

Enhancement in privacy preservation in cloud computing using apriori algorithm

Raniah Ali Mustafa, Haitham Salman Chyad, Jinan Redha Mutar

Department of Computer Science, Collage of Education, Mustansiriyah University, Baghdad, Iraq

Article Info

Article history:

Received Jan 13, 2022

Revised Apr 9, 2022

Accepted May 19, 2022

Keywords:

Apriori algorithm

Associate rule mining

Elgamal cryptosystem

Encryption

Secure plain text equality test

ABSTRACT

Cloud computing provides advantages, like flexibly of space, security, cost optimization, accessibility from any remote location. Because of this factor cloud computing is emerging as in primary data storage for individuals as well as organisations. At the same time, privacy preservation is an also a significant aspect of cloud computing. In regrades to privacy preservation, association rule mining was proposed by previous researches to protect the privacy of users. However, the algorithm involves creation of fake transaction and this algorithm also fails to maintain the privacy of data frequency. In this research an apriori algorithm is proposed to enhance the privacy of encrypted data. The proposed algorithm is integrated with elagmal cryptography and it does not require fake transactions. In this way, the proposed algorithm improves the data protection as well as query privacy and it hides data frequency. Result analysis shows that the proposed algorithm improves the privacy as compared to previously proposed association rule mining and the algorithm also shows 3% to 5% improvement in performance when compared to other existing algorithms. This performance analysis with varying number of the data and fake transactions shows that the proposed algorithm doesn't require fake transactions, like data privacy association rule mining.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Raniah Ali Mustafa

Department of Computer Science, Collage of Education, Mustansiriyah University

Baghdad, Iraq

Email: rania83computer@uomustansiriyah.edu.iq

1. INTRODUCTION

Cloud computing provides a set of the advantages, like flexibly of space, security, cost optimization, and accessibility from any remote location. In cloud computing a database can be outsourced and security and remote access provides an incredible benefit to end users. However, with the advancement in cloud security, privacy preservation is being focused on securing the privacy of personal and organizational level outsourced data. As the private data may contain important personal information like credentials and private information, it should be protected from other users, including administrator, system level users and cloud server itself. Hence database encryption is a mandatory step which needs to be done before database outsourcing. One of the most popular data mining techniques (DM) for determining the relationships between entities in huge data sets is the association rule mining. As the database size is large and it may contain useful hidden information, associate rule mining is used to extract the important relational information from the database. Although associate rule mining extracts useful information but at the same time it compromises the privacy of the stored data. To avoid this issue a new privacy preserving associate rule mining was developed which finds the hidden information of data and protect the privacy of datasets. In regrades to privacy preservation, association rule mining was proposed by previous researchers in order to protect user privacy

[1]–[6]. Results show that the privacy preservation scheme that has been proposed in this research manages to protect the privacy, but those algorithms involve creation of fake transaction and the algorithm also fails to maintain the privacy related to data frequency. By using data frequency, while performing query processing, private information can be obtained even after performing data encryption and query encryption. In this research, a privacy persevering model for associate rule mining is used to access the encrypted dataset in the cloud servers. In this research, an apriori algorithm is used in the associate rule mining to find the relationship among different tuples. It is the mostly preferred algorithm for frequent dataset mining [7]–[9]. A plain text comparison test is carried out in order to check whether two cipher texts are similar. In this way, the proposed algorithm improves the data protection as well as query privacy and it hides data frequency.

To improve the security and privacy related concern in cloud computing, a privacy preserving associate rule mining was proposed. For this problem Wong initially proposed a one tuple to many tuples relational mapping in the cloud datasets, an advantage of this mapping was the converted transaction without deterministic approach. But the major disadvantage of this approach is easily distinguishable fake item from original datasets as the probability of both of the transaction types are the same [10]. In the paper, Giannotti, *et al.* [11] proposed a k-anonymity-based association rule mining algorithm. The proposed algorithm is based on fake transaction, it adds multiple fake datasets in existing datasets and hence, the frequency of every item becomes $k-1$. The drawback of this algorithm is if the fake transactions are known to the attacker then the original datasets can be easily extracted. This algorithm also increases the computational time as there is an extra work to roll out the frequency of fake transaction.

