International Journal of Technical Innovation in Modern Engineering & Science (IJTIMES)
Impact Factor: 5.22 (SJIF-2017), e-ISSN: 2455-2585
Volume 4, Issue 8, August-2018

(Hybrid PSO-Fuzzy Logic) Data Mining Over Semantically Secure Relational Data

Ajitendra Kumar Singh[1], Dr.Akash Saxena [2]

[1]*M.Tech CSE, Compucom Inst. Of Tech. & Management,jaipur*
[2] *Prof CSE Compucom Inst. Of Tech. & Management,jaipur*

Abstract— Data mining is a promising & emerging technology that albums knowledge discovery in huge amounts of data. Numerous abstracted & implemented sachems to process quires on relation data in a secure way have been demonstrated. As cloud computing is gaining momentum, more & more data owners are able to outsource data storage & even data processing capabilities. As user data privacy concern is of primary importance, the sensitive data should be encrypted before being stored on cloud services. Also all the data mining operations, such as clustering are performing on encrypted data only.

This work demonstrated application of hybrid pso (particle swarm optimization) fuzzy logic clustering technique in contrast to effectivively proved KNN method. Cloud data storage & quarry client are implemented in matlab & communication and TCP/IP protocol. All data stored on cloud is encrypted, & all operations on cloud data happen on encrypted data only. The user quarry data is also encrypted, & results are also encrypted, this provides utmost data privacy & security.

Keywords: relational data mining, cloud data, cryptography, privacy preserving (cloud mining), K-NN classifier, hybrid fuzzy, pso classifier.

I. INTRODUCTION

A social database is a gathering of information things composed as an arrangement of formally-depicted tables from which information can be gotten to or reassembled in a wide range of courses without reorganizing the database tables. The social database was imagined by E. F. Codd at IBM in 1970. Initially, databases were level. This implies the data was put away in one long content record, called a tab delimited document. Every passage in the tab delimited record is isolated by a unique character, for example, a vertical bar. Every section contains numerous snippets of data (fields) about a specific question or individual gathered together as a record. The content document makes it hard to look for particular data or to make reports that incorporate just certain fields from each record. Here's a case of the document made by a level database You can see that you need to seek consecutively through the whole record to accumulate related data, for example, age or pay. A social database enables you to effortlessly discover particular data. It additionally enables you to sort in light of any field and create reports that contain just certain fields from each record. Social databases utilize tables to store data. The standard fields and records are spoken to as segments (fields) and lines (records) in a table. With a social database, you can rapidly analyze data due to the course of action of information in sections. The social database display exploits this consistency to assemble totally new tables out of required data from existing tables. At the end of the day, it utilizes the relationship of comparative information to build the speed and adaptability of the database. The "social" some portion of the name becomes possibly the most important factor due to scientific relations. A normal social database has somewhere in the range of 10 to in excess of 1,000 tables. Each table contains a section or segments that different tables can key on to assemble data from that table. For instance By storing this information in another table, the database can create a single small table with the locations that can then be used for a variety of purposes by other tables in the database. A typical large database, like the one a big Web site, such as Amazon would have, will contain hundreds or thousands of tables like this all used together to quickly find the exact information needed at any given time.

Relational databases are created using a special computer language, structured query language (SQL) that is the standard for database interoperability. SQL is the foundation for all of the popular database applications available today, from Access to Oracle.

II. OBJECTIVES OF STUDY

(i). Design of a highly efficient classification algorithm that caters to multiplicity of applications ranging from finance to medicine by hybridization of KNN & Cloud techniques.

(ii). The system will be able to cater to the need of security, privacy in classification system(s), by sending encrypted sample data over internet protocol to cloud server & cater to cloud based classification system.

(iii). Use of advanced low computational cost, encryption techniques such as Blowfish algorithm to encrypt allows for usage of our system on low computational power mobile devices.

(iv). The proposed system will be implemented as two matlab models, one acting as a cloud interface server, which accesses the cloud, performs decryption, KNN classification etc. The other matlab model being the client side, which inputs user query data pattern, sends the query to server wait for the acknowledgment and result from server.

