International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

Confidential and Efficient Hosting Query Services in Public Clouds with RASP Data Disruption

Vanajakshi Devi .K[1], Praveen Kumar .N[2]

[1]Yogananda Institute of Technology & Science, Department of CSE,, Mohan Reddy Nagar, Tirupati, India

[2]Yogananda Institute of Technology & Science, JNTUA, Mohan Reddy Nagar, Tiirupati, India

Abstract: ***As cloud computing offers public sharing services, the data queries and information can be exchanged efficiently in order to enhance scalability and cost-saving. But few concerns with data owners are security and query privacy. A novel technique to overcome this is using RASP data perturbation method to provide secure and efficient range query and kNN query services for protected data in the cloud. This approach provides the private guarantee protocol to cast internet based computing, when compared to the encryption-based approach by enabling much faster query processing. This method combines dimensionality expansion, order preserving encryption, random noise injection, and random projection, to provide strong resilience to attacks on the perturbed data and queries. The kNN-queries are processed by kNN algorithm that works on contrary with RASP data query algorithm. This explains that this will allow users to more intuitively understand the technical advantages of the RASP approach via interactive exploration of the visual interface. These approach analyses the attacks on queries and data with an exactly defined threat model and realistic security assumptions. Highly sophisticated experiments have been conducted to show advantages of the approach.***

Keywords: RASP, disruption, perturbation, kNN algorithm, confidential, CPEL criteria

1. Introduction

With cloud infrastructures, the users can use the services provided by cloud upto certain extent as pay as you go model. But the data services for the owner are still at stake as preserving their data is not simple as possible as they don't have any idea about the underlying architecture. The data owners are not aware that their information is leaked by intrusion which has higher possibility in these kind of conditions. On the other side of the coin, a secured and reliable query service should still provide efficient query processing and significantly reduce the in-house workload to fully realize the benefits of cloud computing. This will be a highly sophisticated and much needed feature because these workloads of query services are highly dynamic, and it will be expensive and inefficient to serve such dynamic workloads with in-house infrastructures. This technique is more famous by maintaining and mining data incurs much higher cost than initial data acquisition.

The significant requirements for constructing a practical query service in the cloud as the CPEL criteria: Efficient query processing, query Privacy, data Confidentiality and Low in-house processing cost. A candid approach for this would be using an encryption for the files before uploading into the cloud. The Random Space Perturbation 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. The RASP disruption is a unique combination of OPE, dimensionality enhancement, random noise injection and also random projection, which provides strong and high confidentiality and privacy guaranteed. The research of secure half-space query transformation method which makes any enclosed range in the original space to an irregularly shaped range in the disrupted space.

RASP Architecture

The confidential and efficient query services building in public hosts is an immense task, which has to coordinate with the data owner and authenticated servers in order to provide it with a secure data through RASP data disruption.

The architecture is partitioned into trusted and untrusted parties. The trusted parties provides schemes to submit queries which includes file owners, in-house proxy servers and authorized servers. The untrusted party is someone who protect the database and cloud provider who submits the queries. RASP perturbation is a novel concept of order preserving encryption (OPE) [1], and many other injection and projections of query space.

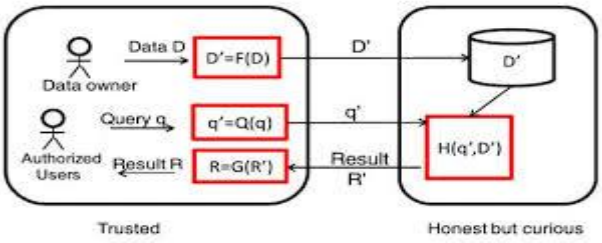

Figure 1: Trusted and Untrusted parties of RASP based query services

1) The details of RASP perturbation can be allowed to observe by the user with perturbation parameter user interface.
2) By indexing and efficient query processing, this approach ensures that the topology of multidimensional range is not altered in secure transformation.
3) The visualization of the two-stage range query processing procedure to understand the transformed query ranges and the query results.
4) The performance of comparison with index-aided processing on encrypted data and RASP query processing.

2. kNN Algorithm

The RASP perturbation is not compatible with the kNN algorithm due to distance orders. Hence, a kNN algorithm is introduced in order to process the queries that uses indexing mechanism and enables faster processing.
Here we use indexes along with inner range expansion which can be given by Binary range search.

1) The user sets the initial outer square range with a certain distance from query point.
2) The algorithm now finds the middle range as it knows inner and outer range in each iteration, in which no. of enclosed points is greater than k, the outer range is replaced by middle; otherwise, the inner range is replaced by outer.
3) The records which are filtered are sent back to client for final kNN algorithm.