Yi *et al.* [2] suggested an associate rule mining based on the k-anonymity but it focuses on the encrypted dataset. El-gamal cryptography technique has been used, which provides an advantage of query protection with data protection. But this algorithm is also relying on the fake transaction and hence additional computation is required to insert fake encrypted transactions. In this algorithm frequency calculation is carried out using conditional gates which are based on the cipher text and its binary array. However, the data frequency of query processing is not encrypted in this algorithm; hence if attacker is aware of the data frequency then original datasets can be easily exposed. In the paper, Saravanan [12] an association rule mining (ARM) method has been proposed, Rob frugal algorithm and frequent item-set rule mining (F-item set RM) has been utilized for identifying fake or reiterate rules in the mining process. The method relies on cryptography is prioritized for database privacy protection. Privacy preserving data mining (PPDM) is a solution for data mining privacy risks. Through a centralized database, a connection is established between the cloud and the data owner, and the unprocessed data is processed using the cloud's solutions. To protect data privacy, a homomorphic encryption technique and a secure comparison scheme are utilized. In the paper, Jabeen *et al.* [13] improved secured association rule mining (ISARM) is proposed for vertical and horizontal database segmentation. Then, for privacy assurance, k-Anonymization approaches such as generalization and suppression depend anonymization are used. Finally, the diffie-hellman encryption technique (DHET) is proposed to protect sensitive data and allow storage service providers to operate with encrypted data. The diffie-hellman method (DHM) is used to improve the system's overall quality by generating safe keys, ensuring that the actual data is protected more effectively. The newly presented approach is tested in a Java simulation environment, which shows that it achieves both privacy and security. In the paper, Kumar *et al.* [14] the principles of DM in cloud computing (CC) and data security are discussed. The present level of cloud computing security has been explored, as well as data privacy analysis, security audits, data monitoring, and other cloud computing security problems. Some difficulties have been partially solved in recent data protection studies on security and privacy challenges in CC. Users can extract large hidden predictive information from a virtually integrated data warehouse using cloud computing, saving storage and infrastructure costs.

In the paper, Priyadarsini *et al.* [15] employed ARM as information mining approach. They employed the apriori formula in particular for ARM. It's been established that the first apriori formula was created for sequential calculation; therefore, utilizing it for the parallel processing doesn't seem like good idea. As a result, they tweaked the apriori formula (FP Growth) to make it more suitable for parallel computing. For cloud computing, they used the CloudSim machine. In the paper, Duc [16] from ARM to classification to clustering, there have been a number of secure methods. Randomization and safe multi-party computation are the two main methodologies in PPDM. The former uses statistical features to disguise sensitive data by adding noise to the underlying values. To prevent adversaries from seeing original data, the latter employs encryption techniques. The methods they offer in their thesis are based on the second approach. To compute the scalar product for several parties, we first introduce CSSP, an efficient privacy-preserving protocol. Thanks to homomorphic multiplicative cryptosystems, the protocol is developed employing caching techniques. CSSP beats prior work in terms of running time while keeping the same level of security when used to association rule mining challenges. Because data is updated on a regular basis, protocols must adapt to the changes. We suggest an incremental PPDMP for ARM with this goal in mind,

allowing participants to conduct mining activities on updated data rather than the complete data set. INCRE is a protocol that scans outdated databases only once, hence lowering computation overheads. They also performed tests to demonstrate the protocol's superiority to existing approaches. In the paper, Lakshmi and Rani [17] a model is provided that uses a sign-rely secure sum cryptography approach to find global association rules with trustworthy parties while maintaining the privacy of individual's data while it is dispersed in a horizontal manner across multiple locations fragment. Data may be partitioned in a variety of ways, including horizontal, vertical, and hybrid partitioning. Each fragment of data in horizontal partitioning consists of a sub-set of the records of a relation R, whereas each fragment of data in vertical partitioning consists of a sub-set of the characteristics of a relation R. Mixed fragmentation is another partitioning approach, in which data is partitioned in a horizontal manner and after that, every partitioned piece is further divided.

