International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 3 | Issue 1 | ISSN : 2456-3307

Data Organization and Knowledge Inference from Sensor Database for Smart Wear

Dr. Nilamadhab Mishra[1*], Kindie Alebachew[2], Bikash Chandra Patnaik[3]

[1*]Post Graduate Teaching & Research Department, School of Computing, Debre Berhan University, Debre Berhan, Ethiopia

[2]School of Computing, Debre Berhan University, Debre Berhan, Ethiopia

[3] Gandhi Institute of Engineering and Technology, Orissa, India

ABSTRACT

Sensor data base has found increasing applications in health care domain. A wide variety of Intensive Care Unit (ICU) applications use sensors such as ECG, EEG, blood pressure monitors, respiratory monitors, and a wide variety of other sensors from where huge amount of physiological signals/data are generated. The main research challenge is how to manage and organize those physiological data with an intention of tracking the condition of patient and providing time critical service information to the patient's smartphone for emergency precautions. The volume and velocity of such data tends to big-data complications and the knowledge inferences from such data need to be performed in a time-critical fashion. In this paper, we discuss the physiological data organization approaches in order to provide the storage for large scale physiological data and the emerging Knowledge inference mechanism for transforming the physiological data into insights in the context of health-care application.

Keywords : Sensor Database, Smart Wear, Knowledge Inference, Physiological Data, Thing To People

I. INTRODUCTION

In the emerging Internet of Things (IOT) each and every real world object can be networked to accomplish the Thing to People (TTP) and Thing to Thing (TTT) communications. In this research, Smart wear is a thing, having five major sensors:-Body temperature sensor, accelerometers on arms and body, lead ECG (Electrocardiography) sensors, lead EMG (Electromyography) sensors, and breathing sensors.

Those sensors are fabricated into the smart wear by using embedded SOCs integrations. A smart phone can be used to be act as gateway and some minor pre-processing. According to the body positions, different sensors are fabricated in different parts of the smart wear and have wired links in between them [1].

The physiological signals/data which sensed from the above sensors are passed through the gateway (Smart Phone) to the internet/extranet cloud via wireless link. The cloud has multiple different types of networked database servers to accommodate different classes of data and its processing activities [2],[3].

A sink must be fabricated in smart wear, which has wired links to all sensors in which all sensors can able to transmit the physiological signals/data to the cloud through the gateway (**figure-1**). The cluster of doctors accesses the data from the cloud to provide necessary services.

Each smart wear must have an embedded chip to store its user localization details and other information to establish emergency contacts. The data fusion has wide spread applications in gathering data from multiple sensors and produces the information of tactical value to the users. Data can come from one or many homogeneous /heterogeneous sources and each data source may be a sensor [4], [5], [6].

Here distributed Data base approach may be preferred to implement, where each sensor acts as a part of database to minimize the amount of data flow to the database server. Here the sink is an active database node to handle instant or long running OLTP queries only and responsible to transmit the sensing data to the DB server [7], [8]. In this paper, emerging Knowledge inference mechanism is discussed for transforming the physiological data into insights in the context to the health-care application.

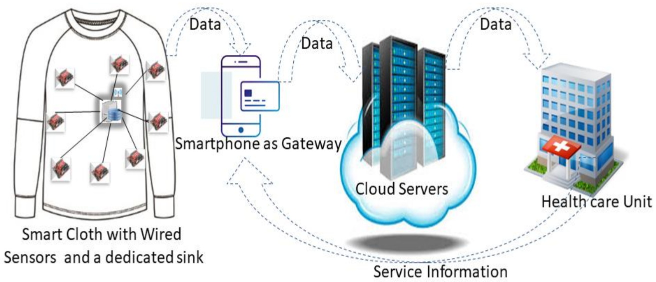

Figure 1 : Data flow structural design from Smart Wear to Healthcare Unit in consort with the flow of Service Information

The emerging research challenges in this context include followings:-

i. Early-stage diseases alarm is a real time issue according to long-term continuous physiological signals.
ii. Provide a new medical service model between patient/health people and doctors.
iii. Design an architecture and database to store physiological signals/data, which are measured from various peoples, who wearing the Smart wear.
iv. Effective data access management in sensor cloud database, so as to preserve the data consistency and integrity to certify that the people will obtain right health information by the right people at the right time to save and maintain a healthy life.

The rest of this paper is organized as follows. Section II discusses the data organization approach for sensor's physiological data. Section III discusses Knowledge inference analysis approach. Finally section IV concludes this paper.

II. DATA ORGANIZATION APPROACH

Based on different types of physiological signals, various databases can be used to place those signals/data. ECG database, EMG database, Clinical database -skin temp, respiration rate, heart rate, Image database-This database comprises several databases like clinical images, normal X-ray, radiology, ECG and multimedia images, and Accelerator database -to detect the motion of arms, head, neck, and back.

