

## Big data and WBAN: prediction and analysis of the patient health condition in a remote area

Deepak Sethi\*<sup>1)</sup> and Jyoti Anand<sup>2)</sup>

<sup>1)</sup>Department of CSE, Amity School of Engineering and Technology, Amity University, Uttar Pradesh, Noida-201313, India

<sup>2)</sup>Department of CSE, University of Engineering & Management, Chomu, Rajasthan, Jaipur- 303807, India

Received 1 April 2019

Revised 17 June 2019

Accepted 18 June 2019

### Abstract

The present research aims at the development of a disease prediction system using a Wireless Body Area Network (WBAN) and big data. WBAN is referred as a dynamic sensor network that is based on the deployment of sensor nodes (SNs) in or around the human body. This makes it feasible to make biotic measurements such as an electroencephalogram (EEG), electrocardiogram (ECG) and electromyography (EMG) among others on human subjects. Big data is based on cloud computing and the concept refers to wide scale distributed data processing applications that generally operate with a huge amount of data. The developed prediction model works in two phases. First, biotic measurements were made on human subjects through the use of body sensors. Second, the obtained data from human subjects was compared with big data to make disease predictions.

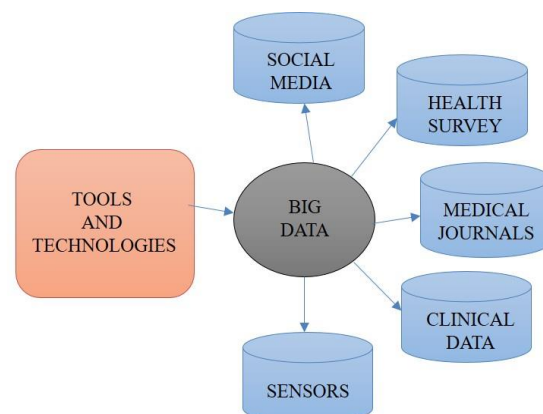
**Keywords:** Wireless body area network, Big data, Sensor nodes, Healthcare, Remote patient monitoring, Energy efficient routing

### 1. Introduction

In the present scenario, the tremendous potential of WBAN and big data can be harnessed for the prediction of many diseases in the healthcare field [1]. These technologies facilitate patient monitoring by analyzing their previous and current health records that may contain information about their blood pressure, heart rate, insulin, and glucose levels among many more. Using big data analytics is quite helpful for the timely prediction of diseases in humans [2-4]. There is an enormous amount of healthcare data that can be used to predict diseases and future health conditions of patients. Big data provides many tools to analyze the data and generate useful information based on the health of a person [1-5].

Generation of data is based on the 5V's: velocity, volume, variety, veracity and value. These are known as the characteristics of big data [6-7]. Velocity determines in what speed the data is being processed, as in case of health care problems, sicknesses tend to increase and so timely health records are important. Volume is the amount of data in GB, TB, PB, and ZB, as electronic health records are generated multiple times and that increases the volume of data. Variety is the type of data. It may be structured or unstructured data in the form of graphs, tables, charts to determine a change in health parameters. Veracity refers to the uncertainty of data for the purpose of making it accurate and of better quality. As data is gathered from various sources, it is necessary to get quality data for better analysis. Value describes meaningful data. Much data is collected, but not all of it is

relevant to a patient's condition. The only data of value is that which can be used to determine that patient has a particular disease. High value data taken by sensors and analyzed in a well-designed framework allows for better prediction of a patient's health. The various sources of healthcare data that may be collected are shown in Figure 1.



**Figure 1** Sources of health care data

WBAN contributes to improve a patient's health, playing a vital role in healthcare applications [8]. It is a wireless technology used to monitor the health condition of patients. This is used to gather real-time information from various

\*Corresponding author. Tel.: +9193 549 41047

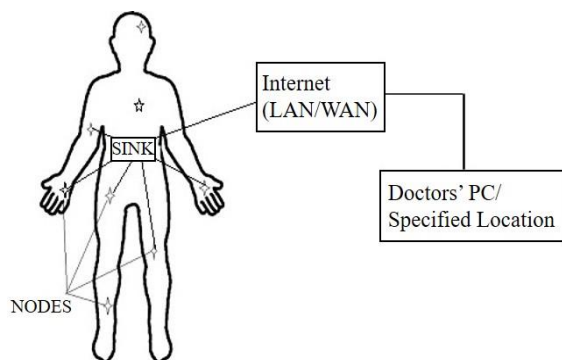
Email address: deepaksethi@live.in

doi: 10.14456/easr.2019.28

sensors. These can be classified as wearable WBAN and implantable WBAN [4]. Wearable WBANs have sensors placed on the surface of the body for a shorter period of time. Wearable WBAN is a fusion of SNs (sensor nodes). IEEE 802.15.4 contains IPv6 over a Low Power Wireless Personal Area Network (6LOWPAN) and other biomedical SNs [9]. 6LOWPAN works on two modes. One is an active mode for transferring biomedical data and other is sleep mode for energy savings [10]. Implantable WBANs are placed deeper within the body such as spinal cord, heart or brain. These monitor the condition of a patient and monitors a critical situation after we proceed for analysis.

