ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 12, December 2015

Survey on Classification for Semantically Secured Encrypted of Data

Ashwini Sarode, Saudagar Barde

Student, Dept. of Computer Engineering, SKN College of Engineering, Savitribai Phule Pune University,Pune, India

Asst. Professor, Dept. of Computer Engineering, SKN College of Engineering, Savitribai Phule Pune University, Pune, India

ABSTRACT: Data mining is the analysis step of the "Knowledge Discovery in database". Data Mining widely used in many fields such as medical research, financial, medication and among govt. departments. For last few years, query processing on relational data has been studied extensively, and many theoretical and practical solutions to query processing have been proposed under various scenarios. Classification is one of the widely applied works in data mining applications. As increasing popularity of cloud computing, users now able to outsource their information as well as the data management tasks to the cloud.
This paper survey various k-nearest neighbor (kNN) query problem over encrypted database outsourced to a cloud: a user issues an encrypted query record to the cloud and the cloud returns the k closest records to the user. This paper is motivated by need of secure classification technique to outsourced data to the cloud.

KEYWORDS: Security, k-NN classifier, Encryption, Cloud Computing.

I. INTRODUCTION

As increasing popularity of cloud computing, users now able to outsource their information as well as the data management tasks to the cloud. Cloud computing has recently emerged as a new platform for deploying, managing, and provisioning large-scale services through an Internet-based infrastructure such as Amazon EC2, Google App Engine, and Microsoft Azure. Last few decades the cloud computing model is changing the organizations way of working their information. Almost all organizations assign their computational functions to the cloud to improve their information.

Regardless of incredible advantages that the cloud environment offers, privacy and security issues in the cloud are thwarting companies to utilize those advantages. When data are private or highly sensitive, the data need to be encrypted before outsourcing to the cloud. However, when data are encrypted, performing any data mining tasks becomes very challenging without ever decrypting the data.

In this paper, survey various problem of secure processing of k-nearest neighbor query over encrypted data (SkNN) in the cloud.

II. LITERATURE SURVEY

P. Williams, R. Sion, and B. Carbunar [2] author provide novel and effective practical scheme with efficient access pattern privacy for remote data storage with correctness. Major aim of invention of this protocols to yield practical computational complexity (to O(log n log log n)) and storage overheads (to O(n)).

Storage client issue encrypted reads, writes without revealing information or access patterns by using proposed mechanism. Proposed scheme is faster than existing system which can execute several queries per second and also offering privacy as well as correctness.

Pascal Paillier [3], this paper identifies the important computational problem called Composite Residuosity Class Problem. To handle this problem author proposed a new trapdoor mechanism and developed three encryption scheme such as a trapdoor permutation and two homomorphic probabilistic encryption schemes.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 12, December 2015

In this paper they introduced a new number-theoretic problem and a related trapdoor mechanism using the concept of composite degree residue. Three new cryptosystems are obtained based on proposed technique, which are providing secure result under the assumption of intractability.

R. Dey, C. Tang, K. Ross, and N. Saxena [4], paper mainly focus on solving the problem of encrypted data classification. Paper proposed a novel PPkNN protocol, a secure k-NN classifier over semantically secure encrypted data. Author proposed a secure k-NN classifier for encrypted data in the cloud. Commonly used scheme in data mining is classification which is used in health-care and business.. The proposed KNN (k Nearest Neighbor) protocol provides protection for the users input query, confidentiality of the data and data access patterns. Efficiency of proposed method gives better result.

In proposed protocol once the encrypted data are handover to the cloud, Alice does not participate in any computations. Thus, no information is revealed to Alice which indirectly achieves privacy.

C. Gentry [5], author proposed a fully homomorphic encryption scheme based on concept of lattice. Proposed work consists of three steps; initial step provide a general result of encryption scheme which permits for evaluation of arbitrary circuit, then a public key encryption scheme based on concept of lattices and final step is to reduce the depth of decryption circuit and used to produce a bootstrappable encryption scheme.

Proposed scheme solve the DMED problem since it allows a third-party to execute arbitrary functions over encrypted data without ever decrypting them. This technique is very expensive and their usage in practical applications has yet to be explored.

D. Bogdanov, S. Laur, and J. Willemson [6], author proposed a novel approach for developing privacy-preserving applications, namely SHAREMIND. SHAREMIND is based on share computing techniques. SHAREMIND is basically a virtual machine for privacy-preserving data processing.
Performance is increased by using proposed method when compared to other similar frameworks. Application development interface developed by SHAREMIND is easy which mainly focus on implementation of data mining algorithm not on privacy issues.

The SHAREMIND framework is considered to be a capable and easily programmable platform for developing and testing numerous privacy-preserving algorithms. By using SHAREMIND anyone can develop secure multi-party protocols without the prior knowledge of all implementation details.

Agrawal and Srikant, Lindell and Pinkas [7], introduced the idea of privacy-preserving under data mining applications. Objective of paper is to build a classifier in order to predict the class label of input data record based on the distributed training dataset without compromising the privacy of data. Author introduced first data perturbation technique to build a decision-tree classifier
To accurately calculate the distribution of original data values author proposed a novel reconstruction procedure. By using these reconstructed distributions, we are able to build classifiers whose accuracy is comparable to the accuracy of classifiers built with the original data.