As per analysis of physio Bank databases, we observe that data can also be divided into several classes like class-1, class-2 and so on. We also think some architectural aspects of the above databases to record the continuous incoming data. The architectural design of database always starts with the users view and passes through the abstract view, logical/conceptual view and physical view. In abstract view, the attributes selection and finalization are important. In logical view, the appropriate logical schema is to be designed by considering its data structure. In physical view, it is important to consider the physical data organization through choosing the appropriate data model to support continuous data recording.

Now we discuss database server organization for the physical data organizations. Front server receives the incoming data and passes to signal DB server. The front server may involve in some pre-processing. Signal DB server (DSP) stores physiological signals /data in predefined format and Image processing and OLTP operation. Event database server records the incoming life events, Transaction processing, Handles the emergency data needs and processing. Finally, the Data warehouse server stores large amount of current and historical physiological data in a standard unified schema format, which will lead to further data mining and data analysis [9], [10], [11]. Here ETL operations are performed to transform the data into various formats–Meta data, information details, raw text/image data, and summary/report.

The physiological data model selections include hierarchical, network, and relational model. IMS-DB implements hierarchical data model that uses tree like ordered structure. We can add data in an increasing manner, but chances of data anomaly, inconsistency and redundancy may degrade the performance. For large and complex physiological data pointers i.e. PCR and sibling pointer management will be difficult. IDMS implements network data model that experiences difficulties to place the continuous incoming data. Both representations i.e. matrix based and linked list based may not be suitable to cop up with. DB2 implements relational data model, which Suitable to store and handle large and bulk amount of continuous incoming data. In DB2, large storage space can be divided into n-number of generation Data Group (GDG) with appropriate version (**figure-2**).

Each GDG can be divided into number of datasets or generation Data Sets (GDS) with appropriate versions. Based on new incoming data from sensors, new GDS can be created with appropriate versions. The multi sensor data fusions can be implemented through GDG and GDS, as because it deals with the associations, correlations and combinations of data and information from multi sensor sources.

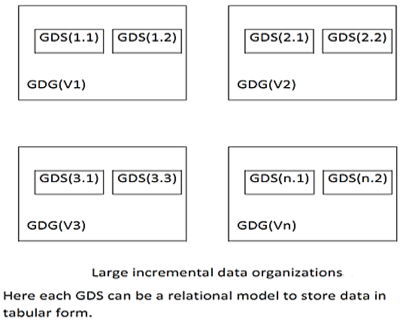

Figure 2: Large scale physiological data organizations in the form of GDG and GDS

Now we focus on the conceivable designs of the logical schema of the database to house continuous store of incoming agile data. For example in ECG database, we consider followings for logical schema.
Rate (normal rate is 60-100/minute, Integer data)
Pattern of QRS complex (regular/irregular, string data) QRS morphology (Narrow complex- sinus, atrial or junction origin, wide complex — ventricular origin, string data)
P waves (absent — sinus arrest, atrial fibrillation, Present — morphology and PR interval may suggest sinus, atrial, junction or even retrograde from the ventricles, string data)
A relationship between P waves and QRS complexes can be studied by AV dissociation (complete/incomplete, Boolean data). Finally, Response to vagal movements can be investigated by sudden terminations/no response.
Hence the above data can be obtained from the ECG measure strip to convey further analysis. The depiction has made on the possible physical schema of the databases to store the incoming data.

- ECG database {sample, time, date, frequency, gain, group, and signal.
- EMG database {patient information, examination information, clinical information, test conditions, parameters, symbolic parameter values, pathophysiological test conclusions.
- Pathophysiological structure conclusions, EMG diagnoses, clinical diagnoses.
- Clinical database {skin temp, respiration rate, heart rate.
- Accelerator database {motion of arms, motion of head, motion of neck, motion of back.

Better to use data warehouse model for agile data rather hierarchical, network, and relational data models. Individual EERD may help to understand the complete process to finalize the number of databases (entities) and their respective Attributes. The logical design of database is highly considered to achieve a true data independency, which is highly desired. The traditional databases are difficult to apply in sensor system because of limited sensor resources. So managing real time distributed data and transaction on high mobility sensor network is a challenging issue. Parallel, complex, and distributed data management on sensor's network with a real time query response is another challenging issue.

