to zero, then the packet will no longer be received or sent from these nodes. The syntax code `DEBUG: node <node-id>` dropping pkts due to energy = 0. Energy modeling used in simulation is adopted from modelling made by Liang Jian Zheng Cuny's University in New York and Vijay Kakadia from the University of Southern California [8].

AODV Protocol Routing

Routing modeling in NS-2 generally consists of 3 parts:

a. *Routing agent* : to replace the packet with a neighbor node
b. *Route logic* : contains information gathered from the routing agent (or a database that contains the topology used in the static ruting)
c. *Classifiers* : located in the node, is used to calculate the ruting table so that packets can be forwarded to the destination node

The implementation of new protocols, eg AODV sometimes does not need all three component above to be implemented. This only happens if previous routing agent that similar behavior with AODV, eg Distance Vector routing protocol. When a new routing protocol implementation includes more than one function blocks, especially when it contains its own

classifier, it is desirable to have another object, which we call a *routing module*, that manages all these function blocks and interface with node to organize its classifiers.

The interaction process of node, routing module, and routing can be seen in Figure 3.

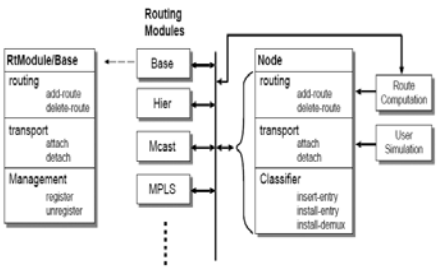


Figure 3. Interaction Between Nodes, The Routing Module, and Routing. Broken Line Shows The Detail of The Routing Protocol [13]

AODV routing agent using the same routing agent as Distance Vector. AODV module used in the simulation is a module that is integrated in the NS-2 version of 2:31.

B. Simulation

Topology used in the simulation:

1. Star Topology

The use of star topology in the simulation was intended to replicate the real conditions of the IEEE 802.15.4 sensor applications in the home. In star topology used 6 nodes :

a. Node 0 as PAN Coordinator
b. Node 1, 2, 3, 4, 5, and 6 are devices that located ± 10 meters from the PAN coordinator. Node 1 is device 1, node 2 is device 2, and so on.

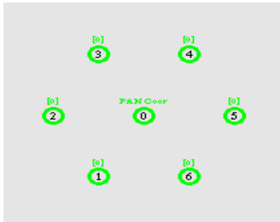


Figure 4. Star Topology on Simulation

Star topology is used in scenario 1 because the topology is sufficient for a home area coverage [3].

2. *Mesh* Topology

Mesh topology is used in the simulation for routing process modelling in the IEEE 802.15.4 network. Topology is selected with more nodes and the wider area of the previous topology. Mesh topology is used in the simulation can be seen in Figure 5.

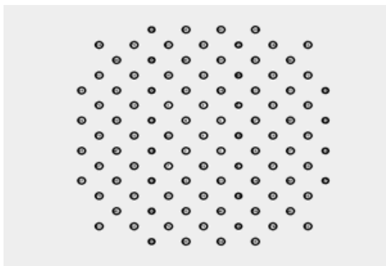


Figure 5. Mesh Topology on Simulation

C. Simulation Scenarios

Scenario 1 :

Star topology is used, and this scenario used to determine energy consumed by each node with different time interval and various distances. The input parameters for the simulation are presented in Table 1.

Table 1. Input Parameter for Simulation

Parameter	Spesification
Distance between nodes	± 10 meter
Packet size	30 bytes, 60 bytes, and 90 bytes
Packet rate	5 Kbps
Initial Energy	13770 Joules
Energy that is used to send the packet (txPower)	0.0744 Watt
Energy used to receive the packet (rxPower)	0.0648 Watt
Energy at idle or sleep	0.00000552 Watt or5.52 µW
Long simulation	3600 s
Error model	Uniform distribution

Routing protocol in this scenario is disable since there is no intermediate node between the sender and receiver.

Time interval used in simulation : 1 seconds, 10 seconds, 30 seconds, 1 minute, 15 minutes, 30 minutes, and 1 hour. Distance used in the simulation : 10 meters, 15 meters, 20 meters, 25 meters, 30 meters, 35 meters and 40 meters.

