Privacy Preserving Utility Verification and Security of Data Published by Non-Interactive Differentially Private Mechanisms

Gouri Namdeo Kale, S. N. Kini

Abstract: ***Service providers have the skill to collect large amounts of user data. Sometimes, a set of providers may attempt to combine their data for particular data mining tasks. In this process, how to keep users' privacy is very critical. Many Users write an own novel, personal story and private dataset, each user need to preserve this data harmless on own site and with user internet amounts of user data. Sometimes, a set of providers may attempt to combine their data for particular data mining tasks. In this process, how to keep users' privacy is very critical. Many Users write an own novel, personal story and private dataset, each user need to preserve this data harmless on own site and with user internet facility user always search a digital publisher. This is the so-called privacy-preserving collaborative data publishing problem. In this paper, we deliberate the collaborative data publishing problem for anonymizing horizontally partitioned data at several data providers. Meanwhile most anonymization methods have bad impact on data utility. However, this task is non-trivial for the reason the utility measuring usually requires the aggregated raw data, which is not exposed to the data users due to privacy concerns. The paper addresses this new threat, and makes several contributions.***

Index Terms: Data Privacy, Data security. double level Encryption, Utility verification

I. INTRODUCTION

In day today life every user is associated with internet and web related movement, now each user want to transaction with internet and internet related things & data. Data writer write data such as novel, Story, article etc. and want to securely publish all this data on Internet publishing site. There are Number of sites provides the feature of a data publishing, but due to unethical activity privacy preserving is turn into issue on every close. The data privacy is major issue at publisher site. Publisher wants to create a faithful model among writer and reader. In this trust model some important point is considered and that are data security, data privacy, Data integrity, Service providers have the capability to gather large quantities of user data. Occasionally, a set of providers may try to collective their data for specific data mining tasks. For example, the hospitals may farm out their medical records to a research group for mining the scattering patterns of any disease. There are many privacy models and corresponding anonymization mechanisms have been projected in the Literature such as *k*-anonymity and differential privacy. *k*-anonymity and its alternatives protect privacy by generalizing the records such as records cannot be separate from another records. Differential privacy is a much more rough privacy model. On removal or addition of single record released data becomes insensitive. For this implementation, matching anonymization mechanisms adds noise to the published data. Apparently, all these *mechanisms of data atomization have serious effects on the utility of data*. causes, the users who have published data frequently strong demands to confirm the real utility of the anonymized data.

II. LITERATURE REVIEW

For a discussion of the guarantees provided by differential privacy and their limitations, see [Kasiviswanathan and Smith 2008; Kifer and Machanavajjhala 2011].As the theoretical foundations of differential privacy become better understood, there is momentum to prove privacy guarantees of real systems.

Several authors have recently proposed methods for reasoning about differential privacy on the basis of different languages and models of computation, e.g. SQL-like languages [McSherry 2009], higher-order functional languages [Reed and Pierce 2010], imperative languages [Chaudhuri et al. 2011], theMapReducemodel [Roy et al. 2010],and I/O automata [Tschantz et al. 2011]. The unifying basis of these approaches are two key results: The first is the observation that one can achieve privacy by perturbing the output of a deterministic program by a suitable amount of symmetrically distributed noise, giving rise to the so-called Laplacian [Dwork et al. 2006b] and Exponentialmechanisms [McSherry and Talwar 2007]. The second result is theorems that establish privacy bounds for the sequential and parallel composition offerentially private programs, see e.g. [McSherry 2009].

In combination, both results form the basis for creating and analyzing programs by composing differentially private building blocks. While approaches relying on composing building blocks apply to an interesting range of examples, they fall short of covering the expanding frontiers of differentially private mechanisms and algorithms. Examples that cannot be handled by previous approaches include mechanisms that aim for weaker guarantees, such as approximate differential privacy [Dwork et al. 2006a], or randomized algorithms that achieve differential privacy without using any standard mechanism [Gupta et al. 2010]. Dealing with such examples requires fine-grained reasoning about the complex mathematical and probabilistic computations that programs perform on private input data.

Revised Version Manuscript Received on July 19, 2017.