In the paper, Khurana and Bahl [18] association rule mining has been employed as a data mining technique in the dissertation. For association rule mining, they used the apriori method in particular. Because the original apriori method was created for sequential calculation, it does not appear to be a smart idea to use it for parallel computation. As a result, they tweaked the apriori algorithm (FP Growth) to make it more suitable for parallel computing. For cloud computing, they used CloudSim Simulator. In the paper, Redekar and Honwadkar [19] investigates the topic of outsourcing association rule mining (ARM) tasks in the cloud environment within corporate privacy-preserving frameworks. The study proposed a new approach that ensures that every converted item from data owner to server is interchangeable when compared to attackers' background information. On an enormously large and genuine transaction database that represents the entire system, their solutions are scalable, effective, and safeguard privacy. This method also recommends using the cloud to enable privacy-preserving mining. In a classical paradigm, it is assumed that the enemy educates the region of items and their exact occurrence and can use this information to recognize cipher items and cipher item sets. In the paper, Gai *et al.* [20] focused on the privacy importance of big data (BD) and take up cloud computing implementations. Suggest a new data encryption method, which is called dynamic data encryption strategy (D2ES). Our suggested method aims to eclectically encrypt data and utilize privacy classification methods under thoroughness constraints. This method is prepared to maximize the privacy protection domain through utilizing an eclectic encryption strategy within the requested execution time (ET) requests.

The achievement of D2ES has been estimated in our experiences, which provides the evidence of the privacy increase. In the paper, Park *et al.* [21] suggeste a privacy-preserving reinforcement learning (PPRL) architecture for the cloud computing platform. For fully homomorphic encryption, the suggested framework uses a cryptosystem based on learning with errors (LWE) (FHE). The suggested PPRL framework's performance is analyzed and evaluated in a diversity of cloud computing-based on intelligent service scenarios. Rajput and Raman [22] suggest a new cloud based privacy-preserving technique for image color improves. Unlike other color correction methods, where the colors of the test image are treated in the plain domain with visible image contents, we suggest that color correcting operations be performed over the cloud in an encrypted domain. As a consequence, superior results are accomplished along with entire privacy confirmation. Furthermore, we suggest a block-based on image encryption approach based on the logistic-tent system and the ElGamal cryptosystem. As a result, when compared to the naive technique, the size of the encrypted image is greatly reduced. The suggested approach is determined to be highly effective when compared to state-of-the-art schemes, according to experimental results. A challenge response game model is used to demonstrate the suggested approach's security strength.

In the paper, Zhou *et al.* [23] we look into ways to deliver multiple granular information views for diverse users. Based on a Galois connection, our technique first creates the association between the keywords and data files. Then we employ data retrieval indexes with changeable thresholds, which allow for granular data retrieval by altering the threshold for various users. We study the issue of privacy-preserving granular data retrieval service (PPGDRS) on cloud where, we present a differentially privacy release approach rely on the suggested index technique to prevent privacy leak. We illustrate the proposed method's privacy-preserving guarantee, as well as the validity of the suggested mechanism through comprehensive tests. In the paper, Rong *et al.* [24] suggested approach k-nearest neighbor (kNN) computation over the databases distributed through multiple cloud environments for privacy-preserving. Unluckily, existent secure outsourcing protocols are either restrictive to a single key setting or perfectly effectual because of repeated client-to-server interactions, making it unpractical for wide implementation. To address these issues, we suggeste a set of secure building blocks and outsourced cooperative kNN protocol. Theoretic analyses illustrates that under the semi-honest scenario, our solution not only protects the privacy of distributed databases and kNN queries, but also hides access patterns. When compared to previous methodologies, the experimental evaluation shows significant efficiency gains. In regrades to privacy preservation, association rule mining was proposed by previous researches to protect the privacy of users. The results shows that the privacy preservation schema proposed in research, manage to protect the privacy but those algorithms involve creation of fake transaction and the algorithm also fails to maintain the privacy related to data

frequency. By using the data frequency, while performing query processing, private information can be obtained even after performing data encryption and query encryption. In the presented study apriori algorithm is suggested to enhance the privacy of encrypted data. The suggested algorithm has been integrated with elgmal cryptography and it does not require fake transactions. The suggested architecture and flow chart of the apriori algorithm and Elgmal encryption algorithm are provided in section 2. The performance analysis will be completely indicated in section 3. Lastly, section 4 provides the conclusions.