III. KNOWLEDGE INFERENCE ANALYSIS

The knowledge inference is a major part of data analytics that studies the historical data to research potential trends, analyses the effect of decisions and events, evaluates the performance of complex problem scenarios, and aims to improve values through gaining knowledge and insights [12]. The knowledge analytic is the science of logical analysis that uses mathematics, statistics, computational intelligence, and other analytic tools to discover the potential knowledge and insights from large scale data sensor's environment. The knowledge analytic mechanism progressively integrates with several new technologies, such as- IoT, big-data, clouds, deep learning, extreme learning machine (ELM), and many more emerging technologies. If we analyse the upcoming prospect of knowledge analytic and inference, we observe that it tends toward delivering both big-data processing and knowledge analytics, which are the most challenging aspects with the growing data dimension and diversity of numerous sensor's applications. The data architecture of sensor's physiological data relies on several NoSQL databases on Hadoop like platforms for batch processing of large scale data that consumes much more time; however the real-time or semi real-time data processing, management and knowledge analytics are much more thought-provoking tasks. Because, the current healthcare platforms need the timely knowledge and insights to transform their data into cognitive decisive goldmines in order to make huge revenues through minimizing the potential upcoming business risks. The NoSQL databases are not designed to execute the knowledge analytic tasks, but it indeed a common minimal requirement for healthcare domain applications.

Several hazards are associated with knowledge analytics. Those hazards are- managing heterogeneous knowledge, transforming the data into knowledge, transforming the knowledge into actions, transforming the actions into cognitive-decisions, and tuning physiological knowledge base to regulate the intellectual healthcare application.

We analyze a small scaled physiological data, where the data selection mechanism uses a pre-structured data munging process, in which the range based natural language query can be effectively used. In this analysis, the data domain with respect to the specific healthcare application scenario is diagnosed for the configuration of representative data set. The configurations of representative data set implements a data munging process that takes the data of healthcare application, select the featured data through an outlier reduction mechanism or any data mining mechanism, and normalize/transform those data into the desired scale/format that can be well fitted to the intellectual healthcare applications. In our work, we analyze several data normalization mechanisms, such as max-min, z-score, and decimal point that transform the featured data into the standardized data for application use. As per the knowledge analytic and inference concerned, we can scale-up and scale-down the original application data into massage data that can be effectively used in several computational analytic frames to generate cognitive insights.

Case analysis:-

In real implementation to human activity supervisions, a human can comfortably use a wireless smart cap that consists of micro EEG sensors, such that those sensors can cooperatively record brain waves through hair directly and send the brain waves to the mobile device wirelessly so as to establish the interactions and relationships in-between human, mobile device, and external applications. The smart cap uses micro EEG sensors to give sufficient freedom and comfort to the human without the risk of critical data loss. We can incorporate the smart device functions into a smart phone or smart watch like portable device for high convenience of the human. Also based on the cognitive discrimination level, the physical ability level, and mobility level of the human, the smart functions may be introduced either as a smart phone like portable device or as a smart cognitive robotic device to regulate the human activities so as to provide the necessary assistance without any other's interventions [1],[13],[14].

The raw electroencephalography (EEG) data are collected from a wireless EEG sensor cap based on the human activities and transformed into mapped EEG data in order to implement into an inference frame. The figure-3 gives an analysis about human inference activities analysis based on the mapped EEG data. Based on knowledge inference identifier, mainly four broad human activities are analyzed, i.e. deep sleep, drowsiness, relax and alert, and full active state.

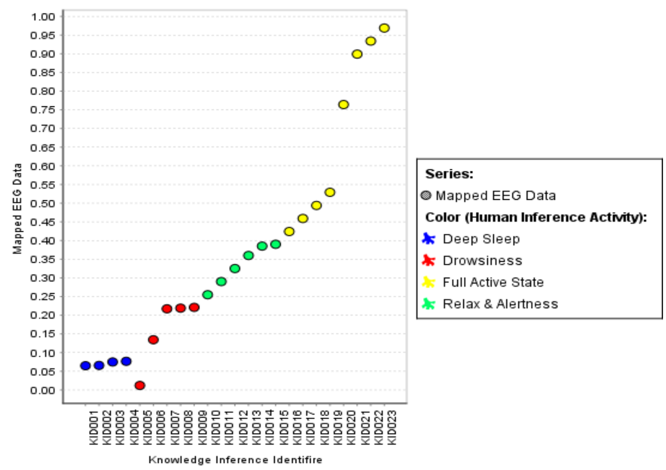

Figure 3 : Human inference activities analysis based on the mapped EEG data.

The analysis tells that the EEG data values vary according to different human activities and with the increasing of EEG data values, human activeness increases. In this case, we consider four broad activities of human that absolutely incorporate the entire activities of human daily living activities.

IV. CONCLUSIONS

In this paper, we discussed an intelligent sensing device that can manage human activities in an effective and efficient way. The physiological data organization approaches are discussed in order to provide the storage of large scale physiological data and the emerging Knowledge inference mechanism is discussed for transforming the physiological data into insights in the context to the health-care application. Our discussion and analysis may be a good companion to the human toward leading a self-regulating life.