Gouri Namdeo Kale, Department of Computer Engineering, Jayawantrao Sawant College of Engineering, Hadapsar Pune-28, Savitribai Phule Pune University, Pune (Maharashtra) India.

Dr. S. N. Kini, Professor, Department of Computer Engineering, Jayawantrao Sawant College of Engineering, Hadapsar Pune-28, Savitribai Phule Pune University, Pune (Maharashtra) India.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

Such reasoning is particularly intricate and error-prone, and calls for principled approaches and tool support. In this article we present a novel framework for formal reasoning about a large class of quantitative confidentiality properties, including (approximate) differential privacy and probabilistic non-interference

III. PROPOSED WORK

A. Apriori Algorithm for Data Processing:-

For frequent item set mining and for learning association rules over transactional database Apriori algorithm is used. It continues by recognizing the frequent individual items in the database and covering them to large item sets as long as those item sets seem suitably regularly in the database. The frequent item sets defined by Apriori can be used to fix association rules which focus general developments in the database.

B. RSA Algorithm for Two Level Encryption and Decryption:-

Modern computer uses RSA algorithm to encrypt and decrypt messages. RSA is an asymmetric cryptographic algorithm. Asymmetric means that there are two different keys. This is also known as public key cryptography one key is shared to everyone and one key kept private. RSA contains two keys public key and private key. The public key is used to encrypt messages. Encrypted Messages can be decrypted by using the private key.

IV. MATHS

Encryption:-

Sender A goes for following steps:-

1. Obtains the recipient B's public key (n, e).
2. Represents the plaintext message as a positive integer *m*, $1 < m < n$
3. Computes the cipher text $c = m^e \bmod n$.
4. Sends the ciphertext *c* to B.

Decryption:

Receiver B goes for following steps:-

1. Uses his private key (n, d) to compute $m = c^d \bmod n$.
2. Extracts the plaintext from the message representative *m*.

Advantages:

1. It is very the fast verification algorithm
2. Data is maintained securely at publisher spot.
3. It Provides Data Integrity , Data Security & Data Privacy
4. Checks all duplicated records & remove that information.
5. As Compared with previous methodology it gives the best utility and efficiency results.

V. ARCHITECTURAL VIEW

The architecture diagram of the system shown below helps us to understand the system.