2. PROPOSED ARCHITECTURE

Current cloud systems consist of data mining protocol in stored dataset which is either partially honest or has an access to personal dataset [25]. Even though a security protocol has been developed to protect the cloud storage from outsider attacks, privacy preservation inside the cloud is equally essential. In the partial honest cloud architecture, system follows the protocol and standards but internal algorithms still have an access to critical data which should be protected and hence it violates the privacy concern of cloud storage. In an unregulated cloud storage model, cloud services only provide a remote access, storage and security and it does not compromise with data privacy protocols and standards. In the proposed research, existing partial honest cloud architecture is adopted and developed further [26].

The proposed architecture consists of four entities, which are: Data provider, two cloud servers, and service user. In this model, data provider is an actual owner of the original data. Service user is an entity looking for data from cloud storage and it is also an authorised entity. In the proposed architecture secure 2 party computational protocols is used. It is used for secure computations between two cloud servers. The workflow of the proposed system architecture is as below. Initially data provider generates a secret and public key using Elgmal encryption algorithm. Once key generation is completed, data provider sends a generated public key and encrypted datasets to first server [27]–[29]. After this, data provider provides an elgmal public-private key to the second cloud server; parallely it sends a generated public key to the service user. After the sharing of key pairs and encrypted data, service users make a request in the form of encrypted query and send it to first server. Secure plain-text equality test is used in the proposed architecture to check if two encrypted datasets are equal or not without actually decrypting it. Using the secure 2 party computational protocol, two cloud servers collaborate together and the first server performs apriori algorithm using secure plain-text equality test [30], [31]. Figure 1 illustrates the proposed cloud system architecture.

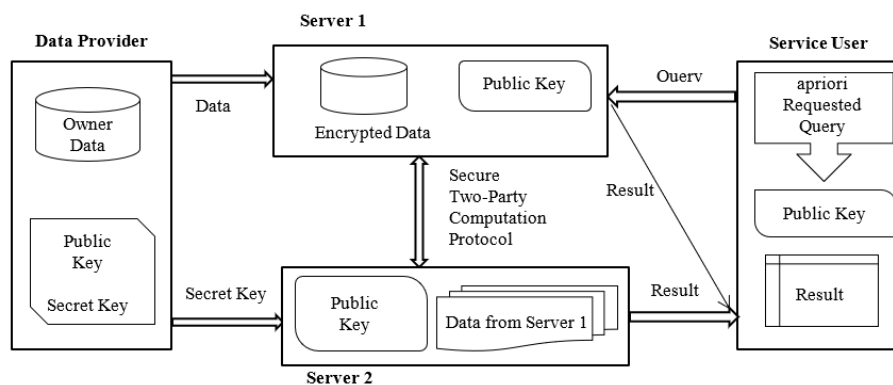


Figure 1. Proposed cloud system architecture

Now apriori algorithm and Elgmal encryption algorithm steps are explained. They are used for enhancement in privacy preservation in cloud computing. Figure 2 explained the flow chart of the apriori algorithm and Elgmal encryption algorithm.

2.1. Secure plain-text equality test

Secure plain-text equality test algorithm returns values based on the input comparison. The output value could either be true or false. It returns true if two plain-text values are identical and false if it is not. This section represents the secure plain-text equality test algorithm used in this architecture. Input of the algorithm is two cipher texts (input 1 and input 2) and the output is either true or false. Initially, server 1 (ES1) generates a composite number t , and sends $input1(t)$ to server 2 (ES2). Using the composite number

input1 (t), input2 (t) and both cipher inputs, using this value the front of S1 (t* input1) & S2 (t* input2) is generated and it is represented as g^{r1} and g^{r2} respectively. This data is then transferred to Server 2. Using the secret key on g^{r1} and g^{r2} Server 2 generates output g^{r1x} and g^{r2x} where x is a secret key provided by data provider after key generation phase. These values are then transferred to server 1, where they generate an output (α) of an algorithm using t, g^{r1x} and g^{r2x} . Output of this step is either true or false which defines whether two cipher-texts are similar or not. Figure 3 explained the flow chart for the secure plain-text equality test algorithm.