H. Hu, J. Xu, C. Ren, and B. Choi [8], studied the problem of processing private queries on indexed data for mutual privacy protection in a cloud environment. Author proposed an encryption scheme based on privacy homomorphism and an efficient solution that comprises a secure traversal framework. The proposed framework is scalable to large datasets by using an index-based approach.

Secure protocols based on this framework is invented for processing typical queries such as k-nearest-neighbor queries (kNN) on R-tree index. Proposed framework is scalable to large datasets. This approach shows various advantages such as feasible, efficient and robust.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 12, December 2015

III. RELATED WORK

Table 1: Survey Table

Sr.no	Paper	Proposed	Advantage
1	Building Castles out of Mud: Practical Access Pattern Privacy and Correctness on Untrusted Storage [2]	Author provide novel and effective practical scheme with efficient access pattern privacy for remote data storage with correctness.	Proposed mechanism is faster and also offering privacy as well as correctness.
2	k-Nearest NeighborClassification overSemantically Secure Encrypted Relational Data [4]	Author proposed a secure k-NN classifier for encrypted data in the cloud.	The proposed KNN (k Nearest Neighbor) protocol provides protection for the users input query, confidentiality of the data and data access patterns.
3	Fully HomomorphicEncryption Using Ideal Lattices [5]	Proposed scheme solve the DMED problem since it allows a third-party to execute arbitrary functions over encrypted data without ever decrypting them.	Decryption overhead is reduced.
4.	Sharemind: a framework for fast privacy-preservingComputations [6]	Author proposed a novel approach for developing privacy-preserving applications, namely SHAREMIND.	The SHAREMIND framework is considered to be a capable and easily programmable platform for developing and testing numerous privacy-preserving algorithms. . By using SHAREMIND anyone can develop secure multi-party protocols without the prior knowledge of all implementation details

IV. ARCHITECTURAL VIEW

Proposed system consists of three modules, such as user, dataset provider and third cloud. Here first query is fired by client in encrypted form is provided to cloud. Then cloud provides k closest records to the client in encrypted form. Dataset provider provides data to the cloud.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 12, December 2015

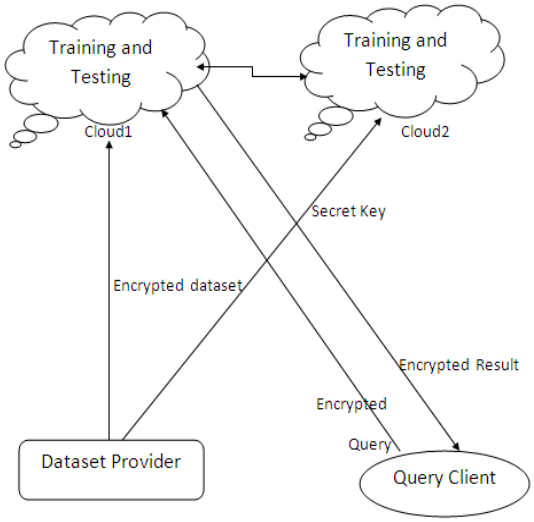

Fig 1. Architectural View

V.CONCLUSIONS

This paper presented an all-inclusive survey of classification data mining that are used to classify the data and a various techniques to protect the information leakage. The main features, the advantages and disadvantages of each recommendation algorithm are described. Classification has features to increase the reach and benefits of data mining technology.As per survey, a strong need to develop secure classification technique to outsourced data to the cloud.

REFERENCES

[1] Bharath K. Samanthula, Member, IEEE, YousefElmehdwi, and Wei Jiang, Member, IEEE, "k-Nearest Neighbor Classification over Semantically Secure Encrypted Relational Data", ieee transaction on knowledge and data engineering, vol.27, no. 5, may 2015.
[2] P. Williams, R. Sion, and B. Carbunar, "Building castles out of mud: Practical access pattern privacy and correctness on untrusted storage," in Proc. 15th ACM Conf. Comput. Commun. Security, 2008, pp. 139–148..
[3] P. Paillier, "Public key cryptosystems based on composite degree residuosity classes," in Proc. 17th Int. Conf. Theory Appl. Cryptographic Techn., 1999, pp. 223–238.
[4] B. K. Samanthula, Y. Elmehdwi, and W. Jiang, "k-nearest neighbor classification over semantically secure encrypted relational data," eprint arXiv:1403.5001, 2014.
[5] C. Gentry, "Fully homomorphic encryption using ideal lattices," in Proc. 41st Annu. ACM Sympos. Theory Comput., 2009, pp. 169– 178.
[6] D. Bogdanov, S. Laur, and J. Willemson, "Sharemind: A framework for fast privacy-preserving computations," in Proc. 13th Eur. Symp. Res. Comput. Security: Comput. Security, 2008, pp. 192–206.
[7] R. Agrawal and R. Srikant, "Privacy-preserving data mining," ACM Sigmod Rec., vol. 29, pp. 439–450, 2000.
[8] H. Hu, J. Xu, C. Ren, and B. Choi, "Processing private queries over untrusted data cloud through privacy homomorphism," in Proc. IEEE 27th Int. Conf. Data Eng., 2011, pp. 601–612.
[9] P. Mell and T. Grance, "The NIST definition of cloud computing (draft)," NIST Special Publication, vol. 800, p. 145, 2011.
[10] Y. Lindell and B. Pinkas, "Privacy preserving data mining," in Proc. 20th Annu. Int. Cryptol. Conf. Adv. Cryptol., 2000, pp. 36–54.