RA Journal of Applied Research
ISSN (e): 2394-6709
||Volume||4||Issue||03||Pages-1514-1518||March-2018||
Index Copernicus ICV: 74.25, DOI: 10.18535/rajar/v4i3.05

A Survey on Real Time Data Analysis from Sensors Using R Programming Techniques

S. Keerthana[1], K. Manish[2], R. Shwetha[3], S. Kayalvizhi[4], S. Kalpana Devi[5]

[1,2,3]UG Student, Computer Science and Engineering, Easwari Engineering College, Chennai, India
[4] Professor ,Computer Science Department, Easwari Engineering College, Chennai, India
[5] Assistant Professor, Computer Science Department, Easwari Engineering College, Chennai, India

Abstract
Health care is a vast ecosystem where automation becomes inevitable. Development of Internet of Things (IoT) in healthcare has drastically changed the conventional model of working towards digital expansion. Transformation of lifestyle has also made a general shift in care from hospital environments to private environments like patient's home. Nowadays, society prefers personalized care and experience towards them for satisfaction along with preservation of privacy. Healthcare data are broadly classified into numerous fields like personal data, Pharmaceutical data, Clinical data etc. This survey paper focuses on obtaining raw information from various sensors and performs data analysis techniques for inferring knowledge about patient's data in real time. All these data obtained from the sensors are analyzed for knowledge inference where clustering and classification tends to be useful. In clustering, k-means algorithm is used because of its ability to produce uniform clusters irrespective of the input size and classify the resultant sets using naive Bayes classifiers. Security of data is also improved by using Cloud Storage with Attribute Based Encryption technique. The aim of this paper is to produce more accuracy and higher level of privacy data sets based on comparison.

Keywords: Clustering, Classification, K-means, IoT, Cloud Storage, Attribute Based Encryption

INTRODUCTION

Human health is being considered as a dynamic process because of its ability to change regularly. Health is a level of functional efficiency of a person's mind, body and spirit. The healthcare transformation has enhanced the entire experience of both patients and doctors. Ability to collect, store and process data is the ultimate development of technological age. Advancement in Internet of things and Data Analytics has made it possible to collect and process huge amount of data which helps in obtaining valuable insights. Development of technology has enabled generation large quantity of data in numerous research fields. Researches in data mining have provided a rise in data storage and manipulation of the earlier stored data for further decision. Mining plays a vital role for uncovering new trends in healthcare organization that in turn is useful for all the parties related to this field. It's a new powerful technology which is of high interest in the modern world. Data mining is a sub field of computer science that uses already existing data in several databases to remodel it into new field of research. For categorizing the patients exhibiting similar kind of diseases or health issues, data mining technologies benefits the healthcare organization by providing them aid in selecting effective treatments. In addition, it is also helpful in predicting the duration of stay of patients in hospital for medical diagnosis and creates a setup for

RA Journal of Applied Research
ISSN (e): 2394-6709
||Volume||4||Issue||03||Pages-1514-1518||March-2018||
Index Copernicus ICV: 74.25, DOI: 10.18535/rajar/v4i3.05

effective data system management. According to a survey, nearly 13% of hospitals are not able to track the records of their patients. This data contains details regarding hospitals, patients, medical claims, treatment cost etc. Electronic Medical Records manages both patients' personal and medical records in an online repository. This helps to provide more focused care, accurate detection and also the ability to identify the medical trends of each individual. As data contains sensitive information, it is to be authenticated and authorized. Thus, we use encryption techniques to preserve privacy.

LITERATURE SURVEY

In paper [12], Sensible objects are made the ultimate building blocks within the development of cyber-physical sensible pervasive frameworks using Internet of Things (IoT). The IoT unifies a kind of application domains, as well as health care. The IoT revolution is redesigning contemporary health care with promising technological, economic, and social prospects. This paper reviews applications and industrial trends in IoT-based healthcare solutions, and also surveys the advances in IoT-based health care technologies. In addition to IoT services, IoT applications deserve proximate attention. Services are used to develop applications, whereas applications are directly used by users and patients. Therefore, categorizing services to be developer-centric and applications to be user-centric. Also, several gadgets, wearables and other healthcare devices presently obtainable within the market are discussed. This product is viewed as IoT innovations that can cause varied healthcare solutions. Subsequent subsections address various IoT-based healthcare applications, including both single- and clustered-condition applications. Standardization, network type, business models, quality of service and health data protection are some of the issue that the paper fails to address instigating further the IOT based researches. Also, approaches for solving IOT based issues practically are not mentioned.

