Everant.org/ETJ

Research Article

Engineering and Technology Journal ISSN: 2456-3358

Reliability with Accuracy: Verifying Correctness of Resultant of the Outsourced Frequent Item Set Mining In Data-Mining-As-A-Service Paradigm

Revansidda Mulge[1], Dr. Kiran KumariPatil[2]

[1]Student, (M.Tech), Dept. Of CSE, RU, Bengaluru, Karnataka, India
[2]Professor and Director, Dept. Of CSE, RU, Bengaluru, Karnataka, India

ARTICLE INFO

Corresponding Author:

Revansidda Mulge[1]
[1]Student, (M.Tech), Dept. Of CSE, RU, Bengaluru, Karnataka, India

ABSTRACT

Nowadays online applications lack reliability for establishing the integrity of user-generated information. Users may unknowingly own compromised devices, or intentionally publish forged information. In these scenarios, applications need some way to determine the correctness of autonomously generated information. Towards that end, this thesis presents a trust-but-verify approach that enables open online applications to independently verify the information generated by each participant. In order to enable independent verification, our framework allowsour application to verify more information from less trustworthy users andless information from more trustworthy users. Thus, an application can trade-off performance for more integrity, or vice versa. We apply the trust-but- verify approach to three different classes of online applications and show how it can enable 1) high-integrity, privacy-preserving, crowd-sourced sensing 2) non-intrusive cheat detection in online games, and 3) effective spam prevention in online messaging applications.

KEYWORDS: ***Trust-but-Verify, high-integrity, privacy-preserving, crowd-sourced sensing, non-intrusive cheat detection, spam prevention.***

I. Introduction

In outsourcing data for data mining computation to third party creates challenges such as correctness and completeness data mining results. Though outsourced computation is cost effective the client of weak computational power cannot recognize whether the server has returned correct or incorrect results. This will create the possibility of incorrect results to be taken in consideration and generating wrong output.

In this paper we focus on verifying correctness of the frequent itemset mining at client end. This paper compares the results of client side with server side. If the results are incorrect then generate a report and send to the server. To verify the result we construct a preorder tree at both client and server end.

II. Background And Related Work

Let’s discuss about the existing system.

Existing system

In the existing system, the system focuses on frequent itemset mining as the outsourced data mining task. Informally, frequent itemsets refer to a

DOI: 10.18535/etj/v2i5.02

set of data values (e.g., product items) whose number of co-occurrences exceeds a given threshold. In the existing system, the system on the problem of verifying whether the server returned correct and complete frequent item sets. By correctness, we mean that all item setsreturned by the server are frequent. Completeness, we think that item set returned are correct result.

Our goal is to design efficient and robust integrity verification methods to catch such server that may return incorrect and incomplete frequent item sets. In particular, we make the following contributions.

First, we design the probabilistic approach to catch mining result that does not meet the predefined correctness/completeness requirement with high probability. The key idea is to construct a set of (in) frequentitem sets from real items, and use these (in) frequentitem sets as evidence to check the integrity of the server's mining results.

Second, we design the deterministic approach to catch any incorrect/incomplete frequent itemset mining answer with 100 percent probability. The key idea of our deterministic solution is to require the server to construct cryptographic proofs of the mining results.

Third, for both probabilistic and deterministic approaches, we provide efficient methods to deal with updates on both the outsourced data and the mining setup.

Last but not least, we complement our analytical results with extensive experiments evaluating the performance of our verification approaches.

III. Problem Formulation

A. Proposed system

Nowadays, due to the large applications of unsure data (e.g., noisy data), uncertain frequent item sets (UFI) mining over uncertain databases has attracted much attention, which differs from the corresponding deterministic problem from the generalized definition and resolutions.

As the most costly task in association rule mining process, it has been shown that outsourcing this task to a service provider (e.g.the third cloud party) brings several benefits to the ownership of data such as less commitment to storage and computation resources and cost relief. However, the correct secure mining result can be dishonest if the service provider is dishonest (e.g., lazy, maliciousetc.).

In this paper, we focus on the integrity and verification issues in UFI mining problem during the process of outsourcing, i.e., how the data owner verifies/authenticates the mining results.

Specifically, we evaluate and enlarge the present work on deterministic FI outsourcing verification to uncertain scenario. For this cause, we expand the existing/present outsourcing FI mining work to undetermined area w.r.t. the two popular UFI definition criteria and the approximate UFI mining methods.

Specifically, we introduce basic/enhanced verification scheme with such a different UFI definition respectively. After that, we further discuss the scenario of existing approximation UFP mining, where we can see that our technique can provide good probabilistic guarantees about the correctness of the verification. Finally, we present the Data Recovery techniques when the result correctness is false.

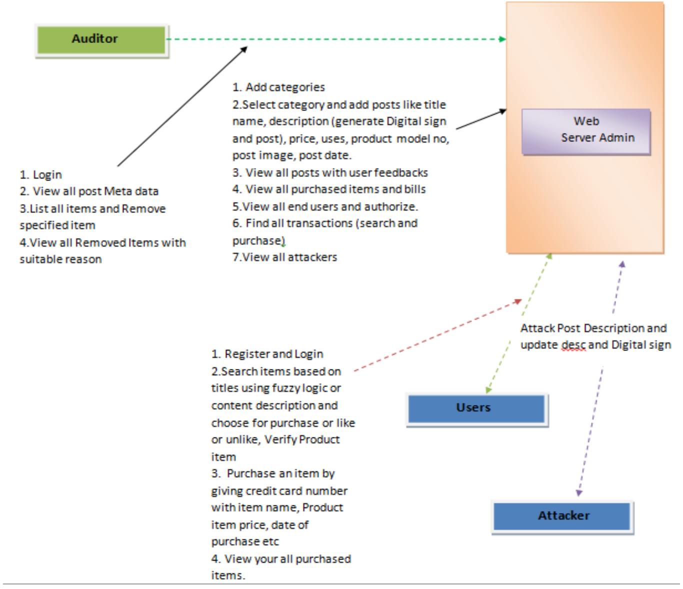
Auditor
1. Add categories
2.Select category and add posts like title name, description (generate Digital sign and post), price, uses, product model no, post image, post date.
3. View all posts with user feedbacks
4. View all purchased items and bills
5.View all end users and authorize.
6. Find all transactions (search and purchase)
7.View all attackers
Web Server Admin
1. Login
2. View all post Meta data
3.List all items and Remove specified item
4.View all Removed Items with suitable reason
Attack Post Description and update desc and Digital sign
1. Register and Login
2.Search items based on titles using fuzzy logic or content description and choose for purchase or like or unlike, Verify Product item
3. Purchase an item by giving credit card number with item name, Product item price, date of purchase etc
4. View your all purchased items.
Users
Attacker

Fig. 1. Architecture diagram

B. Design goals

The process of verification follows the following steps:

1. Web Server Admin

In this module, the Admin has to login by using valid user name and password. After login successful he can perform some operations such as Authorizing users, Adding Categories, Adding Posts based on Categories, Viewing all purchased items by Users, Viewing All Users Feedback on posts, All Search Transactions and viewing all attackers.

A. Adding Categories and Posts

In this module, the admin adds categories, details such as category name. The product posts will be added by selecting category and giving the details. Details such as product image, product title, model of product, uses of product, description and price of that product. The metadata of the products will be stored into the auditor’s database. The metadata details such as, post name, category, price, model number, digital sign. This digital sign will play a major role while verifying the products by the users. The digital sign is created by using the “SHA1” Algorithm. It creates the digital sign based the products description.

B. View Posts with its Feedbacks

In this module, the admin can view all the product posts added by himself. The feedback of that particular product posts can be viewed by the admin which is provided by the users while searching.

C.Purchased Items and Bills

In this module, the admin can see all the products which are all purchased by the users. The details such as product name, model, username, date and the bill on purchased items.

D. Attackers Details

In this, all the details of attackers who all attacked the products description can be viewed. Details such as attacker name, attacker image, attacked product details and the date will be displayed.

2. Attackers

The attacker searches for products by giving some keywords which matches with the description of the products and updates the products description by entering or by removing some data from existing description. Once the data is updated, the new digital sign will be created and will be updated into the products original digital sign, but not updated in the auditor’s database.

3. User

In this module, there are multiple number of users present. User should register before performing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user can perform some operations like viewing their profile details, searching for product posts, like and unlike on post, verifying the product, purchasing the product and viewing all his purchased products.

A. Searching Products

In this module, the user searches for products based on the product description added by the admin. The user can see product details such as, product image, category, model number, price, uses, likes, unlike, description if his searched keyword matches with the description.

Once user get products he can like or unlike or verify or can purchase that product by providing his credit card number. In this, the product verification happens by comparing the digital sign of product created while adding that product by the admin and the digital sign which is present in the Auditor’s database which is stored while adding the posts by the admin. The metadata of the products will be stored in Auditor’s database.

While verifying the products by the users if the digital sign present in the products database and in the Auditor’s database will be get compared. If both the digital signs matches then the results will be shown as “Product is safe and assured” else “Product is Not Safe” will be displayed.

B. Purchased Products

In this module, the user can see all his purchased products. The details such as product name, category, model number, price and the purchased date.

4. Auditor

In this module, the Auditor has to login by using valid user name and password. After login successful he can perform some operations such as Viewing all products metadata, removing the Products by giving suitable reasons.

A. View Posts Metadata

In this module, the auditor can see all the product post's metadata. The details such as product name, model, category, digital sign.

B. Removing Products

In this module, the auditor can remove products by giving suitable reason.

C. Viewing Removed Products

In this module, the auditor can view all the removed products removed by himself. The details such as, product name, model, category and the reasons provided by the auditor while removing products.

IV. Conclusion and Future Scope

In this paper, we demonstrate two reliability affirmation approaches for outsourced visit itemset mining. The probabilistic check approach creates demonstrate (in)frequent itemsets. In particular, we oust somewhat set of things from the principal dataset and install a little course of action of fake trades into the dataset to create affirm (in)frequent itemsets.

The deterministic procedures require the server to assemble cryptographic confirmations of the mining result. The rightness and satisfaction are measured against the confirmations with 100% conviction.

Our examinations exhibit the efficiency and practicality of our approaches. A fascinating course to examine is to build up the model to empower the client to show her check needs in wording of spending arrangement (possibly in cash related setup) other than precision also, audit constrain.

References

1. Srinath Setty, Andrew J. Blumberg, and Michael Walfish. Towardpractical and unconditional verification of remote computations. In Hot OS, 2011.
2. Feida Zhu, Xifeng Yan, Jiawei Han, Philip S. Yu, and Hong Cheng. Mining colossal frequent patterns by core pattern fusion. In ICDE,2007.
3. N. Baughman and B. Levine. Cheat-proof playout for centralized and distributed online games. In IEEE INFOCOM 2001.
4. C. Castelluccia, E. Mykletun, and G. Tsudik. E_cient aggregation of encrypteddata in wireless sensor network.
5. Rosario Gennaro, Craig Gentry, and Bryan Parno. Non-interactive verifiable computing: outsourcing computation to untrusted workers. In CRYPTO, pages 465–482, 2010.
6. Fosca Giannotti, Laks V. S. Lakshmanan, Anna Monreale, Dino Pedreschiand Wendy Hui Wang. Privacy-preserving data mining from outsourced databases. In Computers, Privacy and Data Protection, pages 411–426. 2011.
7. S. Goldwasser, S. Micali, and C. Rackoff. The knowledge complexity of interactive proof systems. SIAM Journal of Computing, 18:186– 208, February 1989.
8. Feifei Li, MariosHadjieleftheriou, George Kollios, and Leonid Reyzin. Dynamic authenticated index structures for outsourced databases. In SIGMOD, pages 121–132, 2006.

DOI: 10.18535/etj/v2i5.02