Medical Cyber Physical Systems and Bigdata Platforms

Don.S
Department of Computer, Information and Communication Engineering
Konkuk University
S.Korea
donsasi@konkuk.ac.kr

Dugki Min*
Department of Computer, Information and Communication Engineering
Konkuk University
S.Korea
dkmin@konkuk.ac.kr

ABSTRACT
Over the past few decades the capabilities of adapting new class of devices for health monitoring system have improved significantly.But the increase in usage of low cost sensors and various communication media for data transmission in health monitoring have lead to a major concern for current existing platforms i.e., inefficiency in processing massive amount of data in real time. To advance this field requires a new look at the computing framework and infrastructure. This paper describes our initial work for Bigdata processing framework for MCPS that combines the real world and cyber world aspects with dynamic provisioning and fully elastic system for decision making in health care application.

Categories and Subject Descriptors
C.3 [**Special-Purpose and Application-Based Systems**]: Real-time and embedded systems; H.4 [**Information Storage and Retrieval**]: Metadata; D.2.8 [**Software Engineering**]: Metrics—*complexity measures, performance measures*; J.3 [**Life and Medical Science**]: [Medical Information systems]

General Terms
Framework, Design, Analysis, Awareness

Keywords
Medical Cyber Physical Systems, Bigdata, Awareness Computing

1. INTRODUCTION
A Medical Cyber Physical Systems (MCPS) are new type of systems that integrate the notion of combining both the physical world aspects with the cyber space. Such systems have grown significant interest in the near future as the current health care system requires a rapid transformation of existing technologies. The systems works with set of embedded devices and control these devices from cyber space with a set of command and control statements. Any changes in physical world directly influence the cyber space. MCPS are commonly used in health care related applications. In the context of remote health care monitoring system, these are small embedded systems that are connected with human body and capture various body conditions, process those informations locally in real times and send only relevant information back to the cyber space through internet or via mobile. Also these datas are in larger volume of size and requires quick response time without much delay. Therefore in the near future it will be an exciting time for the new processing platform such as Bigdata. In this paper,we proposed a Bigdata processing framework for MCPS that incorporates different reusable components which can facilitate dynamic provisioning for loading various functionalities and provide fully elastic to share the data with external source with tight access control.

*Corresponding author

1.1 MCPS for BigData Analysis
Today's medical systems are integration of different classes of devices that can perform various functions in real time. With the emergent of low cost portable devices, monitoring the patient remotely become more common and the number of patients using this become increasingly popular. Storing and processing these data requires good infrastructure and it became huge computational complexity. Therefore there is a great desire to improve the service functionality of current health care monitoring system. The challenges for delivering meaningful information from these data become great demand in health care system. To meet this demand, medical health care industry looked at modernize the existing infrastructure more intelligently. Thus by providing a system that work efficiently,will reduce computation error and fully elastic with on demand. Such system requires various subsystems to perform various functions such as filtering, listening, processing, accelerating and enrichment. Thus establishing a fully awareness system. Studies shows [1] that there has been a growing interest over elderly community to prefer remote healthcare system. There are obvious reason for this.Firstly, due to increased amount of expense in the hospital. Secondly the physical condition of the patient make it very difficult for routine checkup. In remote health care monitoring system, the patient body is connected with various sensors to measure different physiological datas such as ECG, Oxygen level, Hb etc. These datas are then send to the remote application server with-

