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

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Vol. 4, Issue 1, January 2016

Data Attribute Security Using Access Control Mechanism for Distributed Databases

Pradnya Gangavne, Sonal Zinjurte, Rupali Pardeshi

Dept. of Computer Engineering, Dr.D.Y.Patil College of Engineering, Pimpri, University of Pune, India

ABSTRACT- Access control mechanisms protect liable information from unauthorized users. However, when liable information is participate and a Privacy Protection Mechanism (PPM) is not in place, an authorized user can still compromise the privacy of a person leading to sameness disclosure. A PPM can use suppression and generalization of relational data to anonymize and satisfy privacy condition e.g. K-anonymity and l-diversity ,against identity and attribute disclosure. However, privacy is achieved at the cost of accuracy of authorized information. This propose an accuracy-constrained privacy-preserving access control architecture. The access control policies define choice predicates available to roles while the privacy condition is to a satisfy the K-anonymity or l-diversity. An additional constraint that needs to be satisfied by the PPM is the approximate bound for each choice predicate. The techniques for workload-aware anonymization for choice predicates have been discussed in the literature. However, to the best of our knowledge, the problem of satisfying the accuracy constraints for more roles has not been studied before. In our formulation of the aforementioned problem , we propose heuristics for anonymization algorithm and show empirically that the proposed approach satisfies approximate bounds for more permissions and has lower total approximate than the current state of the art**.**

KEYWORDS: Privacy preservation techniques (PPM),Trusted third, party(TTP) protocol,m-privacy techniques, Slicing Algorithm, Sql injection attacks.

I. INTRODUCTION

Privacy preserving publishing of micro data is one of the most important concern in collaborative data publishing. There is a requirement of data characteristic security in disseminated database while preserving solitude. In the proposed work, we infer problem connected in publishing mutual data for anonymizing perpendicularly and parallel partition data. In the proposed system, a hospital collects data from patients and publishes patient records to an external entity. The privacy of data is preserved while sharing of information among hospitals and other providers.

This system is designed to preserve the privacy of an special data in distributed database system by using the slicing algorithm. It helps to improve the data security when data is gathered from different sources and output should be in collaborative fashion. It is used for publishing data in a hostile environment so that the published data remain practically useful while special privacy is preserved.

Privacy preservation techniques are mainly used to reduce the leakage of formation about the particular special while the data are shared and released to public. For this, the sensitive information should not reveal. Data is getting modified first and then published for further process. For this various anonymization techniques are succeed and they are generalization, suppression, permutation and perturbation. By various anonymization techniques data is modified which retains enough utility and that can be released to other parties safely. Organizations need to participate data for mutual benefits or for publishing to a third party. For banking sector want to integrate their customer data for developing a system to deliver better services for its customers. However, the banks do not want to indiscriminately reveal their data to each other for reasons such as privacy protection and business competitiveness.

Main aim is to publish an anonymized view of integrated data, T, which will be immune to attacks (Figure 1.1). Attacker runs the attack, i.e. a single or a group of accidental or internal entities that wants to breach privacy of data using background wisdom.

Collaborative data publishing is carried out successfully with the help of trusted third party (TTP) or a Secure Multi Party Computation (SMC) protocol, which guarantees that information or data about particular special is not reveal anywhere that means it maintains privacy. A more desirable approach for collaborative data publishing is first aggregate then anonymize.

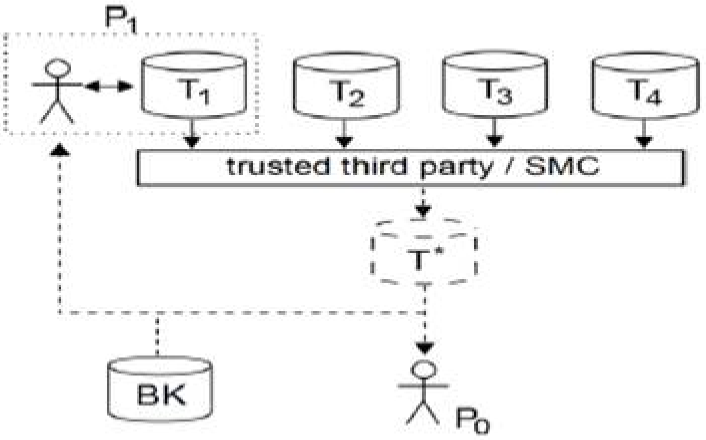

Figure 1: System Module

In above figure, T1, T2, T3 and T4 are databases for which data is provided by provider like provider P1 provides data for database T1. These distributed data coming from dissimilar providers get aggregate by TTP (trusted third party) or using SMC protocol. Then these summation data anonymized further by any anonymization technique. P0 is the authenticate user and P1 trying to infraction privacy of data which is provided by other users with the help of BK (Background knowledge). This type of attack we can call as an “insider attack”. We have to protect our system from such a type of attacks.