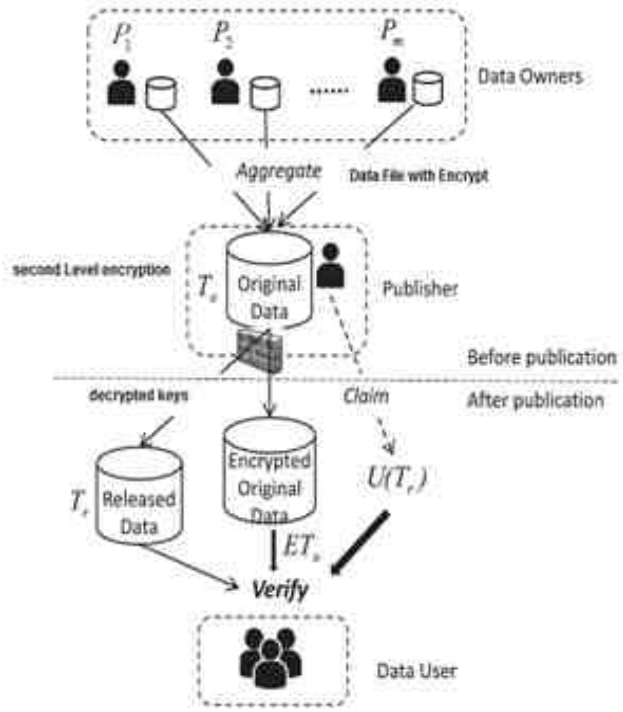

Figure1:-System Architecture

VI. SIMULATION RESULT

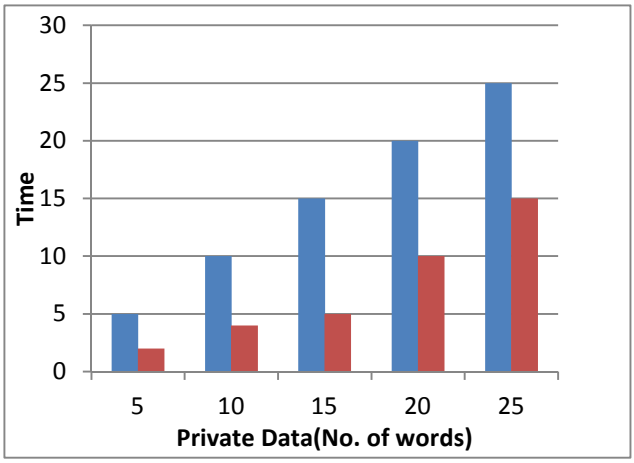

Figure 2:- Finding Private Dataset

The graph shows Private words (number of words) with respect to time.

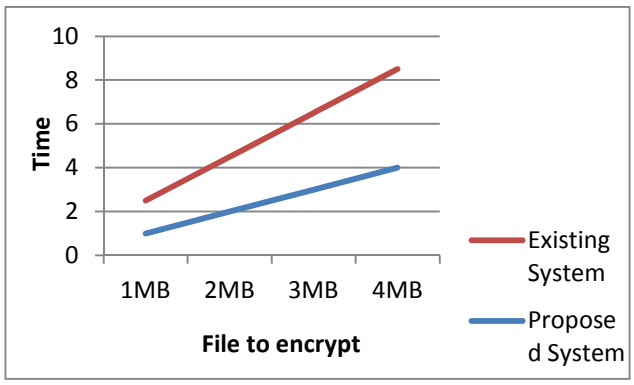

Figure3:-Performance Analysis

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

shows the execution time of existing and proposed methods. We have computed time by subtracting start time from end time for 4 separate file size ranging from 1 MB to 4MB.

Figure 4:- File size & Time Comparison

File size of existing system & proposed system with respect to time is compared.

VII. CONCLUSION

We develop new strategic module for data privacy for data on non-publishing sites, this project provide a very secure Communication trust between Reader, Publisher and writer. We all know that is one prime level. This system provides a very reliable and easy way to protect data from unethical activity. Privacy maintains one prime level. With user this system user fully aware of data security, privacy and data redundancy. So this system are fully satisfied our objective. In future work we want to implement same system on multimedia content and data.

REFERENCES

1. L. Fan, L. Xiong, and V. Sunderam, FAST: Differentially private real-time aggregate monitor with filtering and adaptive sampling, in Proc. ACM SIGMOD Int. Conf. Manage. Data (SIGMOD), 2013, pp. 10651068
2. R. Chen, B. C. M. Fung, and B. C. Desai. (2011). “Differentially private trajectory data publication.” [Online]. Available: http://arxiv.org/abs/1112.2020.
3. D. M. Freeman, Converting pairing-based cryptosystems from composite order groups to prime-order groups, in Proc. 29th Annu. Int. Conf. Theory Appl. Cryptogr. Techn. (EUROCRYPT), 2010, pp. 4461.
4. B. C. M. Fung, K. Wang, R. Chen, and P. S. Yu, Privacy-preserving data publishing: A survey of recent developments, ACM Compute. Surv., vol. 42, no. 4, 2010, Art. no. 14.
5. Y. Hong, J. Vaidya, H. Lu, P. Karras, and S. Goel, Collaborative search log sanitization: Toward differential privacy and boosted utility, IEEE Trans. Dependable Secure Comput., vol. 12, no. 5, pp. 504518, Sep./Oct. 2015.
6. W. Jiang and C. Clifton, A secure distributed framework for achieving k-anonymity, Int. J. Very Large Data Bases, vol. 15, no. 4, pp. 316333, Nov. 2006.

AUTHOR PROFILE

Miss. Gouri Namdeo Kale is currently pursuing M.E (Computer) from Department of Computer Engineering, Jayawantrao Sawant College of Engineering, Pune, India. Savitribai Phule Pune University, Pune, Maharashtra, India -411007. She received her B.E (IT) Degree from SVERI’s college of engineering Pandharpur. Maharashtra, India. Her area of interest is Information Security

Dr. Prof. S.N. Kini. He is currently working as Asst. Prof. (Computer) at Department of Computer Engineering, Jayawantrao Sawant college of engineering Hadapsar Pune.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.