out any loss of information at the receiving side. This would help the medical professional to analyze the data and take the right decision before the condition of the patient goes severe. Since these sensors are low cost devices, the accuracy for detecting any abnormalities depends on the type of hardware it used in the client side and the method proposed in the application side for data analysis. The limitation of these systems are since the devices are recording the data in a daily manner over a period of time, the storage system stores huge amount of datas that include both significant and non-significant data. It will be a tedious execution process for the system to extract the meaningful information from these massive data set. Figure[1] shows the typical scenario diagram for health monitoring systems. For example,if we consider a patient ECG data, it would contain significant information regarding the condition of the heart also normal beat information i.e., less significant for the analysis of heart disease. Each ECG beat waves are represented by PQRST patterns.Where P wave represents depolarization and contraction of the atrium. QRS complex represents the depolarization and contraction of the ventricles and the T wave represents repolarization of the ventricles. The normal QRS duration for each beats are $0.04 \sim 0.12$ sec. QT interval is 0.39 sec. And the normal PR interval range is between $0.12 \sim 0.20$ sec. Variation in any of these can lead to abnormal functioning of the heart. In addition to the limitation mentioned previously i.e., the system requires larger volume of data storage capabilities with increased processing time. This will reduce the performance of the system. Secondly, the type of service provided in the system may not be dynamically configurable to the current scenario, i.e. the system lacks elasticity. Thirdly, there will be lack of awareness between the cyber space and the real world, i.e. the system is not fully adaptable to take decision in real time. Cyber Physical Systems (CPS) is the new paradigm to solve the aforementioned issues. CPS is new type of systems that integrate computing and communication capabilities with the monitoring and control of entities in the physical world. Cyber Physical Systems are commonly used in mission critical scenarios such as military [2], robotics, health care systems, space research etc. The system requirements for developing CPS systems are differ from traditional health monitoring systems. Cyber Physical Systems (CPS) are integrations of embedded devices, networks and physical world aspect[10]. These embedded system have high reliability and thus makes the CPS system differ from traditional monitoring system. In traditional health monitoring system the systems are designed to operate in a controlled environment where the system operate with a set of predefined rules and semantics. For example in the case of video monitoring of a patient, the system will track the objects that come to view area irrespective to the situation. With the help of small embedded software installed in the physical world monitoring components, the CPS can handle and track any unexpected event more intelligently than the traditional systems.

2. MEDICAL CYBER PHYSICAL SYSTEMS

With the rapid transformation for various medical systems, there is a strong requirement for new devices with increased functionalities. The term Medical Cyber Physical systems refers to systems that has combination of embedded devices, software for controlling these devices and communication channel for interaction [3]. For developing safe and effective MCPS requires new design, verification and evaluation techniques due to increase in size and complexity. And the challenges for developing these kinds of systems include executable clinical workflow, model based development, physiological close-loop control, adaptive patient specific algorithms, smart alarms and user centered design and infrastructure for medical integration and interoperations. The application scenario for MCPS varies from patient monitoring, analgesic infusion pumps to implant sensors devices. Cyber physical medical system modeling and analysis is a framework proposed by [4] for safety verification of different applications. The scenario considered for the experiment was analgesic infusion pump control algorithm for keeping the drug concentration in the blood for a fixed level. These systems are example for typical closed loop systems[5]. Any change in the physical world can directly influence it in cyber world and an action will be taken at the physical world based on the instruction given from the cyber space.

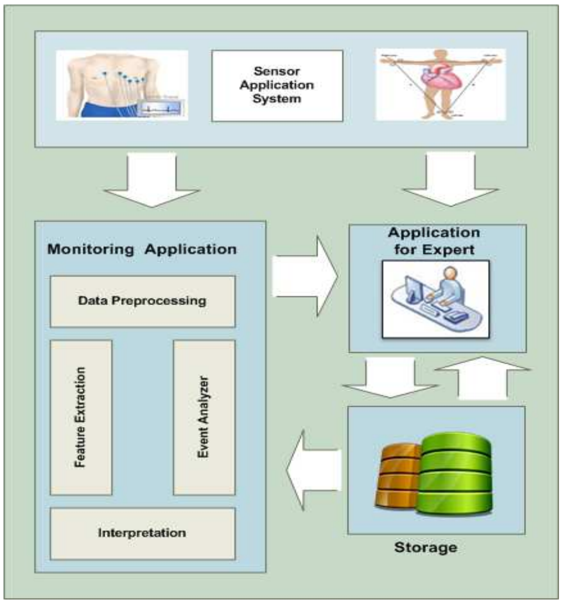

Figure 1: Health Monitoring System Scenario

2.1 Big Data Processing Platforms

Health care monitoring systems are generating loosely structured data from different sensors that are connected to the patient over a period of time. And these are large complex system requiring efficient algorithms to process these raw data's and require huge computational power. Big data[6] refers to the data generated from different sensors this includes medical, traffic and social datas. Some of the characteristics of big data are volume, velocity and value.