## 2. WBAN and Big data

Advancements in topical wireless technology, integrated circuits (IC), electronic devices and autonomous sensing of physiological parameters, gave the birth to short-range wireless communication networks for monitoring patients, known as Wireless Body Area Networks (WBANs) [8,11]. It is one of the major applications of WSN, used in medical and as well as non-medical fields. In WBAN, various tiny, intelligent, low powered and invasive/non-invasive SNs are placed inside or outside the human body to collect data. These SNs are independent devices with limited communication capabilities. To provide real-time feedback to the user, a base station (BS) is placed at the mid-section of the body to aggregate the data and further transmit it to a desired location through the Internet, as shown in Figure 2.



**Figure 2** Deployment of nodes

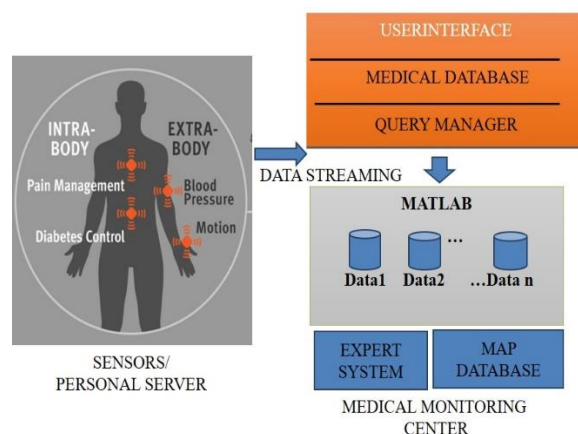
The sink node is supplied with an extensive power supply. Deployed SNs are intended to provide early detection of various types of life threatening diseases viz. cancer and asthma, among others, using different types of biomedical SNs like ECG (electrocardiogram), EEG (electroencephalogram), pulse rate, temperature, and blood pressure [12]. ECG signals check the cardiac status of the human body and require higher bandwidth compared to other signals [13]. There is a recording system to capture the sensed data at various resolutions with different time segments [13]. Mostly, aged and handicapped persons are benefited through remote check-ups, and then doctors prescribed remotely [14]. It also reduces the number of patients in hospitals and reduces health care costs. Diabetes is a common, non-communicable disease that eventually requires control of one's diet and body [15]. From 1980 to 2017, the number of people with diabetes has risen from 108 million to 422 million worldwide. One out of every 3 people have diabetes by age 18 years. This has caused the death of more than a million people. Increased values of parameters

like glucose and insulin causes the diabetic patient to be in a critical situation that can lead to death. WBAN for health monitoring is done by placing sensor nodes in various parts of the body with equal energy, i.e., 0.5 J [16-17]. These SNs sense the various biological parameters such as ECG, EEG, temperature, and blood pressure.

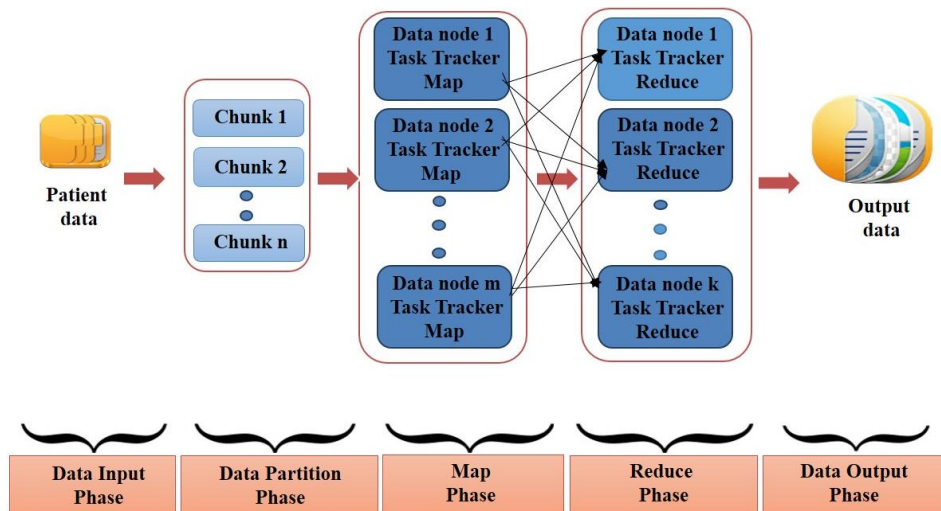
Several methods have been proposed for monitoring patient health conditions using smartphones, wearable sensors and external services. The heart rate and the breathing rate can be monitored in this way. Big data technologies have made it possible for larger datasets of multiple patients to be analyzed in less time with better results. Using these methods provides accurate and useful data. The clusters and correlations, predictive models, statistical and machine-learning models have been very efficient in determining the presence of chronic diseases for better treatment of a patient [18]. Patient care is introduced as "the care that is respectful of and responsive to individual patient preferences, needs, and values" [3]. In the past decade, clinical interactions of patients have been well framed in electronic health records (EHR) which has widened health informatics technologies [14]. One of the most appropriate approaches to gain information about medications of any disease is to tailor knowledge using EHR and give patients personalized access to available facilities. Once available records are collected, further Hadoop and MapReduce techniques can be applied to extract important data regarding a disease. MapReduce automatically processes the input data in parallel as an autonomous entity using MapTask [19]. Approximately 420 research papers have been published from 2010 to 2017 using the abovementioned big data techniques [19].