This algorithm helps to do the search and find the results quickly. Even experiments show that this algorithm is efficient.

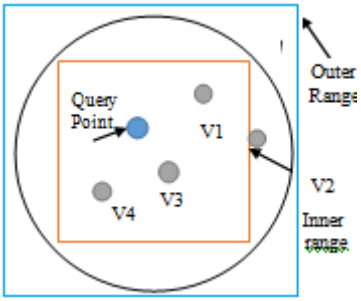

Figure 2: Illustration for kNN-algorithm, where k=3

3. Demonstration and Results

Here we demonstrate unique features of RASP-QS demonstration algorithm and RASP query processing. We also discuss RASP data disruption expense, resilience of OPE enhanced RASP is to the ICA-based attack, efficiency of two stage range query processing and advantages of kNN algorithm.

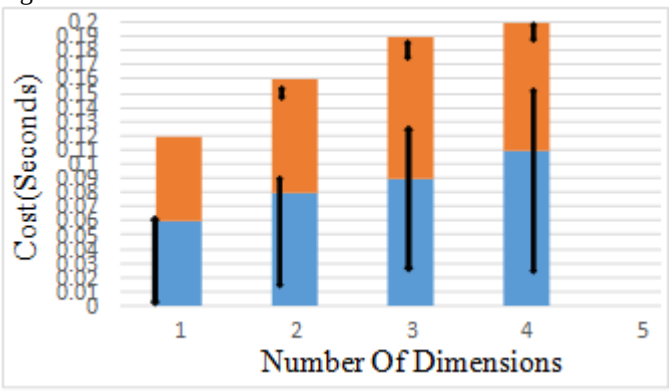

Figure 3: Cost distribution of full RASP scheme

We demonstrate the costs of RASP perturbation here. The costs are of two types: OPE and the rest of RASP. The OPE algorithm distributes the target into buckets and sorted values are aligned proportionally according to the target bucket distribution. The above figure distributes the cost estimation of RASP of 20k records at different no. of dimensions.

4. Conclusion

The RASP perturbation gives the data owner an additional security and privacy from data intruders, information leakage and other security issues. It satisfies the hosting query services in cloud, along with CPEL criteria. This paradigm is the key for the efficiency and scalability of the hosting query services. It is a unique composition of OPE, random noise injection, dimensionality expansion, and random projection, which provides unique security features. This approach will be highly interactive and visual, allowing the users to easily understand the technical details and appreciate the advantages of this demonstrated scheme. The main motive of RASP is to preserve the topology of the queried range in the disrupted space, and allows to use indices for efficient range query processing. To achieve the sub linear time complexity of queries, we introduce kNN algorithm query service based on the range query service. The security and privacy aspects of both the perturbed data and the protected queries is analyzed carefully under a precisely defined threat model. Thus RASP approach provides public hosting of files reliable and strong base towards practical confidential query services. Several experiments are also being conducted to show the efficiency of query services and low cost of in-house processing.

References

[1] 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 Engineering26, 2 (2014).
[2] Curtmola, R., Garay, J., Kamara, S., and Ostrovsky, R. Searchables symmetric encryption: improved definitions and efficient constructions. In ACM CCS (2006), pp. 79–88.
[3] 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.
[4] B. Chor, E. Kushilevitz, O. Goldreich, and M. Sudan, “Private information retrieval,” ACM Computer Survey, vol. 45, no. 6, pp. 965–981, 1998.
[5] B. Hore, S. Mehrotra, and G. Tsudik, “A privacy-preserving index for range queries,” in Proceedings of Very Large Databases Conference (VLDB), 2004.
[6] M. Rudelson and R. Vershynin, “Smallest singular value of a random rectangular matrix,” Communications on Pure and Applied Mathematics, vol.62, pp. 1707–1739, 2009.
[7] H. Hu, J. Xu, C. Ren, and B. Choi, “Processing private queries over untrusted data cloud through privacy homomorphism, “Proceedings of IEEE International Conference on Data Engineering (ICDE), pp. 601–612, 2011.
[8] Vanajakshi Devi.K, Praveen Kumar.N, Ramesh. B, Cost-Efficiency and privacy preserving with EIRQ methods in commercial cloud, IJECS vol 4 issue 3 March 2015

Author Profile

K. Vanajakshi Devi, working as assistant professor in yogananda institute of Tech. & sci. Tirupathi, she has memberships in MISTE and IAENG also serves as academic coordinator and also published journals and attended conferences

N. Praveen Kumar, pursuing Bachelor's Degree in Yogananda Institute of Technology & Science, Tirupati. He also has attended several National & International conferences. He got first place in National Symposium conducted by SVNE, Tirupati.