Paper [8] deals with the growing trends in using cloud for large-scale data storage. Increasing popularity in cloud storage has led many healthcare organizations to move towards Electronic health records (EHRs) in cloud-based storage systems. The cloud server core is a massive database that has enough space to accommodate huge amounts of data from various sensors for long period to trace the history of the user. The database is interfaced to a large set of data analysis algorithms and data are often accessed through internet using dynamic web pages. Several cryptographic access control schemes are planned to safeguard the security of data stored within the cloud by integrating cryptographic techniques with access control models. This paper tends to contemplate a unique access control encryption technique to create a secure and versatile large-scale EHR system where access control policies are enforced in a cloud environment. The benefits of this system are the framework address which has its own unique challenges brought by multiple PHR owners and users, in that it greatly reduces the complexity of key management while enhancing the privacy guarantees compared with previous works. It utilizes ABE to encrypt the PHR data, so that patients can allow access not only by personal users, but also various users from public domain. The disadvantage of the system is that data sets in public cloud must be encrypted efficiently. Encryption of data can be shared only in coarse-grained level.

Paper [1] describes that k-means algorithm has clad to be extremely powerful in making clusters in several sensible applications in developing areas. This algorithm is employed to sporadic K because of its initial center points from the consummate dataset. Then it tends to calculate the Euclidean

RA Journal of Applied Research
ISSN (e): 2394-6709
||Volume||4||Issue||03||Pages-1514-1518||March-2018||
Index Copernicus ICV: 74.25, DOI: 10.18535/rajar/v4i3.05

distance of every data from the initial cluster Centre's, choose the sample that is most proximate, and then assign it to the felicitous cluster. The Centre is updated till the mean square error becomes minimum or the cluster Centre's ceases moving and become available in the Centre. Now, all the data points have minimum distance from the Centre Point. [7] The K-means++ algorithm is used to surmount the matter to arbitrarily choose initial cluster Centre which will increase the amount of iterations. Primary aim of this algorithm is to require those data points as initial Centre's that are as far-off from one another as attainable [11]. Fuzzy C-Means (FCM) is however termed as soft cluster. The FCM algorithm sanctions a data point to belong to any or all the clusters with the membership in between zero and one. If the data point is more proximate to the cluster Centre its membership are more towards a specific cluster. At the surcease of every iteration each the Centre and membership must be updated accordingly [5]. The study shows that K means algorithm takes lesser time than FCM and also produces more effective and tight clusters. Implementation procedures are also simpler than other conventional techniques. But the major drawback of using K Means algorithm is to determine right value of K for productive clustering .FCM produces high complexity than K Means. Time utilization is a limitation due to complexity is also a fallback for K Means++ algorithm.

In paper [14], Medical records within the type of text have patient's detail like name, address, age, zip code, diseases. The sensitive information is name, disease. Here privacy techniques [3] are used to defend the sensitive data from unauthorized users. On the other hand, analyzing data by third party means that a new threat of misusing that privacy. As data contains sensitive information regarding patient and medical organization. The cryptographic algorithm [6] constructs to implement privacy preserving data mining algorithm. The cryptographic techniques are used wherever multiple parties are concerned and to work out the non-sensitive mining results. The process of partitioning a group of objects within the dataset into subsets known as clustering. Each subset may be a cluster, such that attribute in a clusters are similar to each other. K-means clustering algorithm excels in finding accuracy of anonymized table. In K-means clustering algorithm, the term K represents variety of clusters. This algorithm partitions the anonymized dataset into subsets known as clusters. The centroid of a cluster is its center point. The mean of its cluster is calculated based on number of similar objects inside the cluster. The objective of k-means is to minimize the squared error which is done by replacing every object by its nearest centroid. K-Mean is also an good quantization algorithm. The issue is that rescaling of data sets may completely change the result and the order of data might also have strong impact over the output result.

Paper [13] proposes that the age of cloud computing healthcare industry is also moving slowly towards cloud storage. Within the health industry, cloud refers to storing patient data in the kind of electronic record on the remote storage instituted in the computer storage. The security and privacy concern to the patient health record are main problems that limit the adoption of cloud storage in the healthcare. Access control model could be a productive approach to guarantee the data security within the healthcare framework. In ABE scheme [2], data is encrypted and encryption relies on the policy that is based on owner's criteria for accessing data. The user will access data as long as he satisfies attribute policy. In this model, monotonic attributes are used for controlling user access to the system which restricts this ABE model during a real environment. The advantage is that Role based access control maximizes the operational efficiency and provides

RA Journal of Applied Research
ISSN (e): 2394-6709
||Volume||4||Issue||03||Pages-1514-1518||March-2018||
Index Copernicus ICV: 74.25, DOI: 10.18535/rajar/v4i3.05

access tree structure. It also reduces the overall computational overhead of the system. But, privacy and security threats are major concern due to the vulnerability caused by hackers and data without encryption may provide unauthorized access to other users. Selection of authorized clouds also plays a vital role.

