Volume 5, Issue 12, December 2015 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

RASP Data Perturbation Method to Provide Secure and Efficient Query Services in the Cloud

Puja Bhaganagarkar, Prof. Biswa Ranjan Dash

Department of Computer Science, D.Y.Patil College of Engineering,
Ambi, Pune, Maharashtra,
India

Abstract- With the wide readying of public cloud computing infrastructures, victimization clouds to host informationquestion services has become associate appealing resolution for the benefits on quantifiability and cost-saving. we tend to propose the random house perturbation (RASP) information perturbation methodologyto supply secure and economicalvaryquestion and kNNquestion services for protected informationwithin the cloud. The RASP information perturbation methodology combines order conservingencoding, spatial propertyenlargement, random noise injection, and random projection, to supplysturdy resilience to attacks on the flusteredinformation and queries. It additionally preserves three-d ranges, thatpermits existing classification techniques to be applied to speedvaryquestionprocess. The kNN-R formulais meantto figure with the RASP varyquestionformula to method the kNN queries. we'verigorouslyanalyzed the attacks on information and queries underneath a exactlyoutlined threat model and realistic security assumptions. intensive experiments are conducted to indicatethe benefits of this approach on potency and security.

Keywords- Privacy, kNN-R, RASP Algorithm.

I. INTRODUCTION

The extensive exploitation of cloud infrastructures, has made it possible to host services and big data in public clouds. This new paradigm is especially attractive for data intensive query and analysis services for its great scalability and significant cost savings. It is well known that maintaining and mining data incurs much higher cost than initial data acquisition. However, some data might be sensitive that the data owner does not want to move to the cloud unless the data confidentiality and query privacy are guaranteed. On the other hand, a secured query service should still provide efficient query processing and significantly reduce the in-house workload to fully realize the benefits of cloud computing.By moving data services to the cloud, data owners can cut costs in almost every aspect of managing and mining data. However, data privacy is still haunting data owners' minds as the underlying infrastructure is out of their control. In particular, data owners may not be aware of information leakage, which can happen in all kinds of possibilities, if the cloud provider does not want to report the leakage.

A straightforward method is to encrypt datasets before exporting them to the cloud. However, searchable encryption is very challenging, showing limited successes in some specific areas such as document search [4]. Boneh et al. [2] showed that it is possible to construct a public-key system for range query, which is one of the basic database queries.However, it requires a significant amount of storage and computational costs, only applicable to linear scan of the entire database. Database queries such as range and kNN queries normally demand fast processing time (logarithmic or sublinear time complexity) with the support of indexing structures. However, if not impossible, there is no efficient indexing structure developed for encrypted data yet, which renders the current encryption schemes [2] unusable for search in large databases. We recently proposed the Random Space Perturbation(RASP) method [5] for the protection of tabular data, which is secure under the assumption of limited adversarial knowledge - only the perturbed data and the data distributions are known by adversaries. This assumption is appropriate in the context of cloud computing. The RASP perturbation is a unique combination of Order Preserving Encryption (OPE) [1], dimensionality expansion, noise injection, and random projection, which provides sufficient protection for the privacy of query services in the cloud. It has a number of unique features, such as preserving the topology of range query, non-deterministic results for duplicate records, and resilience to distributional attacks [5].

We develop the secure half-space query transformation method that casts any enclosed range in the original space to an irregularly shaped range in the perturbed space. Therefore, we are able to use a two-stage range query processing method : an existing multidimensional index, such as R*-Tree in the perturbed space is used to find out the records in the bounding box of the irregularly shaped range, which is then filtered with the transformed query condition. This processing strategy is fast and secure under the security assumption. To allow the readers to fully appreciate the intuition and the ideas behind the RASP based perturbation and query processing, we propose this RASP Query Services (RASPQS) demonstration system. This system consists of the following major components: (1) the user interface for perturbation parameter generation that allows users to observe the details of RASP perturbation, (2) the visualization of the two-stage range query processing procedure to understand the transformed query ranges and the query results, (3) the visualization of the progressive steps in the kNN query processing that is based on RASP range query processing, and (4) the performance comparison on index-aided processing on non-encrypted data, linear-scan query processing on encrypted data [2], and the RASP query processing.