V. ACKNOWLEDGMENT

The author would like to express thanks to the Post Graduate Teaching & Research Dept., at School of Computing, Debre Berhan University, Ethiopia for supporting this research.

VI. REFERENCES

[1]. Mishra, N., Lin, C. C., & Chang, H. T. (2014). Cognitive inference device for activity supervision in the elderly. The Scientific World Journal, 2014.

[2]. Suryadevara, N. K., & Mukhopadhyay, S. C. (2012). Wireless sensor network based home monitoring system for wellness determination of elderly. Sensors Journal, IEEE, 12(6), 1965-1972.

[3]. Suryadevara, N. K., Gaddam, A., Mukhopadhyay, S. C., & Rayudu, R. K. (2011, November). Wellness determination of inhabitant based on daily activity behaviour in real-time monitoring using Sensor Networks. In Sensing Technology (ICST), 2011 Fifth International Conference on (pp. 474-481). IEEE.

[4]. Mishra, N., Lin, C. C., & Chang, H. T. (2015). A cognitive adopted framework for IoT big-data management and knowledge discovery prospective. International Journal of Distributed Sensor Networks, 11(10), 718390.

[5]. Mishra, N., Lin, C. C., & Chang, H. T. (2014, December). A cognitive oriented framework for IoT big-data management prospective. In Communication Problem-Solving (ICCP), 2014 IEEE International Conference on (pp. 124-127). IEEE.

[6]. Chang, H. T., Mishra, N., & Lin, C. C. (2015). IoT Big-Data Centred Knowledge Granule Analytic and Cluster System for BI Applications: A Case Base Analysis. PloS one, 10(11), e0141980.

[7]. Mishra, N., Chang, H. T., & Lin, C. C. (2014). Data-centric knowledge discovery strategy for a safety-critical sensor application. International Journal of Antennas and Propagation, 2014.

[8]. Mishra, N., Chang, H. T., & Lin, C. C. (2015). An Iot knowledge reengineering framework for semantic knowledge analytics for BI-services. Mathematical Problems in Engineering, 2015.

[9]. Mishra, N. (2011). A Framework for associated pattern mining over Microarray database. International Journal of Global Research in Computer Science (UGC Approved Journal), 2(2).

[10]. Mishra, N., Chang, H. T., & Lin, C. C. (2018). Sensor data distribution and knowledge inference framework for a cognitive-based distributed storage sink environment. International Journal of Sensor Networks, 26(1), 26-42.

[11]. Mishra N, (2017). "In-network Distributed Analytics on Data-centric IoT Network for BI-service Applications", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN: 2456-3307, Volume 2, Issue 5, pp.547-552, September-October.2017.

[12]. Patnaik, B. C., & Mishra, N. (2016). A Review on Enhancing the Journaling File System. Imperial Journal of Interdisciplinary Research, 2(11).

[13]. Chang, H. T., Li, Y. W., & Mishra, N. (2016). mCAF: a multi-dimensional clustering algorithm for friends of social network services. SpringerPlus, 5(1), 757.

[14]. Chang, H. T., Liu, S. W., & Mishra, N. (2015). A tracking and summarization system for online Chinese news topics. Aslib Journal of Information Management, 67(6), 687-699.

Authors Profile

Dr. Nilamadhab Mishra is currently an Assistant Professor in Post Graduate Teaching & Research Dept., at School of Computing, Debre Berhan University, Ethiopia. He accomplishes his PhD in Computer Science and Information Engineering from Chang Gung University, Taiwan. He moreover publishes numerous peer reviewed researches in Thomson Reuter's ranked SCI journals & IEEE conference proceedings, and serves as reviewer and editorial member in peer reviewed Journals and Conferences. Dr. Mishra's research areas focus on Network Centric Data Management and Knowledge Discovery, IoT Data Science and Knowledge Analytics, Business Intelligence, and Cognitive Applications exploration. He has 15 years of Academic Teaching and Research Experience.

Mr. Kindie Alebachew is currently working as a lecturer in College of computing at Debre Berhan University, Ethiopia. He completed his masters of Science in Information Science at Addis Ababa University, Ethiopia. His Research interests comprise Machine Learning, AI, and Data Science. He has 5 years of teaching and research Experience.

Mr. Bikash Chandra Patnaik is currently working as a professor in Gandhi Institute of Engineering and Technology, Orissa, India. He completed his masters of technology in Computer Science at Utkal University, India. His Research interests comprise Machine Learning and Data Science. He has 20 years of academic teaching and research Experience.