HEALTHCARE AND IOT DEVICES: ROLE OF INFORMATION TECHNOLOGY IN THE HEALTHCARE INDUSTRY

Mingalu Pangare Lingkau[1*], Kuo Ling Haoseng[2], Mingalu Pangare Lingkau[3], Mingalu Pangare Lingkau[4], Yong Meng Phaotangu[5]

[1]*Peking University, 5 Yiheyuan Rd, Haidian District, Beijing, China, 100871, mingalu@chinamail.com*
[2]*Tsinghua University, 30 Shuangqing Rd, Haidian District, Beijing, China, 100190*
[3]*Shanghai Jiao Tong University, China, Shanghai, Minhang District, 200240*
[4]*Fudan University, 220 Handan Rd, Yangpu District, Shanghai, China, 200437*
[5]*Nanjing University, Gulou, Nanjing, Jiangsu, China, 210093*

Abstract

Today, wearable health products play a crucial role in most locations, such as constant wellness monitoring of people, street traffic management, weather forecasting, along with smart house. These sensor devices constantly generate massive amounts of data and are kept in cloud computing. This particular chapter proposes Internet of Things design to store and system scalable sensor information for healthcare apps. Proposed architecture comprises 2 primary architecture, specifically, MetaFog-Redirection and Choosing and Grouping architecture. Though cloud computing offers scalable data storage, effective computing platforms must process it. There's a requirement for scalable algorithms to process the big sensor information and recognize the helpful patterns. To conquer this problem, this particular chapter proposes a scalable MapReduce based logistic regression to process such massive quantities of sensor information. Apache Mahout includes scalable logistic regression to system BDA in a distributed way. This particular chapter uses Apache Mahout with Hadoop Distributed File System to process the sensor information produced by the wearable health units.

Keywords

Internet of Things, Big Data Analytics, MetaFog Redirection, Healthcare

JEL Classification

M42, M43

DOI: https://doi.org/10.14311/bit.2022.01.20

Editorial information: journal Business & IT, ISSN 2570-7434, CreativeCommons license published by CTU in Prague, 2022, http://bit.fsv.cvut.cz/

Introduction

The amount and data generation speed of data have increased over the past twenty years in various fields. A statement posted in 2011 by the International Data Corporation states that the general generated and saved information dimensions in the world were 1.8ZB, which enlarged by nearly 9 times within 5 years. Because of the enormous development of earth data, the title of big data is basically used to express substantial datasets. Generally, big data analytics requires superior techniques and tools to shop, process, and evaluate the big volume of information. Large data consists of big unstructured data that need advanced real-time analysis. Consequently, today a lot of the scientists are keen to improve advanced solutions and algorithms to solve the problems when managing big data. To explore hidden values and new opportunities from big data, Yahoo created Hadoop based programs, as well as technologies to keep and process the important information. Today, private businesses will also be enthusiastic about the high potential of big data, along with countless federal organizations declaring important suggestions to accelerate the fundamental information study and uses. 2 top medical journals like Science and Nature also opened unique problems to solve, as well as talk about the difficulties and impacts of big data. Recently, big data played a crucial role in Internet companies, for example Twitter, Facebook, and Google. For instance, Google handles almost hundreds of petabytes, and Facebook creates log data of more than 10 petabytes monthly. A contemporary Chinese business, Baidu, analyzes information of 10 petabytes, and Taobao, a subsidiary of Alibaba, creates information of 10 petabytes for internet trading each day.

"Big data" originally intended the volume, velocity, and range of information, which becomes challenging to evaluate using traditional information processing techniques and platforms [2]. Today, data manufacturing options are quickly developed, like telescopes, sensor networks, high throughput instruments, and streaming machines. These locations create a huge quantity of information. Today, big data plays an important role in various locations, including healthcare, business organization, business, medical investigation, natural resource management, social media, along with public administration. Big data will be classified by 3Vs as follows:

Volume: As Jahan & Sazu (2022) stated the important volume indeed represents great data. Lately, the information model energy sources are augmented, and it brings about a range of information including text, audio, video, and big size pictures. To process the overwhelming amount of information, the conventional data of our processing platforms, as well as strategies, must be increased [5].

Velocity: The speed of the new information has grown dramatically; this velocity provides great information. As Jahan & Sazu (2022) shared the phrase velocity belongs to the information generation speed. The information explosion of social networking has altered, and leads to various information. These days, most people aren't concerned in old posts, as well as notice most popular updates [9].

Variety: The assortment of the information indeed represents big information.

Literature Review

Nowadays, advancement in wireless communication solutions has transformed the standard communication strategies. In the last decade, man-to-machine communication and man-to-man communication are usually used in correspondence environments. The push toward system communications has grown, and machine-to-machine correspondence is recently used in numerous platforms. IoT supplies the platform to support interaction between virtual representations and physical objects [26]. IoT incorporates different technologies and tools like controllers, receptors, and low powered wired and wireless services. Put simply, the web of things is the wired or maybe wireless interconnected different bodily gadgets used to look at, communicate, and transport info with their external environment and inner states. Nowadays, wireless movable sensor systems are usually used

in constant monitoring of healthcare programs, in which sensor systems monitor individuals. This particular article describes the research work associated with the healthcare systems utilizing healthcare sensor networks and wearable sensor equipment [24]. Harvard Sensor Network Lab recently created the CodeBlue task, which strives to keep track of the individuals. CodeBlue task, many health sensors are fixed on the patient's body to sense the patient's health problems. Additionally, these health sensors constantly sense the individual frame and transmit the health problems to the end-user products using wireless technologies [25]. These data are usually used to locate practical patterns to protect the patients from emergency scenarios. The fundamental feature of CodeBlue is extremely simple: a medical doctor or professional queries for affected person healthcare information using their personal digital assistant, which works based upon the publish and subscribe architecture. Lastly, the collected information from the health sensors is creating to a certain channel, and the end user must subscribe to that channel by utilizing their PDA [22]. and laptop Additionally, Wood et al. from the Faculty of Virginia have created the heterogeneous community architecture called AlarmNet. The objective of this particular task is to monitor the individual wellness in the house and assisted living environment. Much more similarly, AlarmNet includes green sensor networks and body sensor networks to effectively sense the particular information. 3 community tiers are utilized in this framework to sense the particular information in the house and assisted living environment. In the first tier, an individual uses a multitude of body sensor products, including SpO2, ECG, and accelerometer, and they feel unique biological details [23]. In the next tier, green receptors, for example dust, motion, temperature, and light, are fixed to the living room to sense the assortment of ecological problems. Lastly, in the 3rd tier, an Internet Protocol based community is created available composed of Stargate gateways named AlarmGate. AlarmNet has used the entire body sensor products to broadcast the single biological details from an individual hop on the next tier [21]. As soon as the next tier obtains the biological information, the fixed sensors ahead the physiological information using the shortest path initially routing protocol on the AlarmGate. The AlarmGate operates like a gateway between IP networks and wireless sensor nodes, and is also attached to a back end server. Ng et al. have created the ubiquitous monitoring atmosphere for implantable and wearable sensors. This particular project is a kind of body sensor system architecture made up of wearable and implantable sensors with the wireless ad hoc system. The primary objective of this particular task is to provide constant monitoring of the patient's health status, as well as predicting emergency problems. Additionally, Chakravorty has created a mobile healthcare project known as MobiCare. This particular task is used to provide timely and continuous monitoring of an individual's biological condition [20].

MobiCare project possibly saves several patients' quality and lives of patient care. MobiCare project involves wearable receptors like ECG, SpO2, and blood oxygen to keep track of the individuals. This particular undertaking appropriate senses the patient's body and transfers the status on the MobiCare client. To send the BSN information on the server, MobiCare customer uses HTTP Post protocol. Additionally, MobiCare server is used to do off-line biological analysis and supports the medical staff for patient care [19].

Blum and Magill have suggested a personalized ambient monitoring challenge to keep track of the patient's psychological health. The PAM task monitors the daily task of individuals with bipolar disorder. Different Bluetooth protocols are used to sign up for mobile cell phones and body sensors. Beyond that, Bluetooth additionally links the personal computer systems to mobile phones. The objective of the mobile cell phones is to aggregate the entire body receptors information and mail it with the personal computers for analysis and storage [18]

Big Data in Healthcare

In recent years, big data analytics also influenced even more in healthcare. As Sazu & Jahan (2022) stated today, healthcare systems are quickly developing medical details that will quickly enlarge the

dimensions of health records, which are accessible electronically. Concurrently, rapid progress and development have been achieved in contemporary healthcare management systems [14]. A recently available analysis expounds 6 use cases of big data to reduce the price of individuals, adverse events, readmissions, triage, and therapy seo for illnesses affecting several organ systems [16]. In one more study, big data, where cases in healthcare were divided into many groups, like medical choice support, delivery and administration, customer behavior, and support services. Jee et al. discussed how to reform the healthcare structure dependent on big data analytics to decide on a proper therapy route, enhancement of healthcare systems, etc [17]. The above mentioned use cases have used the following BDA in healthcare implementation. Patient-centered framework is created based on the fundamental information framework to approximate the quantity of healthcare, patient influence, and dropping readmission rates. Virtual biological human analysis framework is coupled with big data analytics to produce valuable and robust remedies in silico medicine [16].

Digital Epidemiology

Electronic epidemiology allows real time illness surveillance via novel analysis of electronic data. Important analysis and understanding of electronic sources, like social networking, are essential to improve real-time disease surveillance and enable substantial public health solutions [15]. Electronic epidemiology enhances conventional epidemiological scientific studies, like case files, situation reports, ecological scientific studies, cross sectional scientific studies, case management studies, cohort studies, randomized controlled trials, and systematic reviews, as well as meta analysis. While information of electronic epidemiology enhances traditional epidemiology, by conducting computational analytics of real time health and non-health-related electronic sensors and data to derive real time estimates of illness dynamics [14].

Study participants are largely gathered by conventional epidemiological scientific studies to deal with an exploration issue of medical and/or public health significance. Electronic epidemiology additionally uses information energy sources initially collected and/or produced for health and non-health-related requirements [13].

ChatterGrabber is a social networking surveillance program toolkit to determine possible health risks and disease outbreaks, by examining tweets for illness symptoms in particular places. This particular application toolkit is used for illness surveillance in applications that are different, like the EpiDash program, to monitor norovirus outbreaks. Google search queries are examined for surveillance of infectious diseases, like in Google Flu Trends and Google Dengue Trends for dengue and influenza, respectively [9]. These electronic surveillance devices serve as early warning systems for infectious disease outbreaks, as well as complement conventional illness surveillance methods, which have a lag time of dissemination and collection of estimates of illness burden [10]. HealthMap offers real-time infectious disease surveillance by examining internet energy sources of news, list offers, and info provided by worldwide health organizations. It's an automated monitoring instrument for infectious disease outbreaks impacting animal and human health [12].

FRED Software for Disease Modeling

FRED is an ailment modeling program used to deal with a large quantity of information. FRED uses mitigation methods, viral evolution, and private health actions to model the illness outbreak. FRED is an open source framework for pandemic modeling, instead of an unit associated with a specific infectious disease [11]. Geographic regions are utilized in FRED to stand for every individual as agent [8]. Each and every representative has a set of daily behaviors and sociodemographic characteristics, including age, occupation, employment status, sex, and home place, in addition to club membership

for a pair of social contact networks [19]. This particular artificial public information is used to model the illness outbreak of FRED.

NoSQL Databases for Big Data in Healthcare

NoSQL database is used to store large volumes of information in a distributed way. A NoSQL database doesn't comply with any relational schema. NoSQL databases are classified in 4 types, like key value shops, column family repository retailers, document retailers, and graph shops.

Key-value databases store the information based on the answer and value pairs. Generally, key value databases assume the path is the contents and the key is the file. Key-value databases are applicable only for small applications, not for complicated applications. Column family repository stores the massive information in rows as collections of columns. All rows in this database include various columns [7]. A document retailer databases are used to store BDA related to paper format. This type of database is generally used to store semi structured data.

Graph databases include connections, or maybe tips, between nodes. It's a kind of NoSQL database, which uses graph theory to shop, query relationships, and map. Electronic medical record is composed of patient health related info. The following info usually offered in most EHRs are lab results, medication records, billing data, and test specifics. In most cases, laboratory results and billing data can be found as structured "name value pair" data. As Jahan & Sazu (2022) elaborated that lately, an extra quantity of research is attempting to create large data-based electronic phenotype algorithms to recognize illnesses from the EHR. Laboratory information as well as vital signs are mainly organized. It follows a coding program to store the massive amount of lab related information. Nowadays, various algorithms and many dictionaries are developed to decrease the intricacy of laboratory data [22].

To keep the massive quantity of healthcare information, proposed architecture has spent NoSQL based database. Considerable availability of healthcare info has led to increasing quality and accuracy of healthcare delivery [5]. Nowadays, the size and structure of healthcare information are increasing significantly. Hence, the relational database management process isn't ideal for keeping these kinds of massive dimensions of information. Researchers develop a selection of big data solutions to solve such problems [24]. NoSQL databases have considerable benefits, including automobile scaling, much better performance, and substantial accessibility, which will address the limits of relational databases in distributed healthcare systems. Scalable sensor information processing architecture is proposed to this chapter to store and store system body sensor information for healthcare applications. In this proposed structure, electronic health records are collected via medical evaluation, and the outcomes are saved into cloud storage. MapReduce implementation of internet stochastic gradient descent algorithm is used in the logistic regression to build the prediction type. Prior electronic health records are used to train the logistic regression version. Following the conclusion of the instruction procedure, the prediction version will use the present sensor information of the individual to foresee the heart disease status [4].

As Jahan & Sazu (2022) stated the proposed architecture is used for the private health monitoring of people. Each time the respiratory rate, blood sugar, body temperature, blood pressure, and heart rate exceed the normal value of its, the unit sends an alert email with clinical worth to the physician by using wireless community through fog computing. After effectively identifying the authorized user, health information properly shifts to various data clinics offered by various cloud information service suppliers. Meta Cloud Data. Storage architecture is used to transfer the information from uses to cloud information centers, as well as cloud data centers to uses [26].

After the information is transferred from the cloud to the cloud information centers, it's necessary to be saved effectively [6]. Today, data generation options are enhanced, like high throughput instruments, telescopes, sensor networks, and streaming machines, and these locations create

substantial quantities of information. To solve this particular problem, Hadoop Distributed File System, is used in this stage to keep such massive quantities of information. This phase additionally categorizes the information into various stores and levels them into several data centers. Choosing and grouping architecture is stuck with MetaFog Redirection architecture to secure integration of fog to cloud computing to protect great details against intruder. Clinical information is kept in several cloud information centers based on categorization and importance. Information categorization is classified into 3 levels, like sensitive, crucial, and typical. Each categorized information should be kept in various data centers. Proposed architecture is used to redirect the user petition to the correct information center.

MetaFog-Redirection architecture with Grouping & Choosing architecture is suggested to this chapter. The objective of the suggested GC architecture is to integrate fog computing with cloud computing [1]. The next integrations are done in the GC architecture. It provides application integration, information transfer from fog servers to cloud information clinics, and protection systems for interaction between fog computing and cloud computing. Logistic regression is used in the proposed MetaFog architecture to foresee the condition according to the historic documents. Logistic regression is often used in datasets where a reliant variable is dichotomous. Logistic regression is used in this chapter to build the prediction type and find the connection between independent and dependent variables [2]. MapReduce implementation of stochastic gradient descent with logistic regression is displayed below:

MapReduce framework, mappers instantly give M. After splitting into M block sizes, with reducer, get Vi printer. Stochastic gradient descent technique is used in the reducer feature to get the effective weights and minimize the error function. In the above mentioned algorithm, Map and lower technique runs T iterations. Each iteration makes weights kq wherein kϵ. The last weight

θt + one is calculated depending on the average of kq. The average of body weight θt + one should be M described by q t +1 = one kq. M*k =1

Apache Mahout is a library of scalable machine learning algorithms. Apache Mahout is implemented in addition to Apache Hadoop and works with the MapReduce paradigm. Machine learning is a kind of artificial intelligence focused on enabling devices to find out without being explicitly programmed, and it's often used to boost future performance based on prior outcomes [3]. Big information is saved on the HDFS; Apache Mahout is used to perform machine learning algorithms that extract substantial patterns from datasets. The above-mentioned MapReduce based logistic regression may be accomplished with the aid of Apache Mahout. Mahout implementation of logistic regression utilizing SGD supports the next command lines:

• Training the product • Mahout org.apache.mahout.classifier.sgd.TrainLogistic - passes one - speed one - lambda

• 0.5 input heart.csv - capabilities twenty one - output heart.model - target Num - groups two -

• Predictors Thalach Trestbps Fbs - kinds n.

• Testing the product Mahout org.apache.mahout.classifier.sgd.RunLogistic - type in heart.csv - model. heart.model - auc - scores - confusion

Model Development with Cleveland Heart Problems Database

Logistic regression is trained using the previous medical information, as well as sensor information of the individuals. The prediction version can work with the present sensor information of the individual to foresee the heart disease status. The prediction design uses the present sensor information obtained from body sensor products through big data technologies and cloud. Cleveland Heart Disease Database is the de facto collection for heart disorders research. This particular database

is used to teach the proposed prediction version, and it has seventy six attributes. Most published experiments talk about working with a subset of fourteen of them. The Cleveland database is popular by devise learning researchers till date.

Conclusion

This particular chapter proposes Internet of Things design to store and system scalable sensor information for healthcare apps. Proposed architecture comprises 2 primary sub-architectures, specifically, MetaFog-Redirection and Choosing and Grouping architecture. MapReduce-based logistic regression is applied with the aid of Apache Mahout. Logistic regression is educated utilizing the previous medical documents from the Cleveland Heart Disease Database, as well as sensor information of the individuals. The prediction version can work with the present sensor information of the individual to foresee the heart disease status. In this particular analysis, the prediction design uses the present sensor information obtained from body sensor products through the cloud, along with large data solutions.

References

[1] ARUNACHALAM, D., Kumar, N. & Kawalek, J. P., 2018. Understanding big data nalytics capabilities in supply chain. Transportation Research Part E: Logistics and Transportation Review, Volume 114, June , Pages , pp. 416-436.

[2] JAHAN, S. A., & Sazu, M. H. (2022). Role of iots and Analytics in Efficient Sustainable Manufacturing of Consumer Electronics. *International Journal of Computing Sciences Research*, *6*.

[3] BANERJEE, A., Bandyopadhyay, T. & Acharya, P., 2013. Data analytics: Hyped up aspirations or true potential?. Vikalpa, Volume 4, pp. 1-12.

[4] CHEN, Y., 2017. Integrated and Intelligent Manufacturing: Perspectives and Enablers. Engineering, Volume 3, pp. 588-595.

[5] SAZU, M. H., & Jahan, S. A. (2022). How Analytics Can Improve Logistics And Supply Chain In Multinational Companies: Perspectives From Europe And America. *Business Excellence and Management*, *12*(3), 91-107.

[6] ECKSTEIN, J. Et al., 2016. A Comparison of two Predictive Approaches to Control the Longitudinal Dynamics of Electric Vehicles. Procedia Technology, Volume 26, pp. 465-472.

[7] GE, Z., Song, Z., Ding, S. X. & Huang, B., 2017. Data Mining and Analytics in the Process Industry: The Role of Machine Learning. IEEE Access, Volume 5, pp. 20590-20616.