- **Volume:** The amount of data generated by the various medical equipment's are larger in size compared to traditional data.

- **Velocity:** The amounts of data stream by medical network are much less compared to the annual data storage capacity of an entire hospital system.

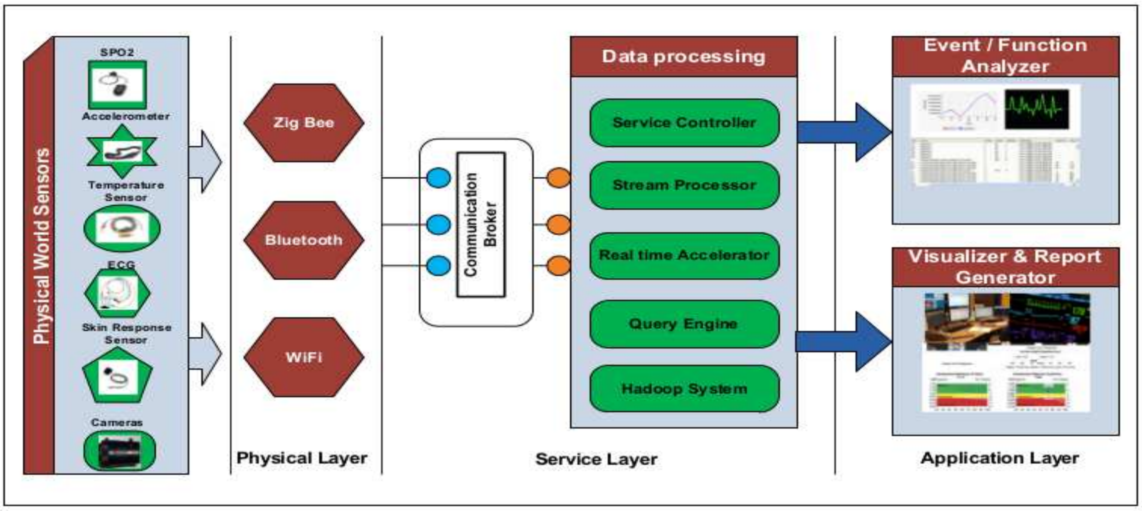

Figure 2: Bigdata Platform

• **Variety:** Traditional data format are less to adapt new type of sensor data types, deployment etc. Whereas non-traditional data such as medical equipment's are easily adaptable to change.

For developing an infrastructure for big data analysis, it requires a mechanism on how to acquire the data, organize these data and process it to extract meaningful information [7]. This can be represented as data acquisition, data organization and data processing.

• **Data Acquisition:** Data acquisition is one of the major challenges in the big data platforms. Since these systems handle large volume of data, the system requires low latency in capturing the data and using simple query to process larger volume of data.

• **Data Organization:** Since the data's are of larger volume of size, the system needs to take and process the data from the original storage location. Apache Hadoop provides a technology to process these larger volumes of data and at the same time keeping the data on the original data clusters.

• **Data Processing:** In big data processing the data must be process in a distributed environment. The requirement for analyzing data such as medical information requires statistical and mining approach for analyzing the data. Delivering the data in a faster response time will be at higher priority.