(v). Use of distributed cloud computing to provide for high computational throughput by utilising multiple processors simultaneously, allows the back bone cloud server, to host a large number of audience.

(vi). Use of matlab cloud storage platform for storage of encrypted relational data in database authentication details and encryption keys.

(vii). Our proposed system fulfils the need of modern day classification system employing high performance distributed KNN comparing, cloud storage, mobile device access, simultaneously. Protecting confidentiality of data user's input query and encapsulates data access parameters.

III. LITERATURE REVIEW

Ms. Anjali J. Rathod Prof. V. S. Mahalle Described That Data mining is a powerful new technique to discover knowledge within the large amount of the data. A number of theoretical and practical solutions to query processing have been proposed under various scenarios. With the recent popularity of cloud computing, data owners now have the opportunity to outsource not only their data but also data processing functionalities to the cloud. Because of data security and personal privacy concerns, sensitive data (e.g., medical records) should be encrypted before being outsourced to a cloud, and the cloud should perform query processing tasks on the encrypted data only. These tasks are termed as Privacy Preserving Query Processing (PPQP) over encrypted data. These protocols protect the confidentiality of the stored data, user queries, and data access patterns from cloud service providers and other unauthorized users. Several queries were considered in an attempt to create a well-defined scope. These queries included the k-Nearest Neighbor (kNN) query, advanced analytical query, and correlated range query. This paper presents protocols utilize an additive cryptography base privacy preserving data mining technique at different stages of query processing to achieve the best performance all computations can be done on the encrypted data .

The cloud computing paradigm has recently revolutionized the organization’s way of operating their data, particularly in the way they store, access, and process data. Therefore, due to the rise of various privacy issues, sensitive data need to be\ encrypted before being outsourced to the cloud. Using encryption as a way to achieve data confidentiality may cause another issue at the cloud during the query evaluation. This paper proposed Privacy-Preserving Data Mining (PPDM) protocols to facilitate different types of queries, namely, the k-Nearest Neighbor (k-NN) query. In the proposed PPDM protocols, once the data owner has outsourced encrypted data to the cloud, the user does not participate in any computations. The proposed protocols utilize additive cryptography based PPDM technique at different stages of query processing to achieve the best performance. Encryption is not the only way to protect data confidentiality, and a variety of different techniques, such as randomization and secret sharing, exist. One can plan to investigate whether these techniques are more efficient and scalable than the encryption based solutions. One can develop PPDM protocols that are secure under the malicious model.