## 3. Hadoop and MapReduce functionality

Hadoop is an open-source framework used to store large datasets and run applications on clusters of commodity hardware [20]. Its important characteristics include partitioning and computation around multiple hosts. A Hadoop cluster extends its estimation scope using simple commodity computing servers. Hadoop clusters work on large datasets which helps in the analysis using MapReduce, Hive, Pig, and HBase, among others. A dataset in this regard refers to patient data about diseases such as diabetes or others related to heart, blood pressure and more.



**Figure 3** Real-time medical response using sensors and big data



**Figure 4** MapReduce tasks

Once the data is available from various sources using electronic health records (EHR), doctor's reports and other specialists then we proceed with further analysis [11]. Collected data is then passed on to a sensor protocol that determines the value that a sensor node will provide for analysis [2]. The database can then be used to search for the information required for analysis and prediction as shown in Figure 3. Hadoop and MapReduce used for analyzing the data and queries are implemented to predict the health condition [21]. Within a learning system, healthcare data is analyzed and combined with clinical reports to determine whether the condition lies in a critical range requiring further medication or not [22].

After the analysis is done, a further step is taken to attain various parameters used to determine the health condition. Now, we use MapReduce for processing and generating parallel datasets to produce a large number of clusters and apply algorithms to process and generate useful information [5]. MapReduce is the earliest and best-known model in parallel and distributed computing. It contributes to segregate data for better analysis. MapReduce functions in two ways. First, it maps a dataset into small chunks in another dataset in parallel [23-24]. Second, it combines the output dataset into smaller datasets, taking the output of mapped data as an input. This MapReduce model generally helps in categorizing massive data in a more granular manner for processing it over multiple computing nodes for precise analysis [12]. These two basic functionalities are shown in Figure 4. A Map job converts a dataset into another dataset where the elements are split into tuples. A Reduce job further takes input from map and breaks those tuples into smaller tuples. MATLAB 2017 was used to perform map-reduce functions on diabetes data and gain information about patients.

Hadoop clusters work on large datasets that help in analysis using MapReduce, Hive, Pig, and HBase, among others. A dataset here refers to patient data with information about their diseases such as diabetes or other diseases related to the heart or blood pressure and many more. Hive is data warehouse for query processing and managing an abundant amount of distributed datasets on Hadoop [25]. The healthcare industry generates enormous amounts of data which can be uploaded to Hive since it is a scalable and productive warehouse [26]. Patient data can be uploaded from diverse sources such as flat files, real-time applications,

databases, web pages. The produced data can generate reports in the form of graphs and charts that doctors can use for treatment and medications of several patients. Here, MATLAB queries are implemented to drag the important data from the database instead of Hive. MATLAB queries are easier to derive results in this particular case, but the priority must be considered.

#### 4. Implementation

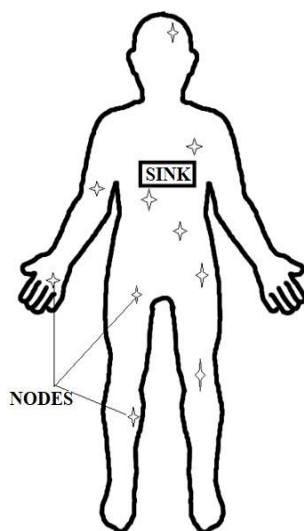
This paper implements WBAN and big data for developing a disease prediction system. Here, WBAN enables a sensing process and big data enables efficient queries to obtain precise results. Using MATLAB with Hadoop, we can easily work on the warehouse and partition the data accordingly. Big data deals with the phenomenon of analysis on massive datasets that are formed using patient data collected using various sensors. MapReduce helps to divide and conquer the data based on the requirements of different fields of doctors and further use it for analyzing the problem for the patient's welfare. An energy efficient routing protocol strategy is used.