Slicing algorithm with which we again used encrypted data which improves precautions. Slicing is the procedure which gives improved result than characteristic simplification and bucketization method. It gives better results for high dimensional data. It can perform permutation within bucket. In slicing we can pool resources sensitive attribute with some quasi identifier. On this sliced data we utilize a confirmation algorithm which verifies that whether information is secured or not.

II. RELATED WORK

We first ceremonial describe our problem setting. Then, we present our data-privacy definition with respect to a privacy constraint to prevent consequence attacks by data-adversary, followed by properties of this new privacy notion. Let T = t1, t2, . . .be a set of report with the identical attributes gathered from n data providers P =P1, P2, . . . , Pn, such that Ti are report provided by Pi. Let AS be a delicate attribute with a domain DS. If thep report contain multiple delicate attributes then, we treat each of them as the sole delicate attribute, while remaining ones we include to the quasi-identifier. However, for our scenarios we use an approach, which preserves more utility without sacrificing privacy. Our aim is to publish an anonymized T* while preventing any data-adversary from inferring AS for Any single report.

An data-adversary is a body of data users with n data providers cooperating to breach privacy of anonymized records. When data are accumulate and join from different data providers, mainly two things are done, for anonymization process. can be used as a component constraint Ci. To protect data from external recipients with certain background knowledge BK, we assume a given privacy condition C is defined as a conjunction of privacy constraints: C1C2...Cw. If a group of anonymized report T* satisfies C, we say C(T*)=true. By definition C() is true and is private. Any of the existing privacy principles .