IV. METHODOLOGY

A. System Block Diagram of Operation.

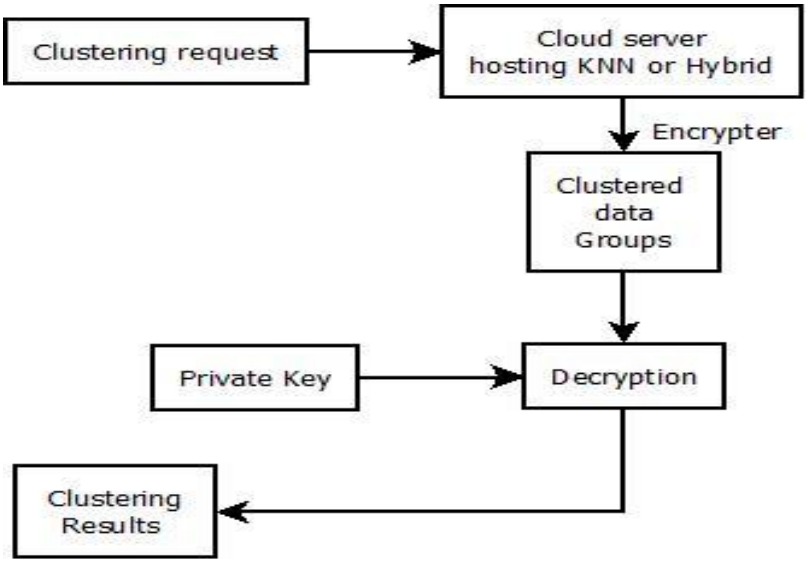


Fig 3.2 System Block Diagram

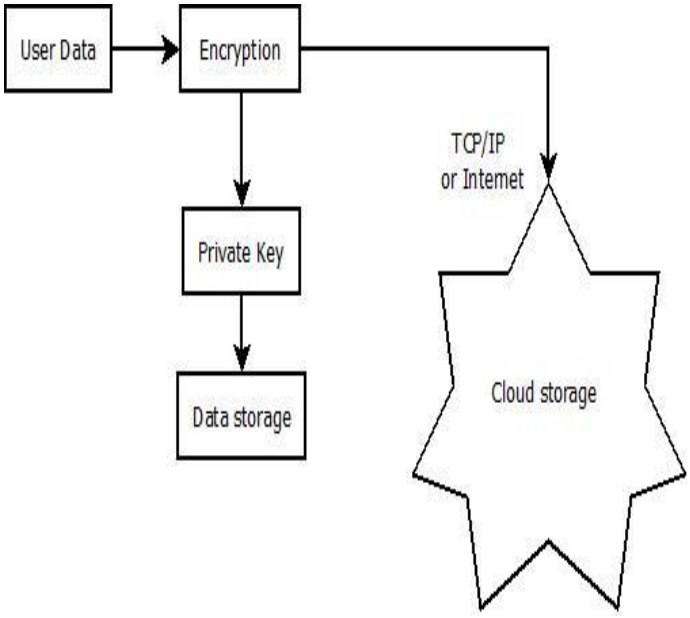
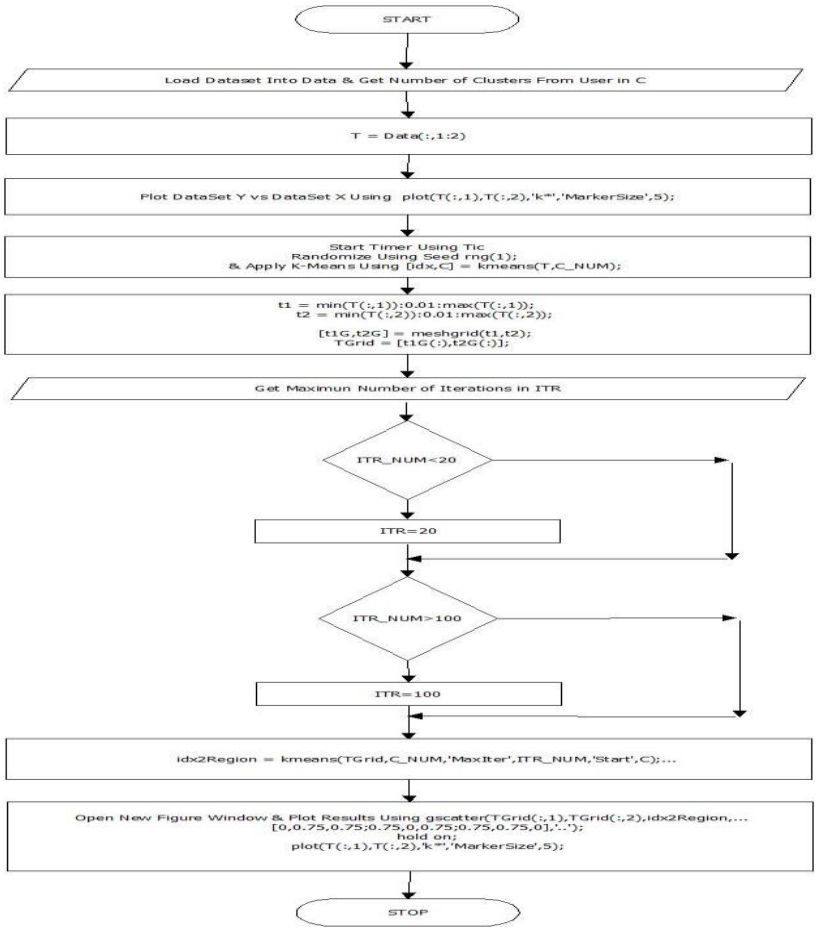