Figure.2 shows the pipe-line architecture of different components that are required inside the Bigdata platform for any medical application systems.The monitoring system consists of various physical world sensors for capturing various biological signals from the patient's body. These include camera sensors for monitoring the motion of the patient. Other various sensors that are attached to patient include ECG for measuring the physiological behavior of the patient, temperature sensors for measuring the body temperature, SPO2 sensor for evaluating the oxygen level of the body and accelerometer sensor for analyzing the body motion. The communication between the client device and the remote application can be established either using Bluetooth, ZigBee or WiFi communication channel. All together call it as components of physical layer. Next one is the service layer where the end machine will provide various services that an application requires. To manage the data received from different patient over a period of time, the communication broker in the service layer act as a mediator between the client and server side for managing the received data, ie. sending to the data processing module for processing and also sending important information back to the client side. The Hadoop system in the service layer is made to scale up and distribute the data into different node, so that the system can process it quickly for a larger collection of data. There are several advantages for introducing Hadoop in Bigdata platform. First it allows the data to be processed in a distributed fashion. It can handle and manage failure of nodes. The data can be processed in any node.Figure.3 shows the hadoop data distribution. Service controller module will manage all the services related to data distribution, protocol , routing etc. It has the capability to dynamically configure the system without interrupting other services. Since the datas are receiving as a stream of datas , the stream processor will have the capabilities to manage various stream of data and process it based on the type of method the platform introduced. Real time accelerator module is used for accelerating the processing time by introducing various window line for the stream processor. When the user queries a request, it will process with the help of query engine. These engines can translate the semantic queries into machine understandable form and retrieve the data accordingly. The application layer contain various user interaction module such as analyzing the events, visualize the important event generating and reporting.

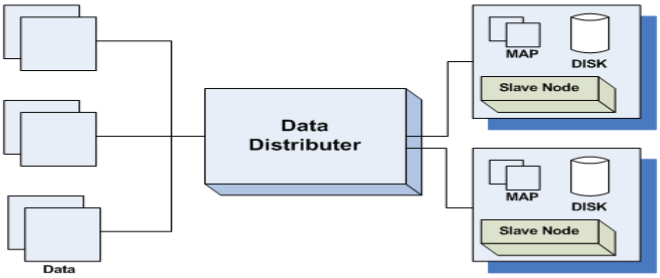

Figure 3: Hadoop Data Distribution

3. FRAMEWORK FOR BIGDATA STREAM PROCESSING

Traditional health care monitoring system are more focused on specific services. For example analyzing patients various physiological conditions in the hospital. However with the increased efficiency in data transmission and technology has made a growing interest in remote home monitoring system [8]. With the increased connectivity of different medical devices, changes continuously the system functionality which makes difficult for the current system to handle on-demand. To handle these on-demand with the current technology we proposed a Bigdata processing framework for medical cyber physical systems that combines the knowledge of both physical and cyber space aspects. The framework for the proposed health monitoring and managing system is depicted in Figure 4.

This framework is a multi-layer framework where each layer provides certain services to the upper layer. The three different layers are component layer, process layer and application layer. Component layer is the level1 in the proposed architecture. This layer provides messaging and distribution services that are required between the system and the application. This level is important for framework construction, as it provide extensive set of functionalities; call it as infrastructure for running an application service. For example how to configure the system to access the datas from different sources and distribute to process layer. These components have the advantage of dynamic reconfiguring at the service level. The component provides functionalities such as routing, protocol,data access and data distribution. Data accesses are means of getting the data from external sources. And these datas are then transferred to the process layer. The services inside the process layer are message listener, message distributer, message processor, accelerator and the awareness module. Message listener listens to different type of sensor datas coming from the component layer as streams. Once it received the datas, it will preprocess the datas to remove unwanted information and consider the resultant data as a stream of events. We call it as messages. These events will have a hierarchical structure so that it will not overlap with other incoming data stream.Data stream processing infrastructure differs from traditional data processing as it is mainly used for processing high volume of data in real time [9]. The main objective for any big data processing platform is to achieve low latency. The data must be processed as a stream rather than store and process mode. The basic requirements for real time stream processing is to