Scenario 2 :

The purpose of scenario 2 is to see the behavior of AODV protocol in IEEE 802.15.4 network. Packet' size used in the scenario: 30 bytes, 60 bytes, and 90 bytes. While the distance ranges from 10 meters up to 80 meters. Background traffic is not used with the assumption that IEEE 802.15.4 network is not used to transmit other data.

4. Result and Discussion

Experiments conducted with 2 scenarios as described in section 3 is used to determine the performance of the protocol on the IEEE 802.15.4 Wireless Sensor Network. Energy was calculated from device node, not from PAN coordinator. Value that is displayed on the graph is the average value of all nodes.

4.1. *Analysis on Scenario 1*

The purpose of this scenario is to give a picture about how much energy (in Joule) issued by the node using the protocol of IEEE 802.15.4. Simulation is done with different time interval, different distance, and average energy of 6 nodes.

4.1.1. Interval Parameter

Simulation tested in star topology with 6 nodes, with each node send a packet every 1 second interval, 10 seconds, 30 seconds, 1 minute, 5 minutes, 15 minutes, 30 minutes, and 1 hour. Node 0 acts as the PAN coordinator .

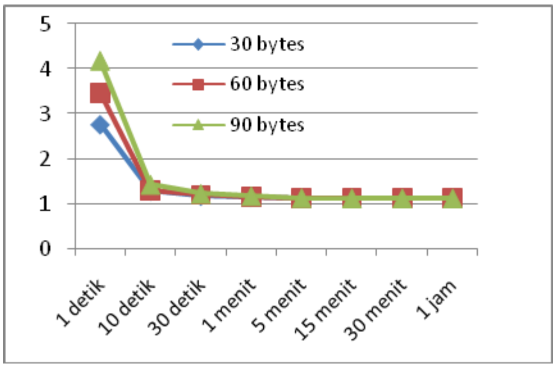


Figure 6. The Relationship of Energy Consumption (Joule) and Time Interval

Based on the results we can obtain that energy is decreasing at 1 second to 10 seconds interval. This is because in time interval bigger than 10 seconds node tends to have more sleep than send/receive data. Thus, the short interval of the packet the more likely that energy used.

4.1.2. Distance Parameter

Simulation is done on the star topology with 6 nodes, with the distance of the node that functions as a device with the PAN coordinator, ranging from 10 up to 40 meters. Energy calculated based on the average energy of 6 sender node.

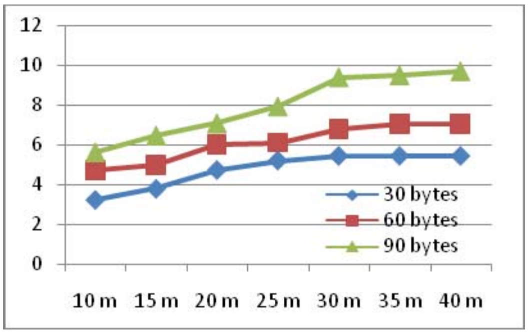


Figure 7. The Relationship of Energy Consumption (Joule) and Distance Interval

From Figure 9, the increase of energy going on the distance of 20 to 30 meters. Longer distance caused increasing energy on each node. The increasing energy is not caused by the packetloss lost since the receiving throughput is 100%.

4.1.3. Delay, Packetloss and Throughput Parameter

Results of delay calculation in scenario 1 can be seen in Table 2.

Table 2. End-to-End Delay in Scenario 1

	10 m	15 m	20 m	25 m	30 m	35 m	40 m
30 bytes	2.5927	2.591967	2.591967	2.591967	2.591967	2.591967	2.591907
60 bytes	3.551967	3.551967	3.551967	3.551967	3.551967	3.551967	3.551907
90 bytes	4.511967	4.511967	4.511967	4.511967	4.511967	4.511967	4.511917

End-to-end delay value from receiver node is not different for the range of 10 to 40 meters. This is caused by the distance that is not significant compared with the electromagnetics propagation waves in the air (3.10^8 m/s). Packetloss for all nodes 0%. While the throughput for the payload of 30 bytes is 1508 kbps, 60 bytes is 3024 kbps, and 90 bytes is 4536 kbps. From the experimental results can be obtained that for star topology with 6 nodes will produce throughput up to 100%.

4.2. *Analysis on Scenario 2*

If scenario 1 using star topology, in scenario 2 mesh topology is used. Performance parameters used in this scenario are more likely with the scenario 1 : time and distance interval, end-to-end delay, packetloss, and throughput. Many nodes are involved in this scenario, so AODV routing protocol is used to facilitate the search path between sender and receiver nodes.

4.2.1. Interval Parameter

Scenario used for time interval : 101 nodes in mesh topology scenario. The distance between sender and receiver is ± 20 meters. Energy displayed in the figure 8 calculated from sender node.

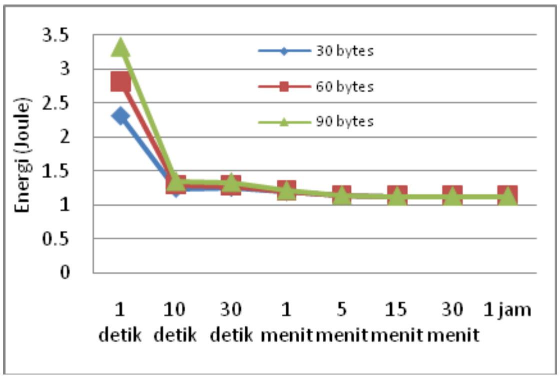


Figure 8. Relations of Energy (Joule) of Parameter Interval

Figure 8 shows that for the node with the packet sent every 1 second (30 bytes of payload) will spend Alkaline Battery type AA for about 8 months, and if the time interval is 1 hour it will consume energy battery for about 16.9 months or 17 months. While for 60 bytes of payload will cause battery life time is last for 6.7 months, and payload of 90 bytes will be out in time-lapse of 5.7 months. For 60 or 90 bytes of payload and an hour sending time interval will cause the battery lifetime for about 16.9 month. In these calculation, 1 month last for 30 days.

From the simulation result we can infer that AA type battery can last up to 2 years with optimal packet size is 30 bytes and transfer rate is 5 kbps

4.2.2. Distance Parameter

Energy measurement is done on the sender node within 1 second sending time interval. In this mesh topology, distance between nodes is set to 10 meters. Result of experiment can be seen on figure 9.

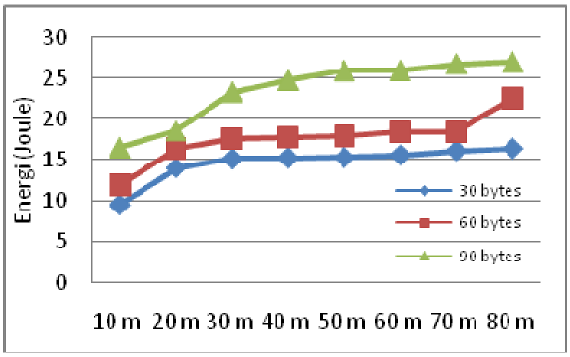


Figure 9. Relations of Energy (Joule) of the distance parameters

The increasing energy on the 60 bytes of payload (distance 70 to 80 meters) due to packet losses. These packet losses caused the node to release more energy for retransmission.

4.2.3. Delay Parameter

Measurement of delay in the scenario 2 is done by sending 3 types of packets: a payload of 30, 60, and 90 bytes. Different distance is set from 10 to 80 meters. Rate that is used by all payloads is 5 kbps.

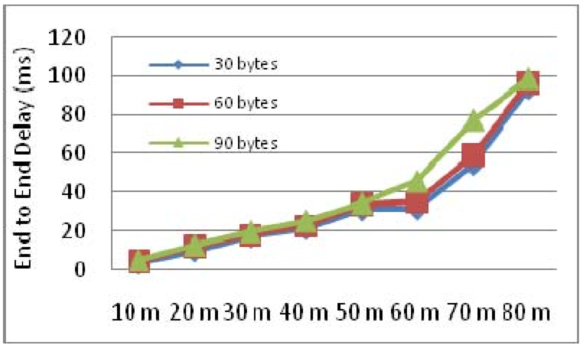


Figure 10. Value of End-to-End Delay on The Distance

The experiment showed that the bigger packets caused bigger delay. 30 and 60 bytes of payload would produce the tolerable delay with optimal distance is 70 meters. The optimal distance for tolerable delay in 90 bytes of payload is 60 meters. Torelable delay for home application is 10 ms/hop or 10 ms/10 meter [3].

This result indicates the optimal distance between sender and receiver in mesh topology ranged between 60 – 70 meters, with payload of 90 bytes.

4.2.4. Packetloss Parameter

Experiment is conducted to see the percentage of packets lost during the delivery process and its causes. bytes. Not included in the experiment is background traffic assumptions, there is no other process of data

during simulation. Error modeling used in this simulation is Uniform distribution with an error rate of 0.2%.

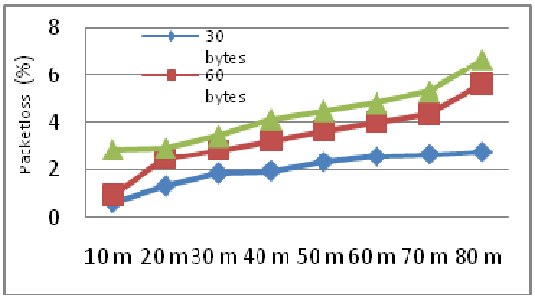


Figure 11. Packet Lost Percentage

Packet lost during transmission caused by many things. Error model used in this simulation caused some of the packets marked as a packet with the broken bit. If the error correction mechanism in the IEEE 802.15.4 protocol is not able to overcome the broken bit in the packets, than it is discarded.

Packetloss was also caused by weak signal produced during data communication, marked as LQI value. Far distance caused the signal power to be decreased, yield that the packet can not be arrive to destination safely.

4.2.5. Throughput Parameter

Throughput is affected by packetloss value. Thus the decrease of throughput value can be caused by the increasing of packetloss value. Results from the calculation of throughput from sender to receiver can be seen in Figure 12.

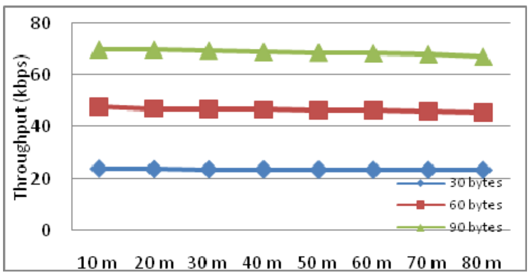


Figure 12. Average Throughput

Simulation results show the decreasing value of throughput caused by increasing distance was not too significant. The discarded packet is relatively small compared with the total packet that is sent during data transmission with the value of 7%. From the results of the simulation can be drawn that large packets under (maximal 90 bytes) with a rate of 5 kbps on the mesh topology is still possible to be applied in home automation.

4.3. *ANOVA Analysis*

Method to observe what kind of parameters that effect to Zigbee protocol is ANOVA 2-Factor-Without-Replication.

Energy on star topology

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*
Rows	0.0925	2	0.0463	1.5051	3.7389
Columns	7.1026	7	1.0147	33.006	2.7642
Error	0.4304	14	0.0307		
Total	7.6255	23			

Energy on mesh topology

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*
Rows	8232.412	2	4116.206	25630.93	3.7389
Columns	7.07604	7	1.0109	6.2945	2.7642
Error	2.2483	14	0.1606		
Total	8241.736	23			

Delay

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*
Rows	333.3229	2	166.6615	91.2521	3.7389
Columns	176.4558	7	25.208	13.8021	2.7642
Error	25.56939	14	1.8264		
Total	535.3481	23			

Packetloss

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*
Rows	21.78085	2	10.89042	53.19921	3.738892
Columns	26.30312	7	3.757588	18.35564	2.764199
Error	2.865943	14	0.20471		
Total	50.94991	23			

Throughput

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*
Rows	223.7855	2	111.8927	6.8044	3.7389
Columns	19547.88	7	2792.555	169.8213	2.7642
Error	230.2172	14	16.4441		
Total	20001.89	23			

Figure 13. ANOVA Analysis Tables

From ANOVA calculation we can infer that time interval effects energy consumption, neither to payload size. Otherwise, energy parameter is also affected by packet size and transmission distance.

Meanwhile, delay, packetloss, and throughput are affected by their distance and payload.

5. Conclusion

Conclusions from the research are :

1. Node that functions as a network device in the IEEE 802.15.4 star topology will be able to

survive for 6 months to 2 years with the power derived from Alkaline Battery type AA.

2. Packet delivery mechanism alternately on the star topology with 6 nodes can increase the throughput compared to the delivering packet simultaneously.
3. Battery lifetime for each device node on Zigbee protocol can stand up to 2 years with optimal packet size is 30 bytes and rate of 5 kbps
4. Percentage of throughput received on the mesh topology is 93% to 99% of the total packet that is sent.
5. Tolerable delay for home automation ranges in the value of 80 ms and at 80 meters distance
6. The cause of packeloss on Zigbee caused by the quality of the signal with indicator Link Quality Indicator (LQI), and full buffer capacity of the receiver node that indicated by the IFQ code
7. Time interval effects energy consumption, neither to payload size.
8. Energy parameter is affected by packet size and transmission distance.
9. Delay, packetloss, and throughput are affected by their distance and payload

References

[1] Anonim. (2008) : Wireless Sensor Network – *Mesh*Netics Learning Courses, downloaded in April 2009 at www.*mesh*netics.com/learning,

[2] Burchfield, T. Ryan dan S. Venkatesan : Maximizing Throughput in Zigbee Wireless Sensor Network through Analysis, Simulations and Implementation. *White paper.* 2006.

[3] Callaway, Ed., *Gorday, Paul., dan Hester, Lance., A. Gutierrez, Jose., Naeve, Marco., Heile, Bob.* : Home Networking with IEEE 802.15.4: A Developing Standard for Low-rate Wireless Personal Area Networks, *IEEE Communications Magazine*, Agustus 2002.

[4] Dermawan, Candra. Zigbee : Teknologi Wireless Masa Depan, downloaded in April 2009 at http://candra.us/blog/?p=75

[5] Gunarto, Thomas Yuni : ANOVA (Analysis of Variation), E-book from Universitas Gunadharma, 2008.

[6] Iqbal, Muhammad Syamsu : Studi Kinerja Protokol IEEE 802.15.4 untuk Jaringan Sensor Nirkabel Adhoc, *Thesis,* Bandung: ITB, 2008.

[7] J. Zheng dan M.J.Lee : A Comprehensive Performance Study of 802.15.4, IEEE Press Book, 2004.

[8] J. Zheng : Zigbee/IEEE 802.15.4 Routing Simulation, Samsung/CUNY, 2003.

[9] Kinney, Patric (Chair of IEEE 802.15.4 Task Group): ZigBee Technology: Wireless Control that Simply Works, *Communications Design Conference,* 2003.

[10] Marandin, Dimitri : Simulation of IEEE 802.15.4/ZigBee with Network Simulator-2 (ns-2), downloaded in Maret 2009 at http://www.ifn.et.tu-dresden.de/~marandin/ZigBee/ZigBeeSimulation.html

[11] Myers, Raymond dan Walpole Ronald : Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuwan, Penerbit ITB, 1995.

[12] Pruim, Randall : ANOVA (Analysis of Variation), Department of Mathematics and Statistics at Calvin College, E-book from http://www.calvin.edu/~rpruim/courses/m143/F00/ downloaded at Mei 2009.

[13] TechTarget : What is WPAN? di unduh dari http://searchmobilecomputing.techtarget.com/sDefinition/0,,sid40_gci837444,00.html], downloaded at Mei 2009.

[14] UC Berkeley, LBL, USC/ISI, and Xerox PARC : The ns Manual, Maret 2008.

[15] Wirawan, A.B. dan Eka Indarto : Mudah Membangun Simulasi dengan Network Simulator-2, Penerbit Andi, Yogyakarta., 2004

Institut Teknologi Telkom

Sponsored by :

Charisma Profil
Furniture & Interior design
Jl. Sadang Asri Tengah No. 38 Blk. GOR Inhoftank Bandung
Telp. 7505482 HP. 0811224773

WASKITA

telkom professional development center
pdc
Center for Continuing Professional Education

PT. je_ferlasthama
architecture + interior consultant
komp. gudang peluru blok f no.147
tebet, jakarta 12830
t. 62 21 83785505 f. 62 21 8294937

KOCITEL

PRECAST AND PRESTRESS CONCRETE
PT BETON ELEMENINDO PERKASA
Jl. Raya Batujajar Km.5 No.56 Telp. (022) 6866031 Fax. (022) 6866033
Website : www.beton.co.id E-Mail : beton@bdg.centrin.net.id