II. LITRATURE REVIEW

Using clouds to host data query services has become an appealing solution for the advantages on scalability and cost-saving. However, some data might be sensitive that the data owner does not want to move to the cloud unless the data confidentiality and query privacy are guaranteed. In addition, a secured query service should still provide efficient query processing and significantly reduce the in-house workload for the purpose of cloud computing. Bearing these criteria in mind, we propose the RASP data perturbation method to provide secure range query and kNN query services for protected data in the cloud. The RASP data perturbation method combines order preserving encryption, dimensionality expansion, random noise injection, and random projection, to provide strong resilience to attacks on the perturbed data and queries. It also preserves multidimensional ranges, which allows existing multidimensional indexing techniques to be applied in range query processing. The kNN-R algorithm is designed to work with the RASP range query algorithm to process the kNN queries. We carefully analyze the attacks on data and queries under a precisely defined threat model and realistic assumptions. Extensive experiments have been conducted to show the advantages of this approach on the balance of performance and security. Secure data intensive computing in the cloud is challenging, involving a complicated tradeoff among security, performance, extra costs, and cloud economics. Although fully homomorphic encryption is considered as the ultimate solution, it is still too expensive to be practical at the current stage. In contrast, methods that preserve special types of data utility, even with weaker security, might be acceptable in practice. The recently proposed RASP perturbationmethod falls into this category. It can provide practical solution for specific problems such as secure range queries, statistical analysis, and machine learning. The RASP perturbation embeds the multidimensional data into a secret higher dimensional space, enhanced with random noise addition to protect the confidentiality of data. It also provides a query perturbation method to transform half-space queries to a quadratic form and, meanwhile, preserving the results of half-space queries. The utility preserving property and wide application domains are appealing. However, since the security of this method is not thoroughly analyzed, the risk of using this method is unknown. The purpose of this paper is to investigate the security of the RASP perturbation method based on a specific threat model. The threat model defines three levels of adversarialpower and the concerned attacks. We show that although the RASP perturbed data and queries are secure on the lowest level of adversarial power, they do not satisfy the strong indistinguishability definition on higher levels of adversarial power. As we have noticed, the indistinguishability definition might not be too strong to be useful in the context of data intensive cloud computation. In addition, the noise component in the perturbation renders it impossible to exactly recover the plain data; thus, all attacks are essentially estimation attacks. We propose a weaker security definition based on information theoretic measures to describe the effectiveness of estimation attacks, and then study the security under this weaker definition. This security analysis helps clearly identify the security weaknesses of the RASP perturbation and quantify the expectedmaintenance of confidentiality under different levels of adversarial power.

III. CONCLUSION

The proposed system is mainly used to perturb the data given by the owner and saved in cloud storage it also combines random injection, order preserving encryption and random noise projection. By using the range query and kNN query user can retrieve their data's in secured manner and the processing time of the query is minimized.Thus our algorithm focuses to further improve the performance of query processing for both range queries and kNN queries and formally analyze the leaked query and access patterns and the possible effect on both data and query confidentiality.

REFERENCES

[1] Agrawal, R., Kiernan, J., Srikant, R., and Xu, Y. Order preserving encryption for numeric data. In Proceedings of ACM SIGMOD Conference (2004).

[2] Boneh, D., and Waters, B. Conjunctive, subset, and range queries on encrypted data. In the Theory of Cryptography Conference (TCC (2007), Springer,pp. 535–554.

[3] Chen, K., and Liu, L. VISTA: Validating and refining clusters via visualization. Information Visualization 3,4 (2004), 257–270.

[4] Curtmola, R., Garay, J., Kamara, S., and Ostrovsky, R. Searchable symmetric encryption: improved definitions and efficient constructions. In ACM CCS (2006), pp. 79–88.

[5] Xu, H., Guo, S., and Chen, K. Building confidential and efficient query services in the cloud with rasp data perturbation. IEEE Transactions on Knowledge and Data Engineering 26, 2 (2014)

[6] J. Bau and J. C. Mitchell, "Security modeling and analysis," IEEE Security and Privacy, vol. 9, no. 3, pp. 18–25,2011.

[7] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, “Privacy-preserving multi-keyword ranked search over encrypted cloud data,”.

[8] K. Chen, R. Kavuluru, and S. Guo, “Rasp: Efficient multidimensional range query on attack-resilient encrypted databases,” in ACM Conference on Data and Application Security and Privacy, 2011, pp. 249–260.

[9] K. Chen and L. Liu, “Geometric data perturbation for outsourced data mining,” Knowledge and Information Systems, 2011.

[10] M. L. Liu, G. Ghinita, C. S.Jensen, and P. Kalnis, “Enabling search services on outsourced private spatial data,” The International Journal of on Very Large Data Base, vol. 19, no. 3, 2010.