The Internet of Things based Medical Emergency Management using Hadoop Ecosystem

M. Mazhar Rathore, Awais Ahmad, Anand Paul
The School of Computer Science and Engineering
Kyungpook National University
Daegu, Korea
rathoremazhar@gmail.com, awais.ahmad@live.com, paul.editor@gmail.com,

***Abstract*—The prevalence of Internet of Things (IoT) in medical health care is bound to generate the massive volume of heterogeneous data due to the millions of medical sensors attached with various patients' body. Therefore, to process such amount of heterogeneous data in real-time to take emergency actions in critical health situation is a challenging task. Therefore, to address such issues, we proposed Hadoop-based medical emergency management system using IoT technology, which involves a network architecture with the enhanced processing features for collecting data received from millions of medical sensors attached to the human body. The amount of collected data is then forwarded to the Intelligent Building to process and perform necessary actions using various units such as, collection unit, Hadoop Processing Unit (HPU), and Analysis and decision unit. The feasibility and efficiency of the proposed system are evaluated by implementing the system on Hadoop using UBUNTU 14.04 LTS coreTMi5 machine. Sample medical, sensory datasets and real-time network traffic are considered to evaluate the efficiency of the system. The results show that the proposed system efficiently process WBAN sensory data.**

Keywords*— Big Data; IoT; Healthcare; Intelligent Building;

I. Introduction

Recently, Healthcare system in IoT recognized as a revolution in ICT since the rapid development has been started at the beginning of 21st century. In Healthcare system's applications, the wireless body area network (WBAN) offers a novel archetype for wireless sensor networks (WSNs) in monitoring biomedical sensors. These sensors can be attached to human body or clothes and can be used to measure the parameters associated with the human body. The measured values can be collected and relayed to the main server using Internet Protocol Version 6 (IPv6) over Low-Power Wireless Personal Area Network (6LoWPAN) [6] by passing through some gateway node in order to generate results. For such application, a suitable example, i.e., ZigBee technology can be employed that uses IEEE 802.15.4 [7].

Many IoT based architectures have been proposed to monitor the health of the people using IoT. Yang et al. [8] proposed iHome open Health-IoT platform based on intelligent medicine box and pharmaceutical packaging (iMedPack). Some other approaches intend to integrate wearable devices to achieve better IoT for the e-healthcare system [9]. However, in any IoT based medical care system, the amount of data collected from the sensor nodes should be accessible anytime and everywhere that requires continuous network connectivity, which results in generating massive volume of data (called as 'Big Data'). According to GSMA, the total number of devices connected to each other will be 24 billion until 2020 [10]. In such circumstances, the Healthcare system will face a critical challenge of available bandwidth.

On the whole, the previously proposed techniques only cover a limited scope of the Healthcare. Moreover, these techniques are not that much efficient to handle the massive volume of heterogeneous data to generate real-time action in case of emergency. Therefore, a desirable system is required that is capable of taking care of the human beings (patients in home, people outside home, and in cars) from all the features, which not only covers the required medications, but can also perform a continuous monitoring of the patients, car accidents, and remote diagnostics and taking actions at real time. The mentioned challenges can only be catered through fast collection and aggregation, parallel and efficient processing of incoming high-speed medical, sensory data. For this reason, Hadoop-based medical emergency management system using IoT technology is proposed. In the proposed system, the human body uses wearable devices or other physical body sensors that measure blood pressure, pulse/heart rate, diabetes, etc. The measured data is then transmitted to the primary medical device (PMD) using Bluetooth or Zigbee IEEE 802.15.4, then to Intelligent Building (IB) through 3G/LTE/WiFi Internet via gateways. IB provides the backbone of the proposed system that process massive volume of the incoming stream of data by capturing and aggregate the data in its Collection Unit. Afterward, the collected data is sent to Hadoop Processing Unit (HPU) for further processing and computations. Finally, the Analysis and decision unit responds to the system based on the results generated by HPU. The main contribution of this work is summarized as follows. The system Architecture is proposed and implemented using Hadoop to process and analyze the massive volume of medical, sensory patient data at real time. Moreover intelligent building concept is introduced, which is mainly responsible for managing, processing, analyzing incoming sensor data and finally make decisions intelligently. This does not only handles the massive volume of data but also gives feedback to the users "anytime-anywhere-anyhow". The whole system is implemented in a real environment using Hadoop on UBUNTU 14.04. Sample medical, sensory datasets are tested to evaluate the proposed system.