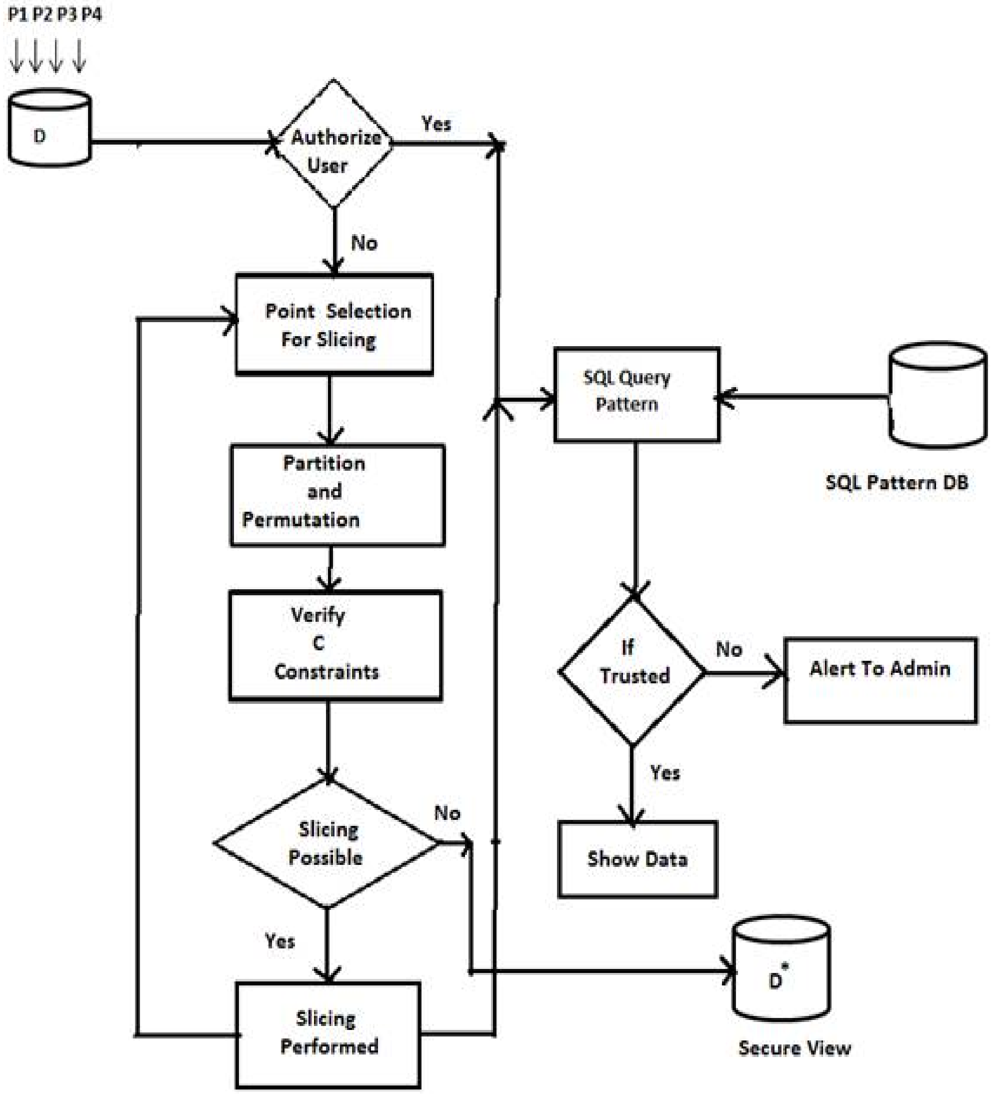

Figure 2: System Architecture

We now formally define a notion of data privacy with respect to a privacy constraint C, to protect anonymized data against data adversaries .The notion explicitly models the inherent data knowledge of an data-adversary, the data report they jointly contribute, and requires that each QI group, excluding any of those report owned by an data-adversary, still satisfies C. Figure 5.1.1 shows our proposed system in which input data is given in encrypted arrangement(attribute name will be in encrypted arrangement). Select point for slicing. Check that input data on privacy constraint C for data privacy. Check beyond is slicing is possible or not. If slicing possible then do it and if not then decrypt data . Our last output T* are anonymized data which will seen only by authenticate user. Any adversary cannot breach privacy of data. In this system we are using horizontal as well as vertical partitioning over database. Slicing algorithm deliver better column partitioning. To understand this properly lets consider hospital management system for test. Let different departments are the providers who provides data from different sources. We consider disease as a AS (sensitive attribute) and age and zip code are QI(quasi identifier).

III. LITERATURE SURVEY

The different authors are presenting the different methods which are previously used for anonymization. We discuss some advantages and limitations of these systems. Privacy preserving data analysis and collaborative data publishing has received considerable attention in current years as promising approaches for sharing data while preserving individual privacy.

3.1.Jing Yang and Ziyun Liu, et, al., A Data Anonymous Method based on Overlapping Slicing, IEEE 2014 International Conference on Computer Supported Cooperative Work in Design. The system also conduct several experiments to confirm that overlapping slicing technology ensures data security and improves the effectiveness of anonymous data at the same time. The overlapping slicing processes high-dimensional data effectively.

3.2.2013 IEEE International Conference on Emerging Trends in Computing, Communication and Nanotechnology ,' AN EFFICIENT TECHNIQUE FOR PREVENTING SQL INJECTION ATTACK USING PATTERN MATCHING ALGORITHM', (ICECCN 2013).

In this paper [2] , Pattern matching is a technique that can be used to identify or detect any anomaly packet from a sequential action. Injection attack is a method that can inject any kind of malicious string or anomaly string on the original string. Most of the pattern based techniques are used static analysis and patterns are generated from the attacked statements. In this paper, we proposed a detection and prevention technique for preventing SQL Injection Attack (SQLIA) using Aho–Corasick pattern matching algorithm. In this paper, we proposed an overview of the architecture. In the initial stage evaluation, we consider some sample of standard attack patterns and it shows that the proposed algorithm is works well against the SQL Injection Attack.

3.3. R. Mahesh and T. Meyyappan, Anonymization Technique through Record Elimination to Preserve Privacy of Published Data, IEEE 2013 International Conference on Pattern Recognition.

In this paper [3], the authors propose a new method to preserve the privacy of individuals' sensitive data from record and attribute linkage attacks. In the proposed method, privacy preservation is achieved through generalization of quasi identifier by setting range values and record limitation.

3.4. Tiancheng Li, Ninghui Li, Slicing: A New Approach for Privacy Preserving Data Publishing, IEEE 2012 Transactions on Knowledge and Data Engineering.