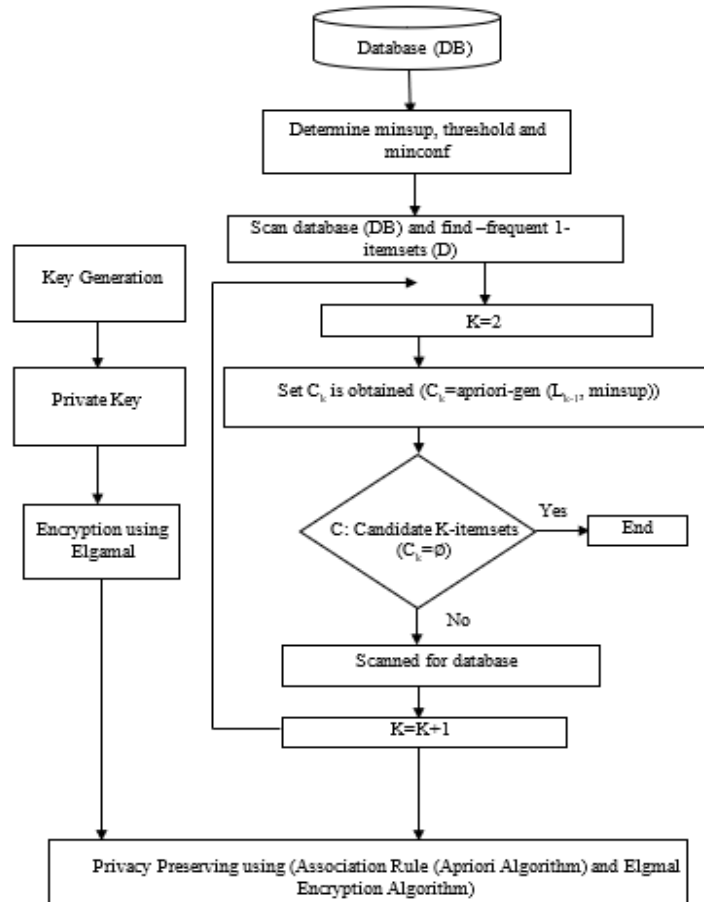


Figure 2. Flow chart of the apriori algorithm and elgamal encryption algorithm

In this research, a privacy preservation in cloud computing using apriori algorithm and secure plain-text equality test algorithms is proposed. The proposed algorithm consists of two important steps which are generation of candidate set and calculation of frequency set.

Algorithm 1:

- 1: Server 2 generate a composite number
- 2: Server 2 send Encrypted composite number to server 1
- 3: Compute g^{r1} and g^{r2} and send to server 2
- 4: Server 2 Compute xx using secret key and send to server 1
- 5: Calculate
- 6: If equals 1
return true
- 7: Else
return false