TABLE ON LIERATURE SURVEY

S.NO	TITLE	AUTHOR	ADVANTAGES	ISSUES
1	The Internet of Things for Health Care: A Comprehensive Survey	S.M.Riazul Islam, Daehan Kwak, MD. Humaun Kabir, Mahmud Hossain and Kyung-Sup Kwak	Advancement in IoT based healthcare technology	Standardization, network type, business models, quality of service and health data protection
2	Secure Role-Based Cloud Storage System For Encrypted Patient-Centric Health Records	Lan Zhou,Vijay Varadharajan, K. Gopinath	Reduces complexity of key management and increases privacy	Data sets in public cloud be encrypted and encryption of the data can be shared solely in coarse-grained level
3	A Comparative Study of K-Means, K-Means++ and Fuzzy C-Means Clustering Algorithms	Akanksha Kapoor ,Abhishek Singhal	K means take lesser time than FCM and provides tight clusters	Determine right value of K, FCM produces higher complexity than K-Means
4	Using K-Means Clustering Algorithm for Handling Data Precision	P. Suganthi,K. Kala, C. Balasubramanian	Minimizes squared error and provides good quantization algorithm	Rescaling of data sets may completely change the result
5	Comparative Study on Attribute-Based Encryption for Health Records in Cloud Storage	Siddhesh Mhatre, Ananth Nimkar	Role based access control maximizes the operational efficiency and provides access tree structure	Privacy, security threats due to vulnerability of hackers.

CONCLUSION

Security of healthcare data is one amongst the essential problems in healthcare industry. IoT brings accurate leads to detection of pulse which is the primary resource for identifying all kinds of diseases. Generally, access control models and encryption based techniques fail to protect health care data in cloud. A survey of existing Attribute based encryption schemes for specific healthcare industry applications is been administered. The classification done by the k means provides precise clusters and helps in prediction of disease in future. This will facilitate healthcare organizations to share their record securely outside trust limit.

REFERENCES

1. Akanksha Kapoor, Abhishek Singhal, "A Comparative Study of K-Means, K-

RA Journal of Applied Research
ISSN (e): 2394-6709
||Volume||4||Issue||03||Pages-1514-1518||March-2018||
Index Copernicus ICV: 74.25, DOI: 10.18535/rajar/v4i3.05

Means++ and Fuzzy C- Means Clustering Algorithms", 3rd IEEE International Conference on "Computational Intelligence and Communication Technology", IEEE-CICT 2017

2. K. Emura, A. Miyaji, A. Nomura, K. Omote, and M. Soshi, "A ciphertext-policy attribute-based encryption scheme with constant ciphertext length," in International Conference on Information Security Practice and Experience. Springer, 2009, pp. 13–23
3. B.Fung, K. Wang, R. Chen, and P. Yu, "Privacy Preserving DataPublishing: A Survey of Recent Developments," ACM Computing Surveys, 2010
4. S. Ghosh, S.K. Dubey, "Comparative Analysis of K-Means and Fuzzy C-Means Algorithms" in International Journal of Advanced Computer Science and Applications, Vol. 4, 2013:pp. 35-39
5. A. Gupta ; H. Shivhare ; S. Sharma, "Recommender system using fuzzy c-means clustering and genetic algorithm based weighted similarity measure" in International Conference on Computer, Communication & Control,2015:pp 1-8
6. Helger Lipmaa," Cryptographic Techniques in Privacy Preserving Data Mining", University College London, Estonian Tutorial 2007
7. K.M. Kumar, Dr. A. R.M. Reddy, "A Fast K-Means Clustering Using Prototypes for Initial Cluster Center Selection", IEEE 9th International Conference on Intelligent Systems and Control (ISCO), 2015
8. Lan Zhou, Vijay Varadharajan, K. Gopinath ,"Secure Role-Based Cloud Storage System For Encrypted Patient-Centric Health Records", The Computer Journal, Volume 59, 3 November 2016
9. K. Lin, X.Li, J. Chen, Z. Zhang, "A K-means Clustering with Optimized Initial Center Based on Hadoop Platform" in The 9th International Conference on Computer Science & Education, 2014:pp. 263-266
10. X. Liu, H. Zhu, J. Ma, J. Ma, and S. Ma, "Key-policy weighted attribute based encryption for fine-grained access control," in Communications Workshops (ICC), 2014 IEEE International Conference on. IEEE, 2014, pp. 694–699
11. Z. Min, Kai-fei, "Improved research to k-means initial cluster centers" in Ninth International Conference on Frontier of Computer Science and Technology, 2015, pp. 349-353
12. S.M. Riazul Islam1, (Member, IEEE), Daehan Kwak2, MD. Humaun Kabir1, Mahmud Hossain3, AND Kyung-Sup Kwak1, (Member, IEEE), "The Internet of Things for Health Care: A Comprehensive Survey", 2169-3536 2015 IEEE.
13. Siddhesh Mhatre, Ananth Nimkar, "Comparative Study on Attribute-Based Encryption for Health Recordsin Cloud Storage", 10.1109/RTEICT.2017.8256677, IEEE,15 January 2018
14. P. Suganthi, K. Kala, C. Balasubramanian, "Using K-Means Clustering Algorithm for Handling Data Precision", IEEE, 10.1109/ICCTIDE.2016.7725374, 31 October 2016
15. M. Yedla, S. R. Pathakota, T M Srinivasa, "Enhancing K-means Clustering Algorithm with Improved Initial Center" in International Journal of Computer Science and Information Technologies, Vol. 1 (2), 2010:pp. 121-125