Fig 3.3 System Block Diagram

3.3 K-NN clustering Flowchart



V. RESULTS

A. Process flow

The proposed system will be implemented as two matlab models, one acting as a cloud interface server, which accesses the cloud, performs decryption, KNN classification etc. The other matlab model being the client side, which inputs user query data pattern, sends the query to server wait for the acknowledgment and result from server. We use of distributed cloud computing to provide for high computational throughput by utilising multiple processors simultaneously, allows the back bone cloud server, to host a large number of audience. We take use of matlab cloud storage platform for storage of encrypted relational data in database authentication details and encryption keys. Our proposed system fulfils the need of modern day classification system employing high performance distributed KNN comparing, cloud storage, mobile device access, simultaneously. Protecting confidentiality of data user's input queries and encapsulates data access parameters.

Main menu in Clint server 8 processes, we will choose the process according step wise and we select fisher iris database first, second k-means database set and after select data wine set user can defined the data set after we encrypted data and upload the data to server k-means clustering is performed and performed adaptive PSO clustering employing fuzzy logic, above two process is taken for user encrypted data prevention.

B. Datasets

To analysis the traits of the particular clustering methods, four experimental datasets are used in this thesis. These datasets are taken from UCI machine learning repository [32] which signifies examples of data with low, medium and high dimension. The description of datasets are defined in table 5.1 and in different sections

TABLE 5.1 Summaries of Datasets

Name of Data Set	Number of classes	Number of features	Size of data set
Fisher Iris	3	4	150
K-Means	4	4	560
Wine	3	13	178

C. Comparative Result

Table 5.2 summarizes the execution time (in seconds) obtained from the three algorithms ie K-Means, APSO and from the four datasets. Result has been obtained by using 100 iterations and 10 population size .The best results are highlighted with bold fonts.

Table 5.2 Execution Time

Dataset	K-Means	APSO
Iris	3.82	**1.05**
K-Means	6.61	**2.84**
Wine	1.19	1.34

Figure 5.21 Execution time comparisons between K- Means, APSO on Iris, K-Means and Wine dataset

As graph shows when we use datasets of low dimension, then APSO algorithm takes less time to execute than the existing k-means and APSO algorithm and when we use the high dimensional dataset than it is again true that APSO takes less time to execute than other two algorithms.

CONCLUSIONS

Cloud computing has gained widespread acceptance in the past few years, and has indeed revolutionized data storage a remote services. As the user data may contain private & sensitive information, the data is generally encrypted before uploading to the cloud. A challenge presented by encrypted data is out sourcing of computation of data, because only data mining operation or quarry operation performed at the cloud end must be performed on encrypted data. K-NN classification is a popular & robust, PPDM (private-preserving data mining) technique, that is studied in contrast to the proposed hybrid fuzzy PSO clustering, As results demonstrated, the hybrid fuzzy-PSO clustering algorithm significantly improves performance & reduces time of classification. Truly random mixed key generation cryptography has been employed to provide for high data security & integrity. The results of clustering are encrypted to be decrypted at the user end using user's private key with which the user encrypted data while uploading to cloud.

FUTURE SCOPES

The proposed system demonstrate privacy preceding data mining (PPDM) over cloud data, but as cloud storage is a relatively nascent & developing technology, & and gaining wide spread popularity, the proposed technique have to adopt & improve to meet future challenges & expectations. One of the key security improvements sought is use of hybrid techniques such as cryptography combined with randomization. Another important adaptation is reduction of space & time complexity of PPDM techniques to effective manage cloud services disk space & power resources. Also the proposed system should be scalable to large volumes of data so as to cater to real life data mining challenge.

