International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 04 Issue: 06 | June-2017 www.irjet.net p-ISSN: 2395-0072

Verifying and Correctness of Frequent Itemset Minining

Suraj Shinde, Akshay Mohite, Sandip Shinde, Prof.Mangesh Manke

Department of computer engineering, Savitribai Phule Pune University MH India.

Abstract - ***Frequent itemsets refer to a set of data values (e.g., product items) whose number of co-occurrences exceeds a given threshold. Frequent itemset mining has been proven important in many applications such as market data analysis, networking data study, and human gene association study. Constructing cryptographic proofs for verification (for deterministic guarantee) and artificial verification objects (for probabilistic guarantee) can be applied to most data mining algorithms. The challenge is that the design of proofs and verification objects has to be customized for different data mining algorithms. Intended method will implement a basic idea of completeness verification and authentication approach in which the client will uses a set of frequent item sets as the evidence, and checks whether the server has missed any frequent item set as evidence in its returned result. It will helps to client detect untrusted server and System will become much more efficiency by reducing time. In authentication process CaRP is both a Captcha and a graphical password scheme. CaRP addresses a number of security problems altogether, such asonline guessing attacks, relay attacks, and, if combined with dual-view technologies, shoulder-surfing attacks.***

Key Words: *Datamining, cloud computing, web base service*

1. INTRODUCTION

Digitalization provide greater efficiency in today's life so we try to all worked done automatically but it generate worst amount of data, storing and operationalize on this data is challenging worked. In data mininig outsourcing this computation to third party service provider it provide limited access resource client advantages like cost effective options this process through we try to take the advantage of data-mining-as-a-service (DMaS) paradigm. In this paper author performed frequent data mining tasks on outsourced data mining, frequent itemset is nothing but a set of data values which is product items it increasing threshold value of number of co-occurrences in set of data values. Frequent itemset mining plays crucial role in many applications like market analysis, human gen association study and in network data study. We consider the server that is potentially untrusted and tries to escape from verification by using its prior knowledge of the outsourced data. In our paper propose system define efficient probabilistic and deterministic verification approaches to check whether the server returned correct and complete frequent itemsets or not? In authentication process CaRP is both a graphical password scheme and a Captcha. CaRP addresses a number of security problems altogether, such as relay attacks, online guessing attacks and also if combined with dual-view technologies, shoulder-sufering attacks. Existing system frequent itemset mining that computationally intensive for that natural solution is the choose computationally powerful service providers for those clients of limited computational resources. It also focuses on the Correctness, completeness and removing the integrity in mining resultset.

2. LITERATURE REVIEW

In previous paper work is more similar to our thesis. These system work are evidence patterns constructed by the encoding methods are easily identify by untrusted user without getting prior knowledge of data. Our probabilistic approach is more efficient because our method only takes 600 seconds to generate 6900 evidence itemsets (i.e., 0.09 second per pattern) while the in existing system take 2 seconds to generate single evidence pattern. Robustness improves about the probabilistic verification approach against the attack shows in the previous paper. For verifying the result integrity of web content data propose an efcient cryptographic approach we also use the set of intersection verification protocol for this purpose. Constructing the proof on server we required less time for computation that means to spend less time on server increase the speed..

Sr. no	Title	Publication	Authors	Description
1	Efficient Verification of Web-Content Searching Through Authenticated Web Crawlers		Michael T. Goodrich, Charalampos Papamantho u, Duy Nguyen.	The search engine returns the cryptographic proof of the query result. Both the proof size and the verification time are proportional only to the sizes of the query description and showing the query result, but do not depend on the number or sizes of the web pages over whic search is shown
				present new authenticated data structures

2	Optimal verification of operations on dynamic sets		Charalampos Papamantho u, Roberto Tamassia, and Nikos Triandopoulo	that allow any entity to publicly verify a proof attesting the correctness of primitive set operations such as intersection, union, subset and set difference.
3	Privacy-preserving data mining from outsourced databases		Fosca Giannotti, Laks V. S. Lakshmanan, Anna Monreale, Dino Pedreschi,an d Wendy Hui Wang	Propose an attack model based on background knowledge and devise a scheme for privacy preserving outsourced mining. This scheme ensures that each transformed item is distinguishable with respect to the attacker's background knowledge, from at least k−1 other transformed items.
4	Audio: An integrity auditing framework of outlier-mining-as-a-service systems		Ruilin Liu, Hui Wang, Anna Monreale, Dino Pedreschi, Fosca Giannotti, and WengeGuo	present AUDIO, an integrity auditing framework for the specific task of distance-based outlier mining outsourcing. It provides efficient and practical verification approaches to check both completeness and correctness of the mining results.

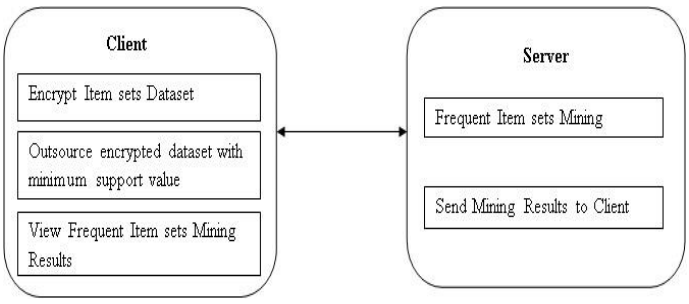

Fig -1: system Architecture

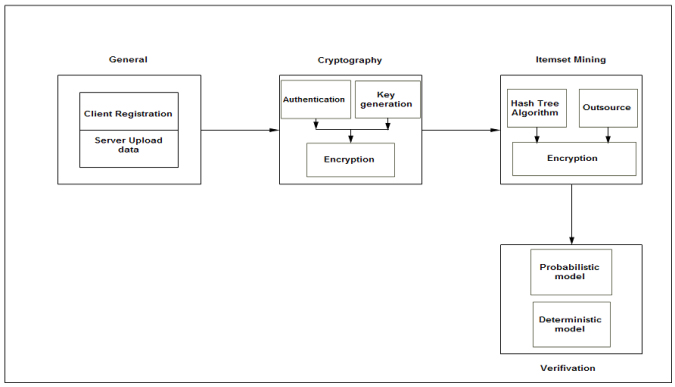

Fig -2:Existing system architecture

3.MODULES

3.1.Register and Login

In this module, the client can be authenticated whether the client is valid client or not. Register client Provide registration form which include client name, password, mobile no and address. Store client details in the database and authenticate client.