#### 5. Energy efficient routing protocols

Numerous efficient routing techniques have been devised by various authors. These techniques are intended to achieve reliable, secure and efficient communications among SNs. With respect to this, [27] introduced packet reception ratio i.e. the number of successful packets received by a sink. In [28], the authors achieved an even consumption of energy using back side routing. This means that SNs are deployed on the back side of the human body with an unequal quantity of initial energy. This enables formulation of a cost-efficient function in each round using the standard deviation. In [29], the authors presented a software-based solution called General Algebraic Modeling System (GASM). It is a mixed-integer programming problem, which balances the energy consumptions among SNs and increases the throughput of the network. In order to increase the lifetime of the network and balance the energy consumptions of SNs, authors designed a Weighted Energy Balanced Efficient routing algorithm in [30]. It is an improvement over the dynamic routing algorithm (DRA), in which Dijkstra's algorithm is used to discover the shortest path. The improved DRA uses

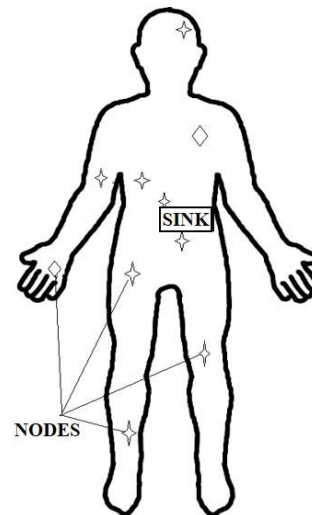
Dijkstra's algorithm to choose a relay node with a minimum hop-count to the sink node. In [31], cross-layer optimization performance was investigated based on a real-life example, by Samiya and others, at the physical and network layers. The best route is selected at the network layer, according to the channel state information from the physical layer. Shortest path routing and cooperative multipath routing are two dynamic routing techniques that are used for reliable data transfer and a lower packet error rate in WBANs. These two routing techniques help in reliable data transfer across WBANs operating near the 2.4 GHz ISM band and a negligible packet error rate is also achieved. To avoid the retransmission of a partially damaged data frame, researchers proposed an approach called Automatic Segment Repeat Request (ASRQ) for IEEE 802.15.4 based WBANs [32]. In this scheme, the data payload is divided into segments and only the damaged data frame is retransmitted. It reduced the traffic load on the network and improved the frame reception rate, and in this way, minimum energy is consumed as well as reliability improved. A new mobility based temperature aware (MTR) routing protocol is proposed by [33], using a store and carry scheme. It also considers the frequent link breaks among neighboring nodes in the network, due to mobility. Two types of nodes are used, static & dynamic. A static node is fixed at the mid-section of the body, while dynamic nodes are attached to the different parts of the body such as legs and arms. To balance the energy consumption among SNs, a new routing path is given to the static node using the mobility characteristics of the dynamic node. In [34], Khan et al. planned an efficient routing technique for WBANs. The data communication method is the same as EERDT.

Sensitive data, like an ECG signal, is transmitted via a single hop. Initially, all SNs calculate a forwarder node to aggregate the data. The distance from a sink node and residual energy of SNs facilitate choosing the forwarder node. As SNs die slowly, network stability has been increased. In [35], the authors give an idea of WBAN with IoT and present Adaptive Thermal-Aware Routing Protocol (ATAR). Its aim is to avoid the issues related to temperature that arises in the human body. This protocol is based on a Multi-Ring routing approach. It finds an alternative route in case of increased temperature. In [36], an Information Centric Networking (ICN) was proposed to secure data communication between doctors and patients.



**Figure 5** Node deployment in a human body (ATTEMPT)

In this paper, two protocols, Energy Efficient and Reliable Data Transfer (EERDT) and Adaptive Threshold-based Thermal-aware Energy-efficient Multi-hop Protocol (ATTEMPT) were compared to analyze their energy efficiency and reliability. The authors chose these protocols for comparison because of their simplicity and flexibility working in both homogeneous and heterogeneous environments. Also, the value sensed by the node will be used to determine the patient's condition using big data. These protocols deploy nodes in a human body as is shown in Figures 5 and 6 to determine the values sensed.



**Figure 6** Node deployment in a human body (EERDT)

In EERDT, ten SNs have been deployed on, in or around the human body. Each SN has equal initial energy and parallel processing capabilities. The sink node is placed on the person's waist. As shown in Figure 6, the SNs for ECG and glucose are appear as diamond and star symbols, respectively. These two SNs transmit data directly to sink.

### 5.1 Algorithm of EERDT protocol

The EERDT algorithm protocol is as follows:

- First, a BS is placed at a fixed position on a person's body. All other SNs are placed at different parts of the body such as a leg, hand, or heart. Each SN has equal energy, i.e., 0.5 J. In this, we propose a concept of hierarchical routing, which helps to reduce packet loss. We are using one-hop routing to improve the quality of WBAN service. One hop routing helps in transferring the most important data first. There is no cluster head because each of the SNs have been initialized with an equal energy.

- In the first round, the value of the cost efficient (CE) function is high and the likelihood of becoming a cluster head (CH) is more. After choosing the CH, all enduring SNs will find out their respective cluster head according to the minimum distance or data is sent directly to BS, whichever is nearer. Then, SNs transmit the data to nearby CH or BS and the consumption of energy will be calculated using (1):

$$\text{Energy Consumption} = (E_{tx} * b) + (E_{amp} * b * d^2) \quad (1)$$

where,  $E_{tx}$  is energy consumption for transmitting the data,  $E_{amp}$  is energy consumption to amplify the data,  $b$  is the number of data bits that need to be transferred, and  $d$  is the distance from a particular node to a CH or BS.

- After collecting all the data, the CHs forward the data to the BS and calculate their energy used during processing from eq.(2):

$$Energy\ Consumption = (E_{tx} + E_{rx} + E_{da}) * b + (E_{amp} * b * d^2) \tag{2}$$

where,  $E_{TX}$  is consumption of energy by a transmitter to send data.  $E_{RX}$  is consumption of energy in receiving data,  $E_{amp}$  is consumption of energy in amplification,  $b$  is the number of data bits that need to be transferred, and distance is the distance from the CH to the BS.

$$CE(i) = Residual\ Energy\ (i) / Initial\ Energy(i) \tag{3}$$

where,  $i$  is the particular node for which cost efficient function is to be calculated. Higher the value of CE, higher the chances to become CH.

- In next round, the SNs will become CHs by calculating their cost efficient function (CE) given in Eq. (3). After selection of the CHs for next round, SNs transmit the data to their respective CHs or BS and evaluate their energy consumption according to Eq. (1). After gathering all essential information, CHs transmit the data to BS and calculate their energy consumption using Eq. (2).

- The above steps will be repeated until the all the SNs deplete their energy or the number of rounds are finished.

In this study, total seven network parameters, (i) initial energy, (ii) amplifier energy, (iii) transmitting energy, (iv) receiving energy, (v) data aggregation energy, (vi) packet size and (vii) number of SNs, are considered to estimate the energy consumption. The values of these seven parameters used here are given in Table 1. A WBAN uses miscellaneous medical sensors such as ECG, EEG, EMG, glucose, temperature, blood pressure and foot pressure as depicted in Figures 5 and 6. These sensors are used to collect data related to the various body parts where they are placed. The values are determined by two routing protocols, EERDT and ATTEMPT. Then a comparison is done and the most efficient value is chosen for analysis using the collected data and large datasets.

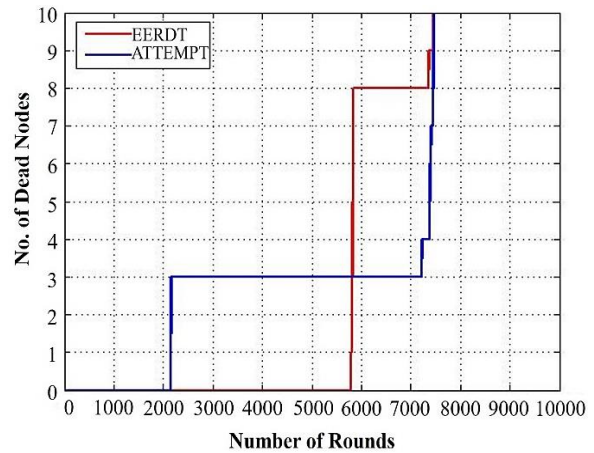
**Table 1** Network parameter values for energy consumption calculation

Simulation Parameters	Value
Initial Energy, $E_o$	0.5 J
Amplifier energy, $E_{amp}$	1.97nJ/b
Transmitting Energy $E_{tx}(elec)$	16.7nJ/bit
Receiving Energy $E_{rx}(elec)$	36.1nJ/bit
Data Aggregation Energy( $E_{da}$ )	5nJ/ bit
Packet size (b)	4000 bits
No. of SNs	10

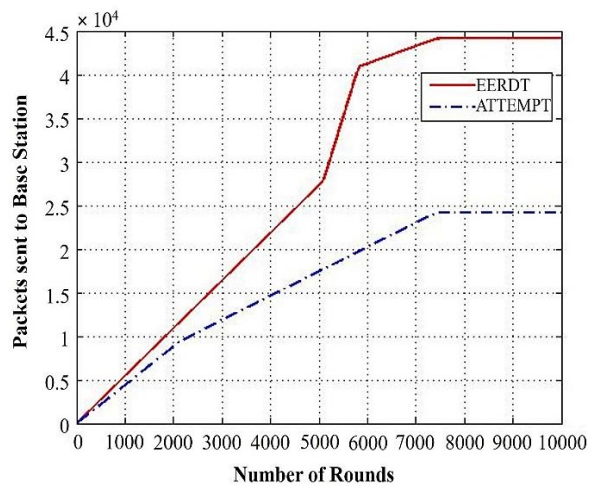
**6. Results**

Figure 7 depicts the number of SNs that lose their energy during the process per to the number of rounds. The stability of the EERDT network protocol has been increased to approximately 5000 rounds compared to ATTEMPT. A total of three SNs of ATTEMPT were dead after around 2000 rounds. Consequently, the EERDT protocol is much more stable in terms of the lifetime of SNs and it functions for a greater number of rounds than the ATTEMPT protocol. Figure 8 shows the throughput of the network. EERDT transfers more data from the cluster head to BS or from the

SNs to the BS. As the numbers of rounds is increased, EERDT will transfer more data packets to the BS than ATTEMPT. In this way, EERDT is more reliable in terms of data transfer as it sends more information



**Figure 7** Network lifetime (ATTEMPT vs EERDT)



**Figure 8** Data transfer (ATTEMPT vs EERDT)

to the BS. Figure 9 demonstrates energy consumption by the SNs in the network.