REFERENCES

1. Fuzzy Logic and Hybrid Approaches to Web Intelligence Gathering and Information Management. Riza C. Berkan and Sheldon L. Trubatch Answer Chase, Inc. 34 Defense Street, Annapolis, 21401 Maryland, USA
2. K-Nearest Neighbor Classification over Semantically Secure Encrypted Relational Data K Sumalatha Y AppaRao, International Journal of Engineering Research in Computer Science and Engineering (IJERCSE) Vol 3, Issue 3, March 2016 All Rights Reserved © 2016 IJERCSE 485 Prasad B II/IV, Associate Professor Department of CSE, Marri Laxman Reddy Institute of Technology and Management (MLRITM) Hyderabad
3. 4. Fuzzy Logic-based Implicit Authentication for Mobile Access Control SAI Feng Yao, Suleiman Y. Yerima, BooJoong Kang, Sakir Sezer Computing Conference 2016 July 13-15, 2016 | London, UK Centre for Secure Information Technologies, Queen's University Belfast Belfast, Northern Ireland, UK fyao02@qub.ac.uk, s.yerima@qub.ac.uk, b.kang@qub.ac.uk, s.sezer@ecit.qub.ac.uk
4. A New Intrusion detection system in data mining & fuzzy logic Balaji.s.c.k, B.Kishore Kumar International Journal of Modern Engineering Research (IJMER) www.ijmer.com Vol. 3, Issue. 6, Nov - Dec. 2013 pp-3425-3428 Assistant professor, SVIT, Hampapuram, Anantapuram (Dt), Andhra Pradesh
5. k -Nearest Neighbor Classification over Semantically Secure Encrypted Relational Data Surya S.R. Swathy J. International Journal of Computer Applications (0975 - 8887) Volume 133 - No.8, January 2016 Review on P G Scholar Department of Computer Science & Engineering College of Engineering, Perumon(CUSAT), Kerala, India Assistant Professor in CSE Department of Computer Science & Engineering College of Engineering, Perumon(CUSAT), Kerala, India
6. K-Nearest Neighbor Classification over Semantically Secure Encrypted Relational Data. Chetan Wankhede, Nikhil Daundka, Rohan Wadmare, Tushar Hinge, Prof. M.A.Ansari Student, Dept. of IT, Rajarshi Shahu College of Engineering, Tathawade, Savitribai Phule Pune University, Pune, India International Journal of Innovative Research in Science, Engineering and Technology (An ISO 3297: 2007 Certified Organization) Vol. 5, Issue 3, March 2016 Copyright to IJIRSET DOI:10.15680/IJIRSET.2016.0503197 3587
7. K-NN Classification Over Secure Encrypted Relational Data In Outsourced Environment Multidisciplinary Akshay Dabi, Arslan Shaikh, Pranay Bamane, Vivek Thorat,Journal of Research in Engineering and Technology, Volume 2, Issue 4, Pg 783-787 M38-2-4-10-2015 Prof. Popat Borse. Computer Engineering. Dr. D.Y Patil School of Engineering Pune, India
8. Original scientific paper A grey-fuzzy approach for optimizing machining parameters and the approach angle in turning Advances in Production Engineering & Management Senthil kumar, N.a, Sudha, J.b, Muthukumar, V. Adhiparasakthi ISSN 1854-6250 Volume 10 | Number 4 | December 2015 | pp 195–208 Journal home: apem-journal.org http://dx.doi.org/10.14743/apem2015.4.202 AISI 1045 steel Engineering College, Melmaruvathur, Tamil Nadu, India College of Engineering, Anna University, Chennai, Tamil Nadu, India Saveetha Engineering College, Chennai, Tamil Nadu, India
9. K-NN Classification over Semantically Secure Encrypted Relational Data C.Vanathy, Ramyaseruba 1,2 Dept of CSE Sree Sastha Institute of Engg & Tech, Chennai
10. A Survey on Intrusion Detection Techniques in MANETs by Shona D, Dr. M. Senthil Kumar Volume 3, Issue 10, October 2015 International Journal of Advance Research in Computer Science and Management Studies Research Article / Survey Paper / Case Study Available online at: www.ijarcsms.com