Asian Journal of Computer Science and Technology
ISSN: 2249-0701 (P) Vol.8 No.S3, 2019, pp.90-93

DOI: https://doi.org/10.51983/ajcst-2019.8.S3.2078

Big Data Analytics and Security in Healthcare Industry

K. Rohitha[1], V. Bhagyasree[2], K. Kusuma[3] and S. Kokila[4]
[1,2,3&4]Assistant Professor,Sreenivasa Institute of Technology and Management Studies, Chittoor, Andhra Pradesh, India
E-Mail: rohitha259@gmail.com

(Received 20 April 2019; Revised 30 April 2019; Accepted 12 May 2019; Available online 20 May 2019)

***Abstract* - Big data analytics plays a major role in today's industry which insisted to use big data analytics for the analysis of previous data. Patient record keeping is very much important to track the history of the patient. According to the patient previous records, decision is made. Large volumes of data are created on a daily basis and this data is used in decision making process. But, health care industry has not sensed the potential benefits from big data analytics. To address this need, four big data analytics capabilities were identified. In addition to four, five capabilities were proposed which providespractical insights for administrator. On the other way, data security plays a key role in health care industry. In order to overcome this, a new architecture is proposed for the implementation to IOT and process scalable sensordata for health care systems. This paper focuses on data security so that we can make use of potential capabilities and benefits of big data analytics in a better way.**
***Keywords:* Big Data Analytics, Big Data Analytics Capabilities, Internet of Things, Healthcare**

I. INTRODUCTION

The big data revolution is transforming the way we live in today's world. The past few years have seen a tremendous generation of data, which impacts our day-to-day life and health care has also been impacted by it. The healthcare industry has absolutely lagged compared to other industries such as banking, retail, etc. in the usage of big data.The healthcare industry maintains large amounts of data about every patient but accessing, managing, and interpreting that data is very critical to creating actionable insights for better care and efficiency. Clinical trends also play an important role in the rise of big data in healthcare. Earlier physicians used their judgments to make their treatment decisions.

The past few years have seen a shift in the way these decisions are being made. Physicians review the patient data and make an informed decision about a patient's treatment. Financial concerns, improved insights into treatment, research, and efficient practices contribute to the need for big data in the healthcare industry.IoT adds additional value to the healthcare industry. Devices that generate data about a person's health and send it to the cloud will lead to a many of insights about an individual's heart rate, weight, blood pressure, lifestyle, and much more. Big data allows real-time monitoring of patients, which leads to intense care. Wireless Sensors and wearable devices will collect patient health data, even from home. This data is monitored by healthcare industries to provide remote health alerts and lifesaving insights to their patients.Smartphone's have added a new dimension. The application enables the Smartphone to be used as a calorie counter to keep a track of calories; pedometers to keep track of how much you walk in a day. All these have helped people live a healthier lifestyle. Further, this data could be shared with a doctor, which will help towards personalized care and treatment.

II. BACKGROUND

A. Big Data

''Big Data'' originally meant the volume, velocity, and variety of data that becomes difficult to process when using traditional data processing platforms and techniques. Especially, big data plays a vital role in healthcare applications [6]. Nowadays, modern healthcare systems are adopting quickly clinical data, which is used for increasing clinical record sizes available online [7]. In addition, new technologies and tools are identified for processing data of large size and gleaning new business insights from that analysis.

As a result, a number of options are identified to use Big Data for reducing the cost of healthcare and to diagnose diseases [9–11]. Some authors have conducted a study on Big Data in healthcare and describe six use cases of Big Data to reduce healthcare cost [12]. In addition, some authors have divided the use cases of Big Data in healthcare into four main categories, namely, management and deliverance, clinical judgment support, sustain services and consumer activities [13]. Similarly, some authors have conducted a study that describes the manner of improvement in the healthcare system based on the Big Data analytics to improve the healthcare systems, select appropriate treatment path, and so on [14–18].

B. Cloud Computing

Cloud computing is shared pools of configurable computer system resources and higher-level services that can be rapidly provisioned with minimal management effort, often over the Internet. In other words,cloud computing is a type of computing used for delivery of hosted services over the Internet to manage the real time applications . In addition, cloud service providers consist of various deployment models such as a private cloud, a public cloud and a hybrid cloud for accessing resources stored in the cloud. Cloud services are categorized into three categories, namely, Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS). Private clouds are used for storing and sharing of data for one organization and do not share physical resources with others.The resources of the private cloud can be provided externally or inhouse.

Public clouds that are offered by third-party providers are used for sharing and processing physical resources and storage. As public cloud providers maintain numerous users, the processing platform is more scalable and more than that of a private cloud. Hybrid clouds combine public and private clouds, enabling drifting of services between the two clouds using orchestration. In the hybrid cloud, cloud bursting is known as the private cloud that can use public cloud resources when the additional computer is necessary.