This paper [4] proposes a novel technique called slicing, which partitions the data both horizontally and vertically. This new approach is used to privacy-preserving micro data publishing. System shows that slicing preserves better data utility than generalization and can be used for membership disclosure protection. Another important advantage of slicing is that it can handle high-dimensional data. It show how slicing can be used for attribute disclosure protection and develop an efficient algorithm for computing the sliced data that obey the ℓ-diversity requirement.

3.5. Tristan Allard, Benjamin Nguyen, Safe Realization of the Generalization Privacy Mechanism, IEEE 2011 Ninth Annual International Conference on Privacy, Security and Trust.

This paper [5] focuses on the organization of the collection and anonymization phases at the data source (i.e., at each SPT) while compromising neither privacy nor data utility compared to a trusted central server approach.

Sr.No.	Reference Name	Work information utilize	Problems found	Publication year
1	"*m*-Privacy for Collaborative Data Publishing".	Propose secure multi-party computation protocols for collaborative data publishing with *m*-privacy.	It remains a question to model and address the data knowledge of data providers .	2013
2	preventing sql injection attack using pattern matching algorithm', (ICECCN 2013).	Pattern matching is a technique that can be used to identify or detect any anomaly packet from a sequential action.	Injection attack is a method that can inject any kind of malicious string or anomaly string on the original string.	2013
3	"Slicing: A New Approach for Privacy Preserving Data Publishing".	Slicing is used for attribute disclosure protection	As randomly generate the associations between column values of a bucket.	2012
4	"m-Privacy for joint data publishing".	Develop methods and tools for publishing data in a hostile environment so that the published data remain practically useful while individual privacy is preserved.	The authorized party who may also play the role of the adversary with the goal of inferring sensitive information from the data received.	2011

IV. ANALYSIS

The purpose of this project is to design, develop and evaluate a file sharing system that proposes a novel solution to the shared file security problem. In detail, this document will provide a general description of our project, including user requirements, product perspective, and overview of requirements, general constraints. In addition, it will also provide the specific requirements and functionality needed for this project - such as interface, functional requirements and performance requirements.

V. CONCLUSION

We consider a potential attack on collaborative data publishing. We have used the slicing algorithm for anonymization and L diversity and verified it for security and privacy by using binary algorithm for data privacy. Slicing algorithm is very useful when we are using high dimensional data. It divides data in both vertical and horizontal fashion. Due to encryption we can increase security. But the limitation is that there could be loss of data utility.

Above system can be used in many applications like hospital management system, industrial areas where we like to protect a sensitive data e.g. salary information of the employee. Pharmaceutical company where sensitive data may be a combination of ingredients of medicines, in banking sector where sensitive data is account number of customer, balance etc. It can be used in military area where data is gathered from different sources and need to secure that data from each other to maintain privacy. This proposed system help to improve the data privacy and security when data is gathered from different sources and output should be in collaborative fashion.

VI. FUTURE WORK

The pruning power of cube materialization algorithms. Further, and demonstrate the ability to surface interesting cube groups as part of the cube computation process.

In future this system can consider for data which are distributed in ad hoc grid computing. Also the system can be considered for the set We can implement the proposed architecture on hadoop base system with cube materialization and map reduce. Also we can identify a subset of holistic measures that are partially algebraic and propose the technique of value partitioning to make them easy to compute in parallel. Design algorithms that partition the cube lattice into batch areas to effectively exploit both the parallel processing power of Map-Reduce valued data.

REFERENCES

1.Jing Yang and Ziyun Liu, et, al., A Data Anonymous Method based on Overlapping Slicing, IEEE 2014 International Conference on Computer Supported Cooperative Work in Design.

2.2013 IEEE International Conference on Emerging Trends in Computing, Communication and Nano-technology ,' AN EFFICIENT TECHNIQUE FOR PREVENTING SQL INJECTION ATTACK USING PATTERN MATCHING ALGORITHM', (ICECCN 2013).

3. R. Mahesh and T. Meyyappan, Anonymization Technique through Record Elimination to Preserve Privacy of Published Data, IEEE 2013 International Conference on Pattern Recognition.

4.Tiancheng Li, Ninghui Li, Slicing: A New Approach for Privacy Preserving Data Publishing, IEEE 2012 Transactions on Knowledge and Data Engineering.

5.Tristan Allard, Benjamin Nguyen, Safe Realization of the Generalization Privacy Mechanism, IEEE 2011 Ninth Annual International Conference on Privacy, Security and Trust.