Input: Encrypted Text 1 and Encrypted Text 2

Output: Boolean

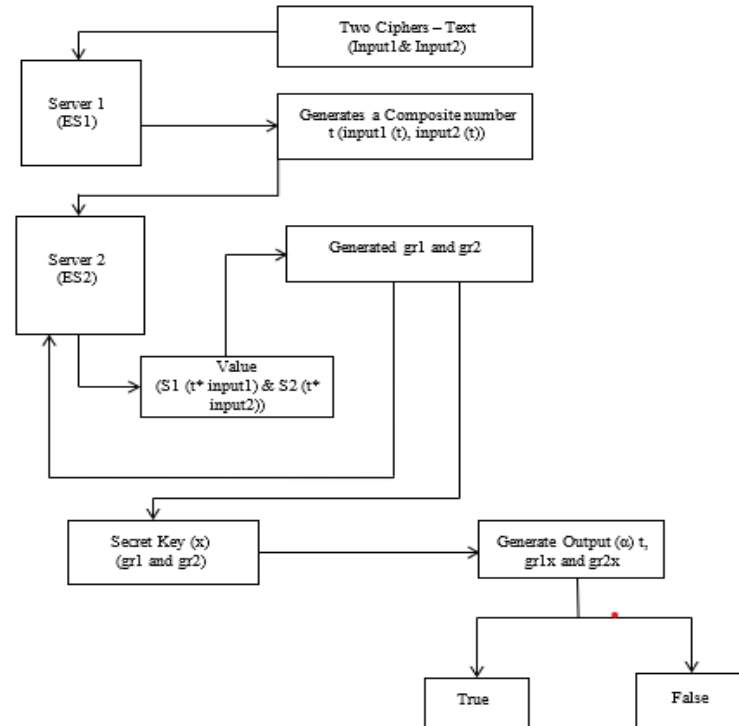


Figure 3. Flow chart of the secure plain-text equality test algorithm

2.2. Generation of candidate set

In this step, a candidate set is generated. In each candidate set, multiple patterns are present. Also, each pattern consists of multiple items. Candidate set generation involves the following steps. Initially a single pair of patterns is selected from $k-1$ frequent dataset. Considering the selected pair of patterns is $\langle p1, p2 \rangle$. In this equation $p1$ and $p2$ are different selected patterns. After this step, a join is performed between items of pattern 1 and pattern 2. The result of this join is added to the candidate set. It is indicated as S_k where k is the number of items. After this join is performed on all pair of excluding items of pattern 1 and pattern 2 and resultant set S_k is transferred to cloud server 1.

2.3. Calculation of frequent set

In this step, a frequency set of candidates set (S_k) is calculated. Algorithm 2 illustrates the pseudo code for the calculation of frequency set. Initially, one pattern from candidate set is selected. After this, secure plain-text equality test is carried out for all items belonging to the transaction or belonging to selected pattern. If this test returns true value then the matched item count is increased by one. Using mathematical functions and arbitrary integer value (g), $E(x, \text{sup})$ and $E(g^{\text{minsup}})$ are calculated. If the value of secure plain-text equality test is true, then frequent attributes of x are added into the frequent set. The same pseudo code is used to find the frequent set for other patterns of candidate set.

Algorithm 2:

```

1: Calculate frequency set
2: Selected one pattern
3: Secure plain-text equality test
4: If test
   return true
5: Increment i value by 1
  
```

2.4. Proposed apriori algorithm

Algorithm 3 represents the Proposed apriori Algorithm. Initially the $L1$ is set to single item set which is received from the data owner. In next step, generation of candidate set algorithm is used, candidate set is generated and later algorithm 2 is used, frequency set of candidate set is calculated. In this stage, if there is no generation of k frequent set then $k-1$ frequent set is returned, where k is a length of itemset. Figure 4 explained the flow chart for the apriori algorithm (PA).

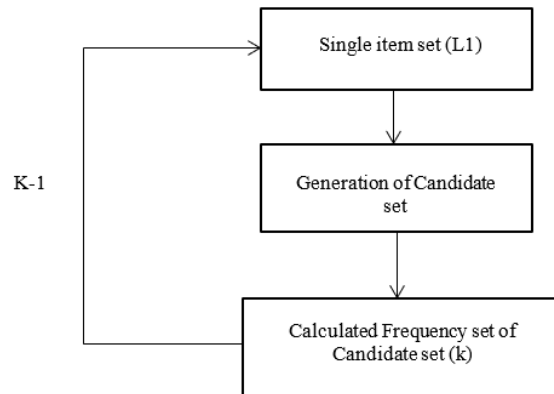


Figure 4. Flow chart for the apriori algorithm

Algorithm 3:

```

1: Receive Frequency item set from data owner
2: Initialize variable i with 2
3: while (true)
4:   var = generate candidate set
5:   if (var == null)
6:     return encrypted frequency set to service user
7:   calculate frequency set
8:   Increment i value by 1
9: end while
  
```

3. RESULTS AND DISCUSSION

3.1. Security analysis

To determine the security vulnerability of the proposed algorithm, security proof analