**Sri Lanka Institute Of Information Technology**

****

Comprehensive Design/Analysis Projects

Project Proposal

**Vigraha High Volume CDR Analyzing Framework**

Title of the project : Vigraha High Volume CDR Analying Framework

Project ID :

Student name : Lasantha Perera

Supervisor name : Mrs.Sujanthi

Co - Supervisor name :

Supervisor signature :

Co - Supervisor signature :

Declaration

I am ensuring that the proposed project work in ” **Vigraha High Volume CDR Analyzing framework**”, is a record of an original work done by me and this is not a copy of anyone.

Date Signature

Abstract

The proposed project is Vigraha High Volume CDR Analyzing Framework. In modern world the major communication method is using mobile phones. So in this system it mainly focuses on developing more useful application for mobile companies. This is a main gate for brand companies and mobile companies. The system is based on distributed computing.

This distributed computing system takes all brand companies requirements as inputs and according to them it generates the messages for customers as output. It informs by a SMS. To select the customers here it uses some log files which already available at mobile companies. It takes all customer details under many log file. In here the system mainly focuses on five log files to gather customer information. The system keeps all records using a single database. Then it analyzes records when brand companies requirements available. Then it generates a SMS for all selected customers and sends the messages. If the system has some failure it sends a SMS to the person who is responsible for the system at that moment. The system is providing reports for the brand companies.

The main target of this system is to analyze millions of records and provide SMS for necessary customers. Also generate reports is another outcome of this system. So this system will be more useful area for communication.

**Table of content**

**Introduction**

Among many communication methods using mobile phones become most powerful factor in today world. As mobile features there are many things such as voice calls, SMS, GPRS, LBS (Location Base System) and so on. In mobile companies they track the customer information regarding these. That means they maintain some log files for these items. Then when the customer uses a voice call or any other feature in mobile phone they track those records to these log files. They do this because to analyzing data. So the proposed project is about application for distributed computing.

This project is more worth for mobile companies, brand companies and mobile users, because when a brand company needs to announce about their promotions or events to their customers they provide this to mobile companies with some details. Then mobile companies select necessary customers to send SMS. Also it provides reports for brand companies. So this is very useful for brand companies. By this the mobile companies can have a very good income. So this is very worth for mobile companies. Also the mobile users able to know the new promotions available in brand companies. So this is worth for mobile users.

The system has been divided into four major components.

* **Data loader**
* **Admin UI**
* **Reporting UI**
* **Rule Engine**

Data loader – use to take records from log files to database.

Admin UI – Has to register and log to the system by administrator and he takes requirements and generates rules.

Reporting UI – Has to generate 30 reports including drill down report.

Rule Engine – Execute rules and select suitable customers by analyzing records.

In many ways the system helps to their clients.

* Provide service for brand companies by mobile companies - The system is useful for brand companies to announce their promotions, events, occasions or any other special functions. They can inform these messages their customers through mobile companies. They only need to tell about their requirements. Then the all tasks will handle by mobile company.
* Provide information to customers - When the mobile company sends SMS for necessary customers they can know about what are the things they can consume from brand companies. Assume mobile company is having requirements from pizza hut and they announce about pizza promotion to customers and at that moment a customer needs to it, actually the message will help him to fulfill his need.
* Handling millions of records - In this system it has to deal with millions of records, since it receives millions of customer details per day by voice calls, SMS, LBS, GPRS and so on. So the system has to maintain each and every record in proper manner and it able to select necessary records according to their customer requirements. Then generate the SMS fro customers. Also for special days such as new year, year end and so on the mobile companies receives more customer information than regular days.
* High availability - In here the server should up and running each and every day, since it deal with customer records. So maintain system’s availability is more important.
* Generate reports - In order to providing the summary report to the brand companies is another objective, because if the brand company need to know to whom mobile company sends SMS obviously the system need to provide a report for them.

In the existing systems it only concern about one log file to analyze records, because if use more than one log file it contains more records to handle. It is not an easy task. Also they do not use hadoop integration to analyze data. The existing system is an industrial project. But in our system it concerns about five log files and uses all records in those log files to maintain data.

**Objectives**

In this system it contains some major objectives as well as some sub objectives.

Main objectives:

* + To analyze huge number of records.
  + To provide services for companies (mobile and brand companies).

Other objectives:

* To provide reports.
* To inform customers about brand companies announcements by SMS.
* To categorize mobile users base on log files.
* To maintain log files, because has two files per log file (SMS, voice calls, GPS, LBS).
* To perform the availability of the database.
* To select necessary customer groups according to brand companies requirements.
* To introduce a new system for mobile companies.

Research questions

1. What are the major contribution of the system and how each component affect to each other ?
2. What methods use to handle data analyzing section and what kind of problems occur when analyze data ?
3. What is the method taken to keep server up and running all the time ?
4. To send SMS to customer what are the approaches that has to be considered ?
5. What databases support for data analyzing and what approach did selected ?
6. What kind of requirements gather from brand companies when select the customers ?

**Methodology**

Among many communication methods using mobile phones become most powerful factor in today world. As mobile features there are many things such as voice calls, SMS, GPRS, LBS and USSD. In mobile companies they track the customer information regarding these. That means they maintain some log files for these items. Then when the customer uses a voice call or any other feature in mobile phone they track those records to these log files. They do this because to analyzing data.

Assume another brand company needs to inform their customer about their new promotion or any other main functions. So simply they tell about it to mobile company with some requirements. The requirement can be age range, the current location of the customer or so on. So the mobile company has to inform this message to customers who match with the requirements. So then analyze their log files and select suitable records and inform about this to customers. But the existing systems which available in mobile companies they only consider about only one log file to analyze records. They not analyze the other sections. But in this system it considers about five sections to analyze records. So the research part is to analyze the customer recodes. It is not an easy task, because they receive millions of records per day. For especial days such as New Year day, Christmas day and so on this record count will be increased. All of them need to handle also.

Here it has separate two files for each log file. It has given some time period for the file and after that time period if has records it sends data to database and for second file. Likewise it sends data records to database. Here it has separate user interface for administrator and for report generating. In administrator section it has a registration and login session. Then insert customer requirements to the system and generate rules according to customer requirements under some categories such as promotions, loyalty programs and so on. Then the rule sends to the database. Then it selects the relevant records to inform customers by SMS. Under report generation section assume if a brand company needs to know to which customers that they inform the message then it able to provide a report.

The representation of high level architecture is representing follow.

Log Files

Data Loader

Admin - UI

Data Store

Reporting - UI

Rule Engine

SMSC

Hadoop Map Reduce

Figure 1

My contribution

In my contribution has to generate reports. Here has a separate Graphical User Interface. When the system executes the rule, it analyze records by hadoop integration and sends SMS to mobile users. Then the system able to provide reports for them. Assume the company needs to know to whom it sends SMS. Then it has to generate information with customer phone number, name of the customer, whether the customer receive that SMS or not.

The system need to generate,

* + Current status reports
  + History reports

Here it contains three kinds of reports. For one kind of report it contains drill down report. That means the system generates reports year wise, month wise, company wise. Here has to generate 30 reports.

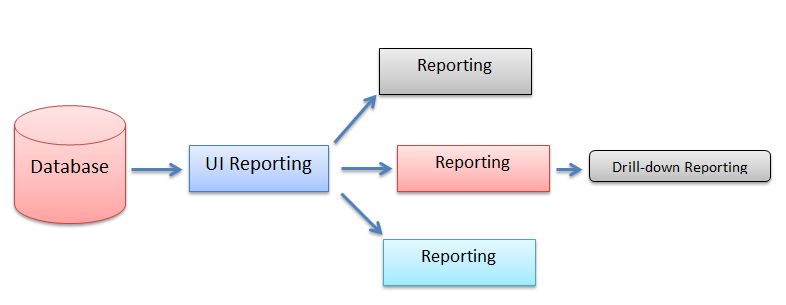
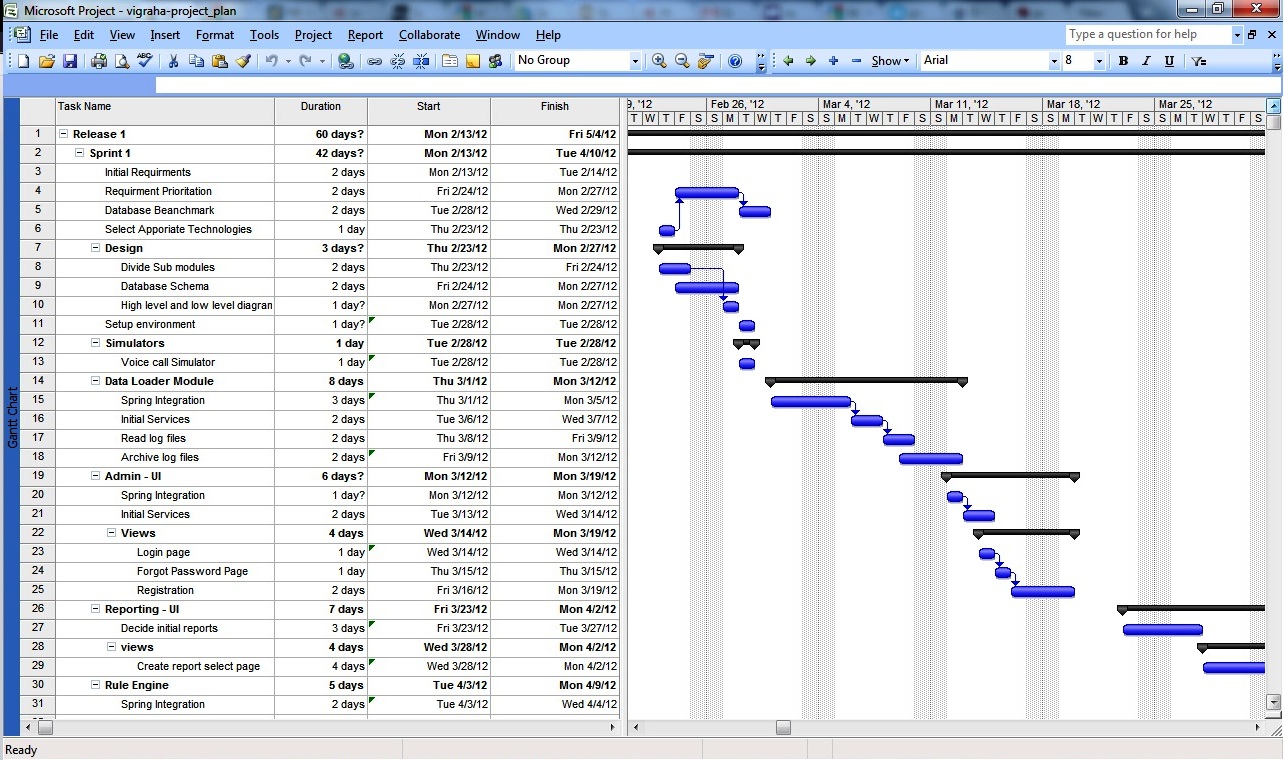


Figure 2

Following indicates the project plan for entire year.

****

Work Breakdown Structure

**Vigraha High Volume CDR Analyzer**

Analyzer

Administrator UI

Maintain log files

Report UI

Create stimulators

Send records to database

Register and log administrator

Get requirements

Generate rules

Execute rule

Analyze records

Send SMS

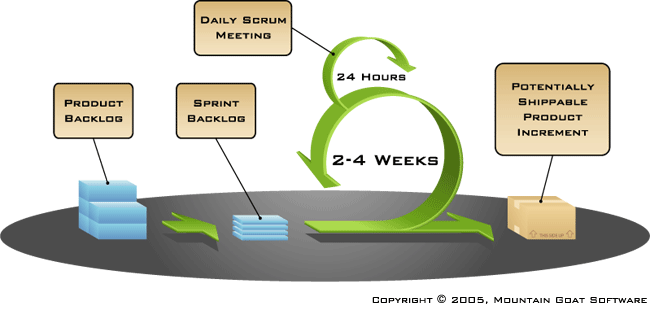
Get requirements

Generate report

Figure 4

Carry out the project

To perform the system we are using agile method. Under this category we use scrum method. In this method has to develop the system for particular time period and then it present to the client. Then he can give comments and according to them we can carry out the rest. So at the end no need to repeat entire project, because time to time we are collecting customer feed backs. If use other models such as waterfall method, iterative model has to face many problems. May be those models cause to waste time, human resources and waste money. But this model solves all these problems.



The scrum process

Figure 5

**Initial User Stories**

In scrum entire system need to describe with user stories. Following are the user stories what we identified.

1. Read log files
2. Archive Log files
3. Summery table uploading schedules
4. Spring Integration
5. Initial Services
6. CAS integration with Single Sign On
7. Open ID Integration
8. Create rules based on programs
9. Search rule
10. View and edit created rules
11. Data analyze with map reduce
12. Configure properties
13. Initial Reports
14. Identify Drill down reports and charts
15. Identify necessary indexes
16. Spring Birt integration
17. Test cases
18. Concurrent access
19. Security
20. SSL for User interfaces

Here are the initial user stories. After Identify user stories we need to estimate those above stories.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| User Stories | | | | | | | | | | | | | | | | | | | | |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| Rajith | 3 | 3 | 5 | 8 | 13 | 8 | 5 | 3 | 3 | 3 | 13 | 5 | 8 | 5 | 3 | 13 | 5 | 0 | 0 | 0 |
| Lasantha | 3 | 3 | 3 | 8 | 13 | 5 | 8 | 5 | 3 | 3 | 13 | 5 | 5 | 5 | 5 | 13 | 5 | 0 | 0 | 0 |
| Thejani | 3 | 3 | 3 | 13 | 13 | 8 | 8 | 8 | 3 | 3 | 13 | 5 | 8 | 5 | 3 | 13 | 5 | 0 | 0 | 0 |
| Naveen | 3 | 3 | 5 | 8 | 13 | 8 | 8 | 5 | 3 | 3 | 13 | 5 | 8 | 5 | 3 | 13 | 5 | 0 | 0 | 0 |
| **Final Extimation** | 3 | 3 | 5 | 8 | 13 | 8 | 8 | 5 | 3 | 3 | 13 | 5 | 8 | 5 | 3 | 13 | 5 | 0 | 0 | 0 |

How collect data?

To build up the log file details such as sender’s mobile number, receiver’s mobile number, the location it required specification of the mobile company. So we receive it from them. Here they maintain some log files such as voice calls, SMS, GPS, LBS and USSD. But when sends SMS to suitable customers they only consider about one log file. But in proposed system it concern about five log files when select the customer groups to send SMS. Also to analyze records has to have a method. So the system uses hadoop integration to do that task. We have to take this and customize according to our system.

Software Requirements for Development

* Net Beans – Here it uses Net Beans 7.0.1 version. The system develops by using java language.
* My SQL

Graphics

* Microsoft project 2007 - This use to develop Gann chart.
* Corel Draw 12 – Use to draw all diagrams of the report.
* Photoshop - To design the interfaces of the system

Technology

* To analyze data use hadoop integration. It has to customize according to our system.

Anticipated conclusions

If consider the real world, the system mainly useful for mobile companies, brand companies and mobile users. When requirements come from brand company to mobile company they select suitable customer group according to requirements to serve their service by providing SMS. In the system it mainly concern about the data analyzing, because the system base on this. This is the most difficult section, because here has to concern about millions of records. Other than this task the system able to provide reports when necessary and the system able to send SMS to mobile customers.

**Description of Personal and Facilities**

By considering each and every fact of the project, it distributes the objectives to each member. To reach this requirement the project split in to modules. So each member assigns with each module and has to complete their module from the beginning to the end. At the end all four modules should be merged together to perform the final product.

According to the project plan, the modules of each group member can be categorized as follows.

|  |  |
| --- | --- |
| Name of the member | Tasks |
| Naveen | * Manage five log files * Log files for Voice call, SMS, USSD, GPRS, LBS * To take customer details * Develop the database to store customer details. * Maintain two files for each log file. * File one – To get data. * File two – Send data after send records to database. |
| Thejani Dineshika | * Develop Graphical User Interface for administrator. * To register * To log to the system * Develop Graphical User Interface for the options. * Promotions * Loyalty programs * Tenure program * Generate the rules for execution. |
| Rajith Delantha | * Execute the rule when necessary. * Use rule engine * Analyze records. * By hadoop Map Reduce * Send SMS for appropriate users. |
| Lasantha Perera | * Develop to generate reports * Contain three reports * For one report contains drill down report * Provide current status. * Provide history reports. |

Budget

References

Appendices

# *Appendix A:* *List of Acronyms and Abbreviations*

CDR – Caller Detail Record

SMS - Short Message Service

LBS - Location Base System

GPRS - General Packet Radio Services

USSD - Unstructured Supplementary Service Data

WBS – Work Break Structure

# *Appendix B: Diagrams and Figures*

**Figure 1**

This diagram represents how the entire proposed project is going on.

* Gather customer records to send to database from log files. Each log file contains two files.
* When take customer requirements it generates the rule.
* Then executes the rule to select matching customers from database
* Send SMS.
* Print reports.

**Figure 2**

From database it takes data and generates three reports. One report contains drill down report.

**Figure 3**

Here it is the plan of the system for one year. It contains the all activities regarding with the proposed system and the time period allocate for them.

**Figure 4**

The Work Break Structure represents the scope of the system. It shows the major contributions of the system and the functions of them. Under those functions it represents the activities relate with it.

**Figure 5**

It is the scrum process which relates with agile methodology.