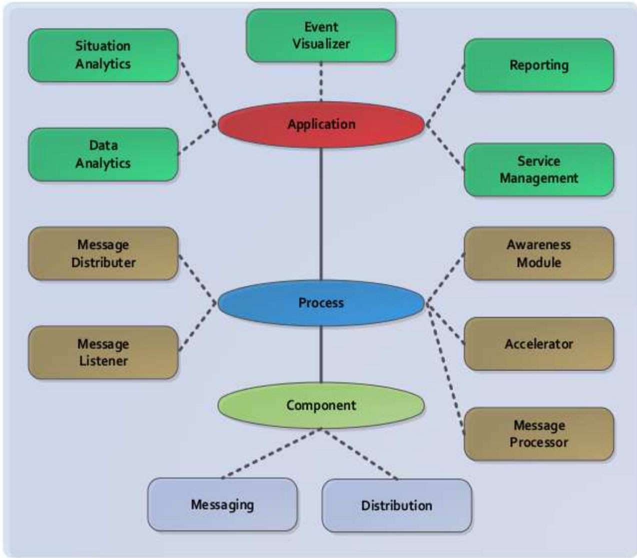

Figure 4: Bigdata Framework for MCPS

process the data 'in-stream'. Another useful requirement is to query these data over continuous stream of data, so that the system can get certain meaningful information whenever needed. Since data stream cannot specify the file ending as it is receiving the data in real time, we need to define a window range over a period of time. And each time process these data inside the window's range. Figure.5 shows the representation of stream processing inside the message listener.

The message distributer distributes the datas into the message processor module based on the arrivals. At this stage the message distributers categorize the pre-processed datas into different groups. The main requirement here is the ability of the system to group different datas into an order. For that we introduced time window. The data will then process it based on the time window. The reason for introducing the time window inside the message distributer is that in real time the data is not in store and process mode. So every grouping must be done within the duration. In some cases the message distributer may receive more data than the fixed time window. In that situation based on the size of the received data volume the window closing duration time will extend. The next main component is the message processor. Here happen the actual execution of various stream events. Each type of events based on the characteristics will have its own message processors. For example if we have ten different groups of datas that we need to process,then the system will have ten different message processors to process it. This will help to improve the performance of the execution. Figure.6 shows how the message processor executes the events.

The received datas are first preserved in appropriate message processing stacks. The messages are then mapped by the message mapping module for processing. These messages are then processed with the help of worker thread. Message analyzer module will analyze each messages based on the

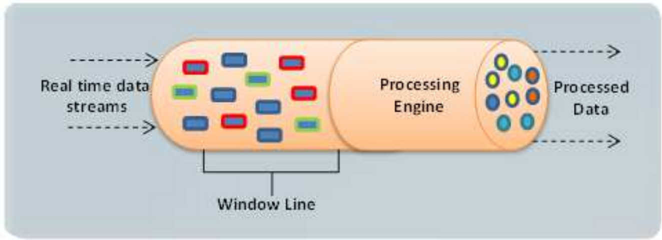

Figure 5: Stream Data Pre-Processing

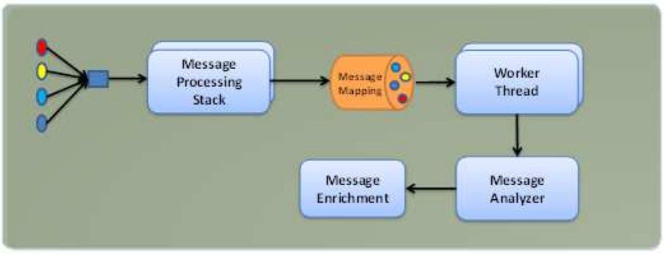

Figure 6: Message Processor Execution Flow

rules defined inside the message analyzer, finally the message enrichment module will identify the importance of each message from the message analyzer. In big data streaming process, handling all the information output by the message processor over a period of time will increase the communication latency between the awareness module and the message processor. To overcome this we introduced the accelerator module. The accelerator module decreases the latency by storing the data as in-memory. This would increase the communication performance. The last component inside the process layer is the awareness module. This module combine the knowledge obtained from the message processor with the semantic knowledge inside the awareness module.