[8] GUNASEKARAN, A., Yusuf, Y. Y., Adeleye, E. O. & Papadopoulos, T., 2018. Agile manufacturing practices: the role of big data and business analytics with multiple case studies. International Journal of Production Research, 56(1-2), pp. 385-397.

[9] SAZU, M. H., & Jahan, S. A. (2022). The impact of big data analytics on supply chain management practices in fast moving consumer goods industry: evidence from developing countries. *International Journal of Business Reflections*, *3*(1).

[10] HUTTER, T., Haeussler, S. & Missbauer, H., 2018. Successful implementation of an order release mechanism based on workload control: a case study of a make-to-stock manufacturer. International Journal of Production Research , 56(4), pp. 1565-1580.

[11] JI, W. & Wanga, L., 2017. Big data analytics based fault prediction for shop floor scheduling. Journal of Manufacturing Systems, Volume 43, pp. 187-194.

[12] KETOKIVI, M. & Choi, T., 2014. Renaissance of case research as a scientific method. Journal of Operations Management, 32(5), pp. 232-240.

[13] KUMAR, A., Shankar, R., Choudhary, A. & Thakur, L. S., 2016. A big data mapreduce framework for fault diagnosis in cloud-based manufacturing. International Journal of Production Research, 54(23), pp. 7060-7073

[14] JAHAN, S. A., & Sazu, M. H. (2022). The Impact of Data Analytics on High Efficiency Supply Chain Management. *CECCAR Business Review*, *3*(7), 62-72.
[15] LINDSTRÖM, J., Larsson, H., Jonsson, M. & Lejon, E., 2017. Towards intelligent and sustainable production: combining and integrating online predictive maintenance and continuous quality control. Procedia CIRP of The 50th CIRP Conference on Manufacturing Systems, Issue 63, pp. 443-448.
[16] SAZU, M. H., & Jahan, S. A. (2022). Impact of big data analytics on distributed manufacturing: does big data help?. *Journal of process management and new technologies*, *10*(1-2), 70-81.
[17] MEHTA, P., Butkewitsch-Choze, S. & Seaman, C., 2018. Smart manufacturing analytics application for semi-continuous manufacturing process–a use case. Procedia Manufacturing, Volume 26, pp. 1041-1052.
[18] PAUL, D., Zhong, W.-D. & Bose, S. K., 2016. Energy efficient cloud service pricing: A two-timescale optimization approach. Journal of Network and Computer Applications, Volume 64, pp. 98-112.
[19] SAZU, M. H., & Jahan, S. A. (2022). Can big data analytics improve the quality of decision-making in businesses?. *Iberoamerican Business Journal*, *6*(1), 04-27.
[20] SADATI, N., Chinnam, R. B. & Nezhad, M. Z., 2018. Observational data-driven modeling and optimization of manufacturing processes. Expert Systems with Applications, Volume 93, pp. 456-464.
[21] SEURING, S. & Müller, M., 2008. From a literature review to a conceptual framework for sustainable supply chain management. Journal of Cleaner Production, 16(15), pp. 1699-1710.
[22] AKTER, J. S., & Haque, S. M. (2022). Innovation Management: Is Big Data Necessarily Better Data. *Management of Sustainable Development*, *14*(2), 27-33.
[23] SHAO, Y., Li, C. & Luo, J. G. :. J. Z. Y., 2018. Efficient jobs scheduling approach for big data applications. Computers & Industrial Engineering, Volume 117, pp. 249-261.
[24] SAZU, M. H., & Jahan, S. A. (2022). How Big Data Analytics Impacts the Retail Management on the European and American Markets?. CECCAR Business Review, 3(6), 62-72.
[25] SHIN, S.-J., Woo, J. & Rachuri, S., 2014. Predictive analytics model for power consumption in manufacturing. Procedia CIRP, Volume 15, pp. 153-158
[26] SAZU, M. H., & Jahan, S. A. (2022). How Big Data Analytics is transforming the finance industry. Bankarstvo, 51(2), 147-172.