3.2. Dataset Encryption

In this module, Client wants to encrypt his dataset, because server was untrusted storage. So he generates the public key and private key for encryption using bilinear pairings. Then he encrypts the transaction dataset using public key. At the same he generates the signature based on Merkle hash trees.

3.3. Outsource Encrypted Dataset

In this module, client can outsource his encrypted dataset to server for frequent itemsets mining. Here client send the frequent itemsets mining request to server with encrypted dataset and support threshold value.

3.4. Frequent Item sets Mining & Verification

In this module, the server performs frequent itemset mining on the received dataset and returns the mining results to the client. Given a transaction dataset D that consists of n transactions, let I be the set of unique items in D. The support of the itemset I ⊆ I (denoted as supD(I)) is the number of transactions in D that contain I. An itemset I is frequent if its support is no less than a support threshold minsup. Clearly the search space of all frequent itemsets is exponential to the number of items in D. The (in)frequent itemsets behave the monotone property. For any given infrequent itemset I, any itemset I0 s.t. I ⊆ I0 must be an infrequent itemset. Similarly, for any frequent itemset I, any itemset I0 ⊆ I must be a frequent itemset.After receive the

frequent itemsets from server, the client apply the proposed efficient probabilistic and deterministic verification approaches to check whether the server has returned correct and complete frequent itemsets.

3. CONCLUSIONS

In this paper, we present two integrity verification approaches for outsourced frequent itemset mining. The probabilistic verification approach constructs evidence (in)frequent itemsets. In particular, we remove a small set of items from the original dataset and insert a small set of artificial transactions into the dataset to construct evidence (in)frequent itemsets. The deterministic approach requires the server to construct cryptographic proofs of the mining result. The correctness and completeness are measured against the proofs with 100% certainty. Our experiments show the efficiency and effectiveness of our approaches. An interesting direction to explore is to extend the model to allow the client to specify her verification needs in terms of budget (possibly in monetary format) besides precision and recall threshold.

REFERENCES

1. S. Goldwasser, S. Micali, and C. Rackoff. "The knowledge complexity of interactive proof systems" *SIAM Journal of Computing*, 18:186– 208, February 1989.

1. Laszlo Babai, Lance Fortnow, Leonid A. Levin, and Mario Szegedy. "Checking computations in polylogarithmic time" *In STOC*, pages 21–32, 1991.

2. Rosario Gennaro, Craig Gentry, and Bryan Parno. "Non-interactive verifiable
3. computing: outsourcing computation to untrusted workers" *In CRYPTO*, pages 465–482, 2010.

4. Charalampos Papamanthou, Roberto Tamassia, and Nikos Triandopoulos. "Optimal verification of operations on dynamic sets" *In CRYPTO*, 2011.

5. Hakan Hacig¨um¨us¸, Bala Iyer, Chen Li, and Sharad Mehrotra. "Executing sql over encrypted data in the database-service-provider model" *In SIGMOD*, pages 216–227, 2002.

6. Feifei Li, Marios Hadjieleftheriou, George Kollios, and Leonid Reyzin. "Dynamic authenticated index structures for outsourced databases" *In SIGMOD*, pages 121–132, 2006.

7. Ian Molloy, Ninghui Li, and Tiancheng Li. "On the (in)security and (im)practicality of outsourcing precise association rule mining" *In ICDM*, pages 872–877, 2009.

8. Fosca Giannotti, Laks V. S. Lakshmanan, Anna Monreale, Dino Pedreschi, and Wendy Hui Wang. "Privacy-preserving data mining from outsourced databases" *In Computers, Privacy and Data Protection*, pages 411–426. 2011.

9. W. K. Wong, David W. Cheung, Ben Kao, Edward Hung, and Niko Mamoulis. "An audit environment for outsourcing of frequent itemset mining" *In PVLDB*, volume 2, pages 1162–1172, 2009.

10. S. Wiedenbeck, J. Waters, J. C. Birget, A. Brodskiy, and N. Memon, "PassPoints: Design and longitudinal evaluation of a graphical password system," Int. J. HCI, vol. 63, pp. 102–127, Jul. 2005.

11. Bin B. Zhu, Jeff Yan, Guanbo Bao, Maowei Yang, and Ning Xu, "Captcha as Graphical Passwords-A New Security Primitive Based on Hard AI Problems," IEEE Transactions on Information Forensics and Security, Vol. 9, No. 6, pp.891-904, June 2014.

BIOGRAPHIES

MR.SURAJ MILIND SHINDE pursuing the Bachelor of Engineering Degree in Computer Engineering at Dr. D.Y. Patil Institute of Engineering and Technology, Ambi, Pune

MR.AKSHAY SUBHASH MOHITE pursuing the Bachelor of Engineering Degree in Computer Engineering at Dr. D.Y. Patil Institute of Engineering and Technology, Ambi, Pune

MR.SANDIP DAULAT SHINDE pursuing the Bachelor of Engineering Degree in Computer Engineering at Dr. D.Y. Patil Institute of Engineering and Technology, Ambi, Pune