This semantic knowledge is written in the form of ontologies. The awareness module act as knowledge enricher. From the message processor we can retrieve only low level information from each data. With this we cannot fully understand the behavior of the system. Thus by combining both message processor and the awareness module, the framework can provide the full functionality that a big data processing platform requires and use in various medical domains. The application layer provides services such as data analytics responsible for analyzing the data characteristics. Situation analytics responsible for handling important events occurred in the physical space. Event visualizer module will visualize the type of event happening at the application layer. Service management monitors and handles the top level services the system requires. With the help of human the system will report the important event based on the observation and future examination. Through different test scenarios the framework can explore and examine various actions that happen in the physical space and take necessary steps in due course. The system examines the current situation and dynamically enriches the knowledge based on the current knowledge gained with the semantic data stored in the storage. We believe that this kind of platform can improve the decision making in any real time scenarios, because the framework is adaptable and reconfigurable based on the application and the situation.

4. CONCLUSION

Currently available systems for health care domains limits the functionalities due to the ever increases in demands. As a result the integration of new technologies is necessary to cope up with on-demand. To solve this problem, we designed a framework that can dynamically handle the situations in real time. The MCPS system introduced in this paper gives a new paradigm for integrating various cyber and physical aspects and thus provides a complete solution for handling various clinical issues such as examination, prediction and handling larger volume of datas in real time. In the future, further studies will be conducted for improving the big data platforms through theory and experiments.

5. ACKNOWLEDGMENTS

This paper was supported by NRF International Research Collaboration Program (2011-0030918) from National Research Foundation (NRF) of the Ministry of Education, Science and Technology of Korea. This research was supported by the MKE(The Ministry of Knowledge Economy), Korea, under the ITRC(Information Technology Research Center) support program (NIPA-2013-(H0301-12-4014)) supervised by the NIPA(National IT Industry Promotion Agency)

6. REFERENCES

[1] Pantelopoulos, Nikolaos G, A survey on wearable sensor-Based Systems for health Monitoring and Prognosis, IEEE Tran Sys Man and Cybernetics, vol 40(1), pp: 1-12, 2010.

[2] Zhoung Liu,Dong-sheng Yang,Ding Wen and Wei-ming Zhang,Cyber-Physical-Social Systems for Command and Control,IEEE Intelligent Systems,Vol 26,no.4,pp 92-96,2011

[3] Lee Insup and Oleg Sokolsky, Medical Cyber Physical Systems,in Proc.of DAC,USA 2012.

[4] Banerjee, Ayan and Gupta, Sandeep K. S. and Fainekos, Georgios and Varsamopoulos, Georgios,Towards modeling and analysis of cyberphysical medical systems,Proc, ISABEL'11,pp:154-158,2011.

[5] Arney, David and Pajic, Miroslav and Goldman, Julian M. and Lee,Insup and Mangharam, Rahul and Sokolsky, Oleg, Toward patient safety in closed-loop medical device systems, ICCPS'10,pp:139-148,2010.

[6] Sam Madden, From Databases to Big Data, IEEE Internet Computing, Vol 16, No 3, pp 4-6,2012.

[7] Dean, Jeffrey and Ghemawat, Sanjay,MapReduce: simplified data processing on large clusters,OSDI,2004.

[8] Gilles Virone, Majd Alwan, Siddharth Dalal, Steven W. Kell, Beverely Turner, John A. Stankovic, and Robin Felder, Behavioral Patterns of Older Adults in Assisted Living,IEEE Tran InfoTech in BioMed,Vol 12,No.3,pp:387-398,2008.

[9] Michael Stonebraker, Cetintemel U, Zdonik Stan,The 8 requirements of real time stream processing, SIGMOD pp:42-47,2005.

[10] Lee, Edward A,Cyber Physical Systems: Design Challenges,ISORC,pp 363-369,2008.