From Figure 9, it can be observe that ATTEMPT lost more than half of its energy in 2000 rounds, whereas EERDT lost half of its energy around 3000 rounds. After around 4000 rounds, both ATTEMPT and EERDT had lost equal amounts of energy and all SNs are dead before 7500 rounds in both protocols. Table 2 compares the network lifetime of ATTEMPT and EERDT protocols. The first SN in ATTEMPT died at around 2149 rounds, where the as first SN in the EERDT protocol died at around 5791 rounds. All SNs were dead a little earlier in EERDT (7450 rounds) compared to ATTEMPT (7471 rounds). But, overall the EERDT protocol shows much more stability than ATTEMPT.

**Table 2** Network lifetime comparison achieved from ATTEMPT and EERDT protocols

Protocols/SNs	ATTEMPT (Rounds)	EERDT (Rounds)
First node died	2149	5791
All SNs died	7471	7450

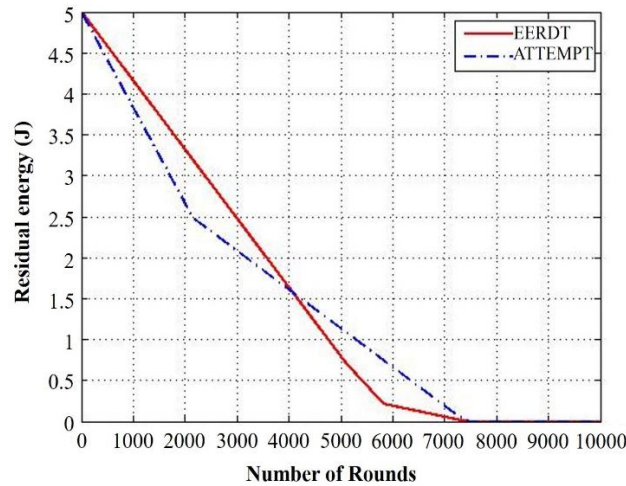


Figure 9 Energy consumption (ATTEMPT vs EERDT)

```

1 ds = datastore('diabetics.csv');
2 ds.SelectedVariableNames = 'diabetesMed';
3 ds.SelectedFormats = '%C';
4 preview(ds)
5 outds = mapreduce(ds, @countMapper, @countReducer);
6 readall(outds)
7 function countMapper(data, ~, intermKV)
8 % Counts insulin in each chunk.
9 a = data.diabetesMed;
10 %b = data.glucose;
11 c = num2cell(countcats(a));
12 keys = categories(a);
13 addmulti(intermKV, keys, c)
14
15 %d = num2cell(countcats(b));
16 %key_s = categories(b);
17 addmulti(intermKV, key_s, d)
    
```

(a) Code for MapReduce in MATLAB

```

Starting parallel pool (parpool) using the 'local' profile ...
connected to 2 workers.
Parallel mapreduce execution on the parallel pool:
*****
*      MAPREDUCE PROGRESS      *
*****
Map  0% Reduce  0%
Map 50% Reduce  0%
Map 100% Reduce  0%
Map 100% Reduce 100%
    
```

(b) Output for fig. 10(a) in MATLAB

Figure 10 Applying MapReduce on a dataset

Data drawn from the queries implemented on MATLAB determines the patient’s details about a person’s health status. Various disease categories include diabetes, high blood pressure and insulin levels. All these give insight into the medical conditions of patients. Queries can be customized to access the required data for each individual patient [37-39]. Routing protocols can merge big data functions such as map-reduce to partition and categorize them based on the required information as shown in Figure 10. Big data and WBAN collectively contribute to

analyse and generate data in the field of healthcare. Combining both helps in implementing the algorithms of protocols and map-reduce in a concise manner for efficient outcomes [40].

Implementation of these queries as shown in Figure 11 determines important data about particular patients. Figure 10 depicts that if a person is having a high glucose value, the patient should seek immediate medical attention. Similarly, other parameters can be integrated to get more accurate results.

```

Command Window
New to MATLAB? Watch this Video, see Demos, or read Getting Started.

[150] 'XX' 'Male' [54] '120/80' 'normal' 'Up' [355] 'Yes' [1x32 char] [97.2000]
'seek immediate medical attention'

[254] 'CCC' 'Female' [46] '210/120' 'hypertension-4' 'Down' [355] 'Yes' [1x32 char] [97.2000]
'seek immediate medical attention'

[358] 'SSSS' 'Female' [54] '130/85' 'high' 'Steady' [355] 'Yes' [1x32 char] [97.2000]
'seek immediate medical attention'

[498] 'HH' 'Male' [87] '90/60' [1x22 char] 'Steady' [355] 'Yes' [1x32 char] [97.2000]
'seek immediate medical attention'

[602] 'XXX' 'Female' [40] '180/110' 'hypertension-3' 'Up' [355] 'Yes' [1x32 char] [97.2000]
'seek immediate medical attention'

[742] 'M' 'Female' [47] '130/85' 'high' 'Steady' [355] 'Yes' [1x32 char] [97.2000]
'seek immediate medical attention'

[846] 'HHHH' 'Male' [32] '175/105' 'hypertension-3' 'Down' [355] 'Yes' [1x32 char] [97.2000]
'seek immediate medical attention'

[986] 'MMMM' 'Female' [42] '60/40' 'serious' 'No' [355] 'Yes' [1x32 char] [97.2000]
'seek immediate medical attention'

[1090] 'RRRR' 'Male' [52] '140/90' 'hypertension-1' 'Steady' [355] 'Yes' [1x32 char] [97.2000]
'seek immediate medical attention'

```

**Figure 11** Results showing query results based on patient data

Further, the results are saved to keep track of the health of the patient. Information provided can be used for further analysis by doctors to give treatment accordingly.

## 7. Conclusions

This paper describes integration of big data with WBAN using MATLAB as a front end for implementation. It is observed that big data can easily be fused with many networks to achieve effective results. The use of an efficient WBAN algorithm as ATTEMPT and EERDT along with MapReduce gives effective outcomes. Including a big data framework with MATLAB enhances the use of large datasets for efficient results. Such an algorithm gives better accuracy to accumulate more granular details from the sensors. Using precise values from various datasets gives accurate outputs. Healthcare that merges big data and sensors can give even better results if Hadoop frameworks like Hive and Pig contribute. Future work includes examining other WBAN algorithms and fusing them with Hive and scooping queries for more efficient results.

## 8. References

- [1] Andreu-Perez J, Poon CCY, Merrifield RD. Big data for health. *IEEE J Biomed Health Informat.* 2015; 19:1193-208.
- [2] Asif M, Khan S, Ahmad R, Sohail M, Singh D. Quality of service of routing protocols in wireless sensor networks: a review. *IEEE Access.* 2017;5:1846-71.
- [3] Carroll JG. Crossing the quality chasm: A new health system for the 21st century. *Qual Manag Health Care.* 2002;10:60-1.
- [4] Darwish A, Hassanien AE. Wearable and implantable wireless sensor network solutions for healthcare monitoring. *Sensors.* 2011;11: 5561-95.
- [5] Dean J, Ghemawat S. MapReduce: a flexible data processing tool. *Comm ACM.* 2010;53:72-7.
- [6] Gandomi A, Haider M. Beyond the hype: Big data concepts, methods, and analytics. *Int J Inform Manag.* 2015;35:137-44.
- [7] Olaronke I, Oluwaseun O. Big data in healthcare: Prospects, challenges and resolutions. In *Future Technologies Conference (FTC); 2016 Dec 6-7; San Francisco, USA. USA: IEEE; 2016. p. 1152-7.*
- [8] Negra R, Jemili I, Belghith A. Wireless body area networks: applications and technologies. *Procedia Comput Sci.* 2016;83:1274-81.
- [9] Singh D. Secure 6lowpan networks for e-healthcare monitoring applications. *J Theor Appl Inform Tech.* 2015;76(2):143-51.
- [10] Singh, D. Mobility and energy efficient mechanism for 6lowpan devices to support global healthcare system. *J Theor Appl Inform Tech.* 2014;66(1):315-29.
- [11] Hripsak G, Albers DJ. Next-generation phenotyping of electronic health records. *J Am Med Informat Assoc.* 2012;20:117-21.
- [12] Lämmel R. Google's MapReduce programming model-revisited. *Sci Comput Program.* 2008;70(1): 1-30.
- [13] Bakul G, SinghD, KimD. Optimized WSN for ECG monitoring in ubiquitous healthcare system. *The 4th International Conference on Interaction Sciences; 2011 Aug 16-18; Busan, South Korea. USA: IEEE; 2011. p. 23-6.*
- [14] Peek N, Abu-Hanna A. Clinical prognostic methods: trends and developments. *J Biomed Informat.* 2014;48:1-4.
- [15] Nadeem Q, Javaid N, Mohammad SN, Khan MY, Sarfraz S, Gull M. SIMPLE: Stable increased-throughput multi-hop protocol for link efficiency in wireless body area networks. *IEEE 8th International Conference on Broadband and Wireless Computing,*

- Communication and Applications (BWCCA'13); 2013 Oct 28-30; Compiègne, France. USA: IEEE; 2013. p. 221-6.
- [16] Sethi D, Bhattacharya PP. A study on energy efficient and reliable data transfer (EERDT) protocol for WBAN. 2016 Second International Conference on Computational Intelligence and Communication Technology (CICT); 2016 Feb 12-13; Ghaziabad, India. USA: IEEE; 2016. p. 254-58.
- [17] Anand J, Sethi D. Comparative analysis of energy efficient routing in WBAN. 2017 3rd International Conference on Computational Intelligence & Communication Technology (CICT); 2017 Feb 9-10; Ghaziabad, India. USA: IEEE; 2017. p. 1-6.
- [18] Viceconti M, Hunter P, Hose R. Big data, big knowledge: big data for personalized healthcare. *IEEE J Biomed Health Informat.* 2015;19:1209-15.
- [19] Shabestari F, Rahmani AM, Navimipour NJ, Jabbehdari S. A taxonomy of software-based and hardware-based approaches for energy efficiency management in the Hadoop. *J Netw Comput Appl.* 2019;126:162-77.
- [20] Khan P, Hussain A, Kwak K. Medical applications of wireless body area networks. *Int J Digit Content Tech Appl.* 2009;3(3):185-93.
- [21] Halde SV, Khot ST. Efficient collection of big data in WSN. 2016 International Conference on Inventive Computation Technologies (ICICT); 2016 Aug 26-27; Coimbatore, India. USA:IEEE; 2016. p. 1-5.
- [22] Rani S, Ahmed SH, Talwar R, Malhotra J. Can sensors collect big data? an energy-efficient big data gathering algorithm for a WSN. *IEEE Trans Ind Informat.* 2017; 13(4):1961-68.
- [23] Yahya O, Hegazy O, Ezat E. An efficient implementation of a-priori algorithm based on hadoop-MapReduce model. *Int J Rev Comput.* 2012;12:59-66.
- [24] Shvachko K, Kuang H, Radi S, Chansler R. The hadoop distributed file system. 2010 IEEE 26th Symposium on Mass Storage Systems and Technologies (MSST); 2010 May 3-7; Incline Village, USA. USA: IEEE; 2010. p. 1-10.
- [25] Chennamsetty H, Chalasani S, Riley D. Predictive analytics on electronic health records (EHRs) using hadoop and hive. 2015 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT); 2015 Mar 5-7; Coimbatore, India. USA: IEEE; 2015. p. 1-5
- [26] Sahoo PK, Mohapatra SK, Wu SL. Analyzing healthcare big data with prediction for future health condition. *IEEE Access.* 2016;4:9786-99.
- [27] Jafari R, Effatparvar M. Cooperative routing protocols in wireless body area network: a survey. *Int J Comp Info Tech.* 2017;5(1):43-51.
- [28] Ha I. Even energy consumption and back side routing: an improved routing protocol for effective data transmission in wireless body area networks. *Int J Distr Sensor Network.* 2016;12(7):1-11.
- [29] Effatparvar M, Dehghan M, Rahmani AM. Lifetime maximization in wireless body area sensor networks. *Biomed Res.* 2017;28(22):9805-12.
- [30] Shimly SM, Smith DB, Movassaghi S. Experimentally-based cross-layer optimization across multiple wireless body area networks. *arXiv:1701.08605.* 2017:1-15.
- [31] Kim B, Lee B, Cho J. ASRQ: automatic segment repeat request for IEEE 802.15.4-based WBAN. *IEEE Sens J.* 2017;17(9):2925-35.
- [32] Kim BS, Kang S, Lim J, Kim KH, Kim KI. A mobility-based temperature-aware routing protocol for wireless body sensor networks. 2017 International Conference on Information Networking (ICOIN); 2017 Jan 11-13; Da Nang, Vietnam. USA:IEEE; 2017. p. 63-6.
- [33] Ara A, Al-Rodhaan M, Tian Y, Al-Dhelaan A. A secure privacy-preserving data aggregation scheme based on bilinear ElGamal cryptosystem for remote health monitoring systems. *IEEE Access.* 2017;5:12601-17.
- [34] Khan RA, Mohammadani KH, Soomro AA, Hussain J, Khan S, Arain TH, et al. An energy efficient routing protocol for wireless body area sensor networks. *Wireless Pers Comm.* 2018;99(4):1443-54.
- [35] Jamil F, Iqbal M, Amin R, Kim D. Adaptive thermal-aware routing protocol for wireless body area network. *Electron.* 2019;8(1):1-28.
- [36] Naik BB, Singh D, Samaddar AB, Lee HJ. Security attacks on information centric networking for healthcare system. 2017 19th International Conference on Advanced Communication Technology (ICACT); 2017 Feb 19-22; Bongpyeong, South Korea. USA:IEEE; 2017. p. 436-41.
- [37] Sharma R, Ryait HS, Gupta AK. Performance analysis of attempt, simple and deec routing protocols in WBAN. *Int J Latest Trends Eng Tech.* 2015;6(2):133-39.
- [38] Wang Y, Li PF, Tian Y, Ren JJ. A shared decision-making system for diabetes medication choice utilizing electronic health record data. *IEEE J Biomed Health Informat.* 2017;21:1280-87.
- [39] Ibarra E. QoS-aware energy management in body sensor nodes powered by human energy harvesting. *IEEE Sens J.* 2016;16:542-49.
- [40] Rios LG. Big data infrastructure for analyzing data generated by wireless sensor networks. 2014 IEEE International Congress on Big Data; 2014 Jun 27-Jul 2; Anchorage, USA. USA: IEEE; 2014. p. 816-23.