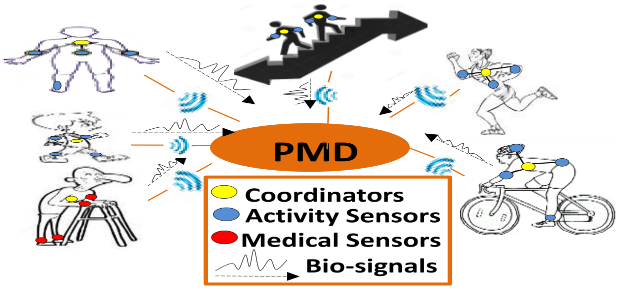

Fig. 1. Deployment scenario for proposed system model

II. HADOOP-BASED PROPOSED SYSTEM

In this section, we deliberate the proposed scheme, which includes sensor deployment describing the functionalities of the medical and activity sensors, followed by the network architecture including intelligent building, and the algorithm.

A. Sensors deployment and Data aggregation

Various sensors including activity sensors, medical sensors, and coordinator sensors are attached to human body parts, as shown in Fig. 1. The human body parts include the wrist, ankle, heart, chest, and a helmet (while cycling). These sensors collects heterogeneous medical data, such as, diabetic and blood pressure data (elderly age people), body temperature, heartbeat, blood pressure, sweating, glucose level, and other medical measurements while performing various activities, such as, physical exercise, sitting, walking, climbing upstairs or downstairs, and cycling. In order to collect and aggregate data from various sensors and making system energy efficient, we used coordinator, which works as a sink node or relay node that collects the data and transmits it to the primary medical device (PMD) through Bluetooth or Zigbee.

The coordinator receives health readings from various sensors attached to the body, aggregates them and then encapsulate them into a single block in a proper format and sequence. It adds the packet header, which mainly contains the GDID (globally unique device identified) and U_ID (user id), which uniquely identifies the sensors and the user (or patient). Later, it enlarges it by adding various sensors readings in the following sequence, depending upon the available readings: 1) Glucose level reading 2) Blood pressure reading 3) Pulse rate 4) Temperature 5) Heart rate and 6) Breath rate. Since every sensor is not required to transmit data directly to PMD, therefore, the energy is saved. Fig. 2 shows the various packet structures at coordinator node depending on the number of sensors attached to the body.

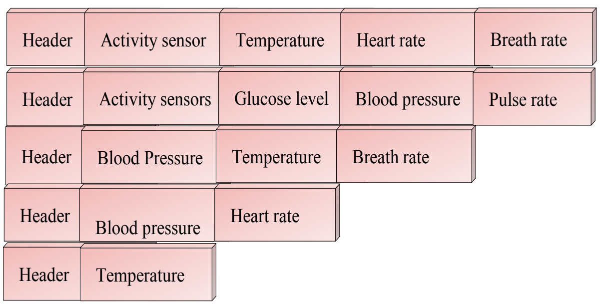

Fig. 2. Various packet structure of sensors readings sent by the coordinator

B. Proposed architecture

Fig. 3 illustrates the concept of the proposed IoT based Medical Emergency Management. We assume that all the persons are equipped with smart devices. In the proposed scenarios, the sensors, attached to the patient body, transmit the measured data to the agent (e.g., raspberry-pi). Raspberry-pi is a device used to convert sensory data to mobile readable data. After the conversion, the mobile readable data is forwarded to the primary mobile device (PMD) using ZigBee or IEEE 8.2.15.4. PMD is attached to the Intelligent Building via Internet (LTE/3G/WiFi). Intelligent Building is a smart block used for storing, processing, and execute certain actions depending on the context of data.

Intelligent building is the central component with complete intelligent system that handles incoming high-speed Big Data using parallel processing by Hadoop Ecosystem. The system handles capturing, collecting, processing, and analyzing health sensors data from a large number of people with body area network (BAN) by continuously monitoring their health parameters. The intelligent building, we can say intelligent medical emergency management system, mainly composed of collection unit, processing unit having Hadoop system equipped with intelligent medical expert system, sensor health measurement patient database, aggregation result unit, and finally the application layer services.

Collection unit is the entrance point, which continuously collects the data from each registered person of the BAN network. It captures data using high speed capturing devices such as RF_RING and TNAPI [11]. It is composed of a single server performing all functionalities of the collection, filtration, and load balancing or have multiple servers for each functionality depending on the complexity of the system. We considered the multiple servers for collection unit in which collection server collects high-speed incoming WBAN sensors. It extracts the required information from each packet such as GDID, U_ID and all sensors measurements squeezed in a single packet. Moreover, it filters all necessary data by discarding all unnecessary repeated readings. It uses Hadoop libraries such as Hadoop-pcap-lib, Hadoop-pcap-serde, and Hadoop Pcap Input to process the incoming packet and generate sequence file corresponding to each user by extracting sensors' readings encapsulated in the packet. When the sequence file is reached its size or time threshold, it is sent to the Hadoop processing units to process the sequence file by analyzing and calculating statistical parameters. The load balancer, which is sometimes the master node in the Hadoop ecosystem, decides which data nodes will process the sequence file. Each processing server or data node has its GDID range for which it processes sequence file.

Hadoop processing Unit is composed of various master nodes and many data nodes. It uses Hadoop Distributed File System (HDFS) on multiple parallel data nodes to store the data in blocks. Each data node is equipped with same proposed algorithm using MapReduce implementation to process the sequence files either by calculating statistical parameters or by analyzing sensors measures in sequence file to generate intermediate results for decision-making. Various Mapper and Reducer functions work in parallel to achieve the efficiency.

The mapper function initially decides whether each sensor reading is normal and need not be analyzed deeply or have some abnormal values that require analysis and emergency actions. It compares sensors values with their corresponding normal threshold values, and if satisfied, then just stored them in the database without further time-consuming analysis. However, when any sensor value from WBAN sensor satisfied its corresponding serious threshold value such as BP: below 40 and above 200, diabetes: below 40 and above 400, pulse above 100 for man and above 130 for a kid, etc., it generates an alert directly to application layer service for a quick response. Application layer perform quick action depending on the sensor, its value, and patient such as, call police in case of accidents, call doctor, ambulance in case of heart attack, serious diabetes, BP, pulse reading. However, when the sensor's reading is neither normal nor too dangerous, it requires analysis. The readings which lies in mentioned range are processed by calculating statistical parameters or performing other calculations depending on the algorithm to generate intermediate results for a final decision. Finally aggregation unit of Hadoop processing system aggregates the results using Reducer from various parallel processing servers (data nodes) and send them to the final Decision Server.

Decision servers are equipped with the intelligent medical expert system, machine learning (ML) classifiers, and other complicated medical problem detection algorithms for further analysis and decision-making. It analyzes the current results, received from the processing unit, depending on the previous history of the patient using complex medical expert systems, machine ML classifiers, etc. The concrete detail of the medical expert systems is out of the scope of the paper due to its vastness. In out scenarios, we are using machine learning classifiers like REPTree, which is more efficient and accurate, at decision server for various normal disease detection.

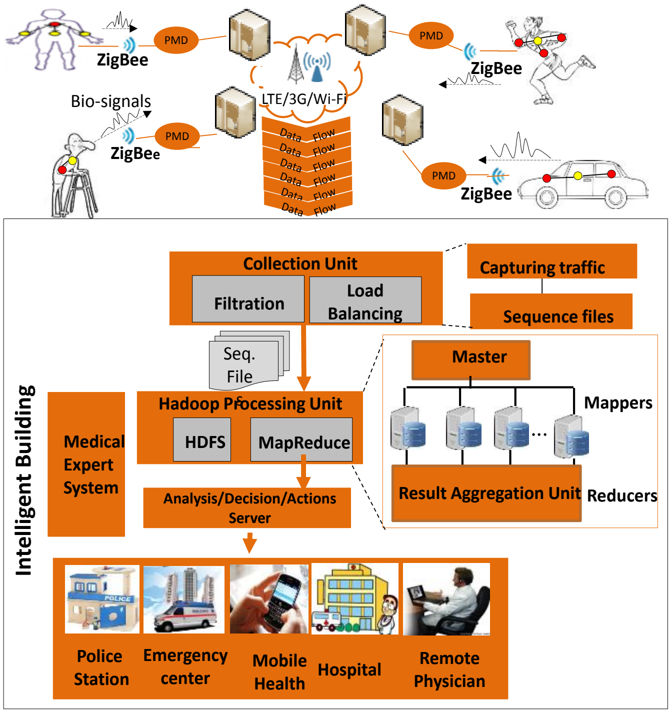

Fig. 3. Application scenario for proposed Healthcare IoT System

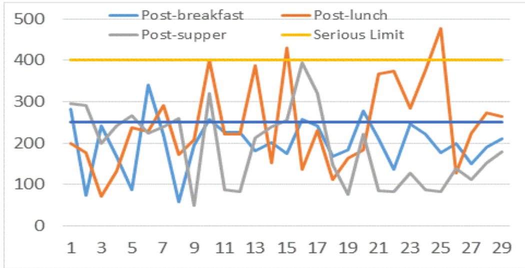

Fig. 4. Post-meal diabetes measure of a patient

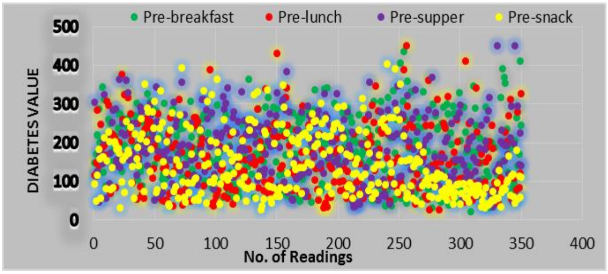

Fig. 5. Pre-meal diabetes measure of a patient

III. SYSTEM IMPLEMENTATION

The proposed system is implemented using Hadoop single node setup on UBUNTU 14.04 LTS coreTMi5 machine with 3.2 GHz processor and 4 GB memory. MapReduce is used as a front-end programming with Hadoop-pcap-lib, Hadoop-pcap-serde, and Hadoop Pcap Input libraries for network packets processing and generating sequence file at collection unit. The Map function of our implementation maps U_ID and corresponding sensors values. It compares the values with thresholds and generates action or alerts. Moreover, it generates intermediate results and sent it to Reduce function as U_ID as a key and results as value at aggregation unit. The Reducer aggregates the results, sort and organize them. Finally, the decision are made based on the results. Hadoop MapReduce implementation of the whole system on various data nodes makes the algorithm be processed in a parallel environment efficiently.

Three UCI Datasets are collected, i.e., diabetes dataset [12] ICU dataset [13] and WISDM lab activity datasets [14] for evaluating the system. Diabetes dataset contains the 13437records of diabetes outpatient in various time slots. ICU dataset holds the ICU sensed data of an 8.5-month-old biliary atresia deceased kid, which contains the 7931 record heart rate (bpm), respiration rate (breath/ min), Arterial pressure – Mean (mm Hg), Arterial pressure Systolic (mm Hg), Arterial pressure Diastolic (mm Hg), Arterial O2 saturation (%), tidal volume, PIP (cm H20), etc. WISDM Lab dataset contains the measures corresponding to the user activity such as jogging, walking, upstairs movement, downstairs movement. Moreover, health care measures from dataset files are replayed to the system as network traffic to check the real-time efficiency of the system.

Considering the diabetes patient, blood glucose measures at regular post-breakfast, post-lunch, post-supper, pre-breakfast, pre-lunch, pre-supper, and pre-snacks timings are taken for analysis. The graph of the first 30 blood glucose measures of the patient after the meal is shown in Fig. 4. The patient is insulin dependent diabetes patient. Thus, the system considers 250 as an average threshold. However, when the received measure crosses the serious threshold, i.e., 400 as at 25th and 15th readings, the emergency action is taken either admitting the patient in the hospital or requesting him to increase insulin dose. While analyzing the post-meal patient measures, most of the post super readings crosses the average threshold but rarely after- breakfast. For this reason, the patient is asked to change the food for dinner and also increase the insulin dosage. Moreover, the Pre-meal diabetes analysis is also performed as shown in Fig. 5. Nine times, the glucose test crosses the severe threshold when quick action is taken in which patient are suggested to increase insulin dose. Most of the abnormal actions are at pre-lunch and pre-supper timings, which shows the rise in glucose level at PM timings.

IV. System Evaluation

The proposed system mainly focus on the processing of the high-speed, large amount of WBAN generated data. For this reason, the assessment of the system is done by considering the average processing time to process one record of various datasets. The efficiency results, as shown in Fig. 6, the average processing time is less than 15ms. Most of the dataset contains less than 3ms average processing time for 1 record. Since the diabetes datasets contain records from serious diabetes patient and the size of the record file are too small as compared to other files therefore lot of input, output, and switching is done because of MAP and Reduce function. For this reason, the processing time is higher, nearly 15ms per record.

Furthermore, when we increase the number of sensors per record (per person or packet), the processing time is also going to be increased, as illustrated in Fig.7. The rise in the average processing time is because of the increase in overall threshold comparisons due to the enlargement in some sensors.

V. Conclusion

In this paper, we proposed medical emergency management system based on Hadoop ecosystem using IoT technology. The proposed system involves in the different aspect of Hospitals, emergency services, first aid, and police stations. The proposed system provides a promising solution for suppositories nonacquiescence issues by automatically remind the persons about their prescriptions, helping and supporting them on various occasions (e.g., first aid, remote physician, police station, etc.). With a view to developing a continuous follow-up and monitor the persons vital sign's (anytime-anywhere-anyhow), a flexible system is designed, which is based on the Intelligent Machine. Intelligent Machine receives data from various persons, processes and analyzes it using Hadoop and generates output for decision-making. Based on the output, the machine executes individual actions (e.g., first aid, remote physician, reminding patient about the doctor prescriptions, car accident, and so on so forth). The performance of the system is tested on Hadoop using UBUNTU 14.04 LTS coreTMi5 machine with 3.2 GHz processor and 4 GB memory. The evaluation shows that the performance of the proposed network architecture fulfills the required desires of the persons connected to it, whether the input data is a real-time as well as offline while taking actions at real time.

Fig. 6. Average processing time of the proposed system on various Datasets

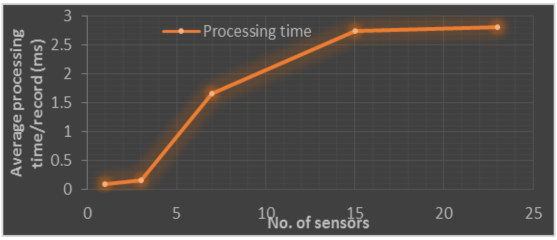

Fig. 7. Efficiency of the system with respect to number of sensor per record

Acknowledgment

This study was supported by the Brain Korea 21 Plus project (SW Human Resource Development Program for Supporting Smart Life) funded by Ministry of Education, School of Computer Science and Engineering, Kyungpook National University, Korea (21A20131600005). This work is also supported by the IT R&D Program of MSIP/IITP. [10041145, Self-Organized Software Platform (SoSp) for Welfare Devices].

References

[1] N. Kushalnagar, G. Montenegro, and C. Schumacher, "IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs): Overview, Assumptions, Problem Statement, and Goals," IETF RFC 4919, Aug. 2007.

[2] M. P. R. Sai Kiran, P. Rajalakshmi, Krishna Bharadwaj, Amit Acharyya, "Adaptive Rule Engine Based IoT Enabled Remote Healthcare Data Acquisition and Smart Transmission System" 2014 IEEE World Forum on Internet of Things (WF-IoT).

[3] Geng Yang, Li Xie, Matti Mäntysalo, Xiaolin Zhou, Zhibo Pang, Li Da Xu, Sharon Kao-Walter, Qiang Chen, and Li-Rong Zheng," A Health-IoT Platform Based on the Integration of Intelligent Packaging, Unobtrusive Bio-Sensor, and Intelligent Medicine Box", IEEE Transactions on Industrial Informatics, Vol. 10, No. 4, November 2014

[4] P. Castillejo, J.-F. Martinez, J. Rodriguez-Molina, and A. Cuerva, "Integration of wearable devices in a wireless sensor network for an E-health application," IEEE Wireless Commun., vol. 20, no. 4, pp. 38–49, Aug. 2013.

[5] [ONLINE] http://gigaom.com/2011/10/13/internet-of-things-will-have-24-billiondevices-by-2020/

[6] F. Fusco and L. Deri, "High Speed Network Traffic Analysis with Commodity Multi-core Systems", ACM IMC 2010, Nov. 2010.

[7] [Online]. Available: UCI Machine Learning Repository: Diabetes Data Set, "https://archive.ics.uci.edu/ml/datasets/Diabetes," Accessed on 31 January 2015.

[8] [Online]. Available: UCI Machine Learning Repository: ICU Data Set, https://archive.ics.uci.edu/ml/datasets/ICU," Accessed on 31 January 2015.

[9] [Online]. Available: WISDM Lab: Dataset, "www.cis.fordham.edu/wisdm/dataset.php," Accessed on 31 January 2015.