C. Internet of Things

The Internet of Things (IoT) is a connection of physical objects with network connectivity that used to collect and exchange data. The IPv6 Internet is one of the most important connectivity of the IoT, as it is not possible to add billions of devices to the IPv4 Internet. In order to achieve an efficient communication between the devices in the internet, a layered architecture is identified with different layers such as Application Layer, Communication Layer, Security Layer, Embedded Layer, Hardware Layer, Integration Layer and DB Layer (Fig. 1).The following technologies are used in the IoT, namely RFID tags, sensors, actuators and mobile phones. These objects interact with each other using unique addressing schemesfor reaching a common goal. Many protocols have been developed in all the layers of ISO stack to enable the operations of IoTdevices.Constrained Application Protocol (CoAP) Routing Protocol for Low- Power and Lossy Networks (RPL) messaging protocols are more familiar. These protocols are designed with energy preservation in mind, along with low compute and memory requirements.

As shown in Fig. 1, the focus of the secure layer is on providingsecurity to things, communications and applications. Embedded security layer focuses ondeveloping security functions for embedded things likesensors, controllers, wearable things and nonwearable things. Communication security layer is used for securing data transfer between things. Database security layer is focused on data protection and securing DB connections to remote/cloud servers. Application security layer is the major security concern in IoT layers. Layer applications are deployed in this and protected with the help of different web security tools. The focus of DB layer is on the database connections from IoT applications to Fog/Cloud remote data centers. Java web services, JDBC, AWS EBS, and AWS EMR are used for providing database connections between IoT applications.

III. RELATED WORK

Wireless Mobile Sensor Network (WMSN) with Internet of Things is playing an important role in the continuous monitoring of individual health conditions [14]. Nowadays, many researchers are using IoT in healthcare application which includes monitoring of continuous heartbeat, blood pressure, blood sugar level and body temperature [15].

Recently, CodeBlue project has been developed by Harvard Sensor Network Lab to monitor individual health [16]. This project uses the following sensors namely EKG, pulse oximeter, EMG, and SpO2 sensor and Mica2 motes to monitor individual health condition [17]. In general, IoT devices are attached with the patients' body to monitor health status in a continuous manner. The fig.1 shows the layered architecture of Internet of Things. It consists of Integration layer, secure layer, communication layer, embedded layer, Hardware layer and Application layer.

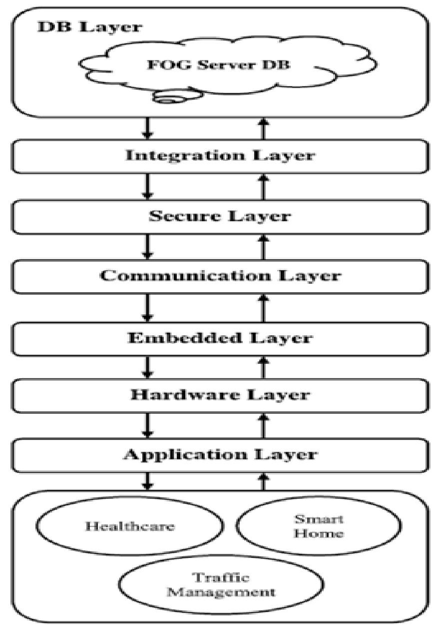

Fig. 1 The layered architecture of Internet of Things

The IoT devices sense the patients' health status and then transfer the patient data to doctors and care holders. This data is most often used for disease diagnosis and patientcare. The useful information mined from the clinical database is used for preventing and protecting the patient health during situations of emergency [18]. University of Virginia, has recently developed the Alarm-Net infrastructure to monitor patient health in an assisted living environment. Various sensors and actuators are used in the Alarm-net framework for observing the health condition of the patient. This framework follows three-tier architecture to observe the patients' health on a continuous basis. BP sensor and ECG sensor are used for monitoring the physiological condition of the patient.

Tier2 architecture is used for observing the environmental parameters such as dust, heat, movementand light. The third and final tier-3 is used for transmitting the original observed signal to the destination using wireless communication protocols and networks infrastructure.

The requirement of integration platform between Cloud computing and Internet of Things based healthcare system, big data processing framework. A tools are required to process he huge data generated from the various IoTdevices, Transmission cost should be reduced for transfer of clinical data between the sensors and the server [19], the need for efficient optimization techniques between the consignment and the header information, the requirement of effective fault tolerance system to manage network failure and emergency situations , the requirement of authentication and authorization schemes to protect the clinical data against the intruders and unauthorized access, MIPv6 and ICMPv6 based network communications and IoT network protocols are required to overcome network delay and latency and star and mesh topologies are also required to manage high loads and single- and multi-hop routing frameworks .

IV. BIG DATA ANALYTICS CAPABILITIES AND IT'S POTENTIAL BENEFITS

A. Capabilities of Big Data Analytics

1. Analytical Capability of Patterns of Care: Analytical capability refers to the analytical techniques typically used in a big data analytics system to process datawith an immense volume, variety and velocity via unique data storage, management, analysis, and visualizationtechnologies. Analytical capabilities in healthcare can be used to identify patterns of care and discover association'sfrommassive healthcare records, thus providing a wide view for evidence-based clinical practice. Healthcare analytical systems provide solutions that fill clinical need and allow healthcare industries to parallel process large volumes of data, manipulate real-time, or near real time data, and capture all patients' visual data or medical records.

2. Unstructured Data Analytical Capability: An analytical process in a big data analytics system starts by acquiring data from both inside and outside the healthcare industries, storing it in distributed database systems, filtering it according to specific discovery criteria, and then analyzing it to integrate related outcomes for the data warehouse.

3. Decision Support Capability: Decision support capability emphasizes the ability to produce results about daily healthcare services to assist managers' decisions and actions. In general, this capability yields sharable information and knowledge such as historical reporting, executive summaries, drilldown queries, statistical analysis, and time series comparisons. Such information can be utilized to provide a comprehensive view to support the implementation of evidence-based medicine, to detect advanced warnings for disease surveillance, and to develop personalized patient care. Some amount of information is deployed in real time (e.g., medical devices' dashboard metrics) while other information (e.g., daily reports) will be presented in summary form.

4. Predictive Capability: Predictive capability is the ability to build and assess a model aimed at generating accurate predictions of new observations, where new can be interpreted temporally and or cross-sectionallydefines predictive capability as the process of using a set of sophisticated statistical tools to develop models and estimations of what the environment will do in the future. By definition, predictive capability emphasizes the prediction of future trends and exploration of new insights through extraction of information from large data sets. To create predictive capability, organizations have to rely on a predictive analytics platform that incorporate data warehouses, predictive analytics algorithms (e.g., regression analysis, machine learning, and neural networks), and reporting dashboards that provide optimal decisions to users. This platform makes it possible to cross reference current and historical data to generate context-aware recommendations that enable managers to make predictions about futureevents and trends.

5. Traceability: Traceability is the capability to track output data from all the system's IT components throughout the organization's service units. Healthcare related data such as activity and cost data, clinical data, pharmaceutical R&D data, patient behavior and sentiment data are commonly collected in real time from payers, healthcare services, pharmaceutical industries, customers and stakeholders outside healthcare.Traditional methods are insufficient when faced with the volumes experienced in this context, which results in unnecessary redundancy in data transformation and movement, and a high rate of inconsistent data. Using big data analytics algorithms, on the other hand, enables authorized users to gain access to large national or local data pools and capture patient records simultaneously from different healthcare systems or devices.

B. Potential Benefits

The use of Big Data is becoming common these days by the companies to perform their peers. Big Data helps the organizations to create new opportunities and entirely new categories of companies that can combine and analyze organization data. The big data analytics has the potential to reduce system redundancy (10 occurrences) and transfer data quickly and securely at different locations (7 occurrences). For example, to aggregate data from approximately 50,000 patients, 6700 appointments, and medical staffs within the hospitals for building the predictive model to tackle the problem of overbooking appointments, Mental Health Center of Denver used a mining table with 3474 attributes to classify the characteristics of appointment for each patient. This mining allows to record patient and appointment information accurately and avoid data duplication and in turn increase a predictive quality.

V. SECURITY IN BIG DATA ANALYTICS

In recent years, IoT devices are continuously generating voluminous data which is often called big data (structured and unstructured data). In general, it is difficult to process and analyze big data for finding meaningful information.

Moreover, data security is a key requirement in healthcare big data system.

Healthcare industries store, maintain and transmit huge amounts of data to support the delivery of efficient and proper care. Nevertheless, securing these data has been a daunting requirement for decades. Complicating matters, the healthcare industry continues to be one of the most affected to publicly disclosed data breaches. In fact, attackers can use data mining methods and procedures to find out sensitive data and release it to public and thus data breach happens. While implementing security measures remains a complex process, the spikes are continually raised as the ways to defeat security controls become more sophisticated.

As a result, it is crucial that organizations implement healthcare data security solutions that will protect important assets while also satisfying healthcarecompliance. To overcome this issue, one has to propose a new architecture for the implementation of Internet of Things (IoT) to store and process scalable sensor data (big data) in health care applications. MF-R with GC architecture for Internet of Things and big data ecosystem was proposed by [20].

VI. CONCLUSION

Through analyzing big data cases, our research has provided a better understanding how healthcare organizations can leverage big data analytics as ameans of transforming IT to gain business value. However, like any other study, ours has limitations. The primary limitation of this study is the data source. One challenge in the health care industry is hat its IT adoption usually lags behind other industries, which is one of the main reasons that cases are hard to find. Although efforts were made to find cases from different sources, the majority of the cases identified for this study came from vendors. Further and better discovery could be done through collecting and analyzing primary data. Given the growing number of healthcare organizations adopting big data analytics, the sample frame for collecting primary data becomes larger. In conclusion, the cases demonstrate that big data analytics could be an effective IT artifact to potentially create IT capabilities and business benefits. Through analyzing these cases,we sought to understand better how healthcare organizations can leverage big data analytics as a means to create business value for health care.We also discussed five strategies that healthcare organizations could use to implement their big data analytics initiatives. In addition to, the better security has to be provided in order to properly utilize the capabilities and potential benefits of Big Data Analytics. For Future work, a secured architecture will be designed to provide data security.

REFERENCES

[1] J. Bertot, "Big data open government and e-government: Issues policies and recommendations", *Information polity*, Vol.19, No.1, pp. 2, 2014.

[2] A. Cardenas, P. Manadhata, and S. Rajan, "Big data Analytics for security", *IEEE Security & Privacy*, Vol. 11, No. 6, pp. 74-76, 2013.

[3] Chen Deyan, and H. Zhao, "Data security and privacy protection issues in cloud computing", *Computer Science and Electronics Engineering (ICCSEE) 2012 International Conference*, Vol. 1, 2012

[4] R. Eynon, "The rise of Big Data: what does it mean for education technology and media research?", *Learning Media and Technology*, Vol. 38, No. 3, pp. 237-240, 2013.

[5] R. Ghosh, "Healthcare Analytics: Data governance and management Analytics Magazine Informs", 2017.

[6] S. Kaisler, F. Armour, J. Espinosa, and W. Money, "Big Data: Issues and Challenges Moving Forward", *System Sciences (HICSS) 46th Hawaii International Conference*, pp. 995-1004, 2013.

[7] H. Patil, and R. Seshadri, "Big data security and privacy issues in healthcare", *2014 IEEE international congress on big data*, pp. 762-765, 2014.

[8] S. Ryu, and T. Song, "Big Data Analysis in Healthcare", *Healthcare Research Information*, 2014.

[9] A.S. Sukumar, R. Natarajan, and R. Ferrell, "Quality of Big Data in health care", *International Journal of Health Emerald*, Vol. 28, No. 6, pp. 621-634, 2015.

[10] *Apache Hadoop*, Sep. 2016, [Online] Available: http://hadoop.apache.org.

[11] W. Raghupathi, and V. Raghupathi, "Big data analytics in healthcare: promise and potential", *Health Information Science and Systems*, Vol. 2, No.1, pp. 3, February 2014.

[12] D. O. Olguin, P. A. Gloor, and A. S. Pentland, "Wearable sensors for pervasive healthcare management", *Proc. IEEE. The 3m International Conference on Pervasive Computing Technologies for Healthcare PervasiveHealth 2009*, pp. 1-4, April 2009.

[13] M. Viceconti, P. Hunter, and R. Hose, "Big Data Big Knowledge: Big Data for Personalized Healthcare", *IEEE Journal of Biomedical and Health Informatics*, Vol. 19, No. 4, pp. 1209-1215, July 2015.

[14] B. Marr, *How Big Data Is Changing Healthcare", Forbes*, 2015, [online] Available at: http://onforb.es/lbfRQ0b.

[15] H. KupwadePatil, and R. Seshadri, "Big Data Security and Privacy Issues in Healthcare", *2014 IEEE International Congress on Big Data*, pp. 762-765, 2014.

[16] W.C. Figg, and H.J. Kam, "Medical Information Security", *International Journal of Security*, Vol. 5, No. 1, pp. 22-34, 2011.

[17] E. Srimathi, and K.A. Apoorva, "Preserving Identity Privacy of Healthcare Records in Big Data Publishing Using Dynamic MR", *International Journal of Advanced Research in Computer Science and Software Engineering*, Vol. 5, No. 4, pp. 968-973, 2015.

[18] A. Tewari, "Security and Privacy in E-Healthcare Monitoring with WBAN: A Critical Review", *International Journal of Computer Applications*, Vol. 136, No. 11, pp. 37-42, 2016.

[19] F. A. Khan, A. Ali, H. Abbas, and N. A. H. Haldar, "A cloud-based healthcare framework for security and patients' data privacy using wireless body area networks", *Procedia Computer Science*, Vol. 34, pp. 511-517, 2014.

[20] N. Mohammed, S. Barouti, D. Alhadidi, and R. Chen, "Secure and private management of healthcare databases for data mining", *Proceedings - IEEE Symposium on Computer-Based Medical Systems 2015*, Vol. 2015, pp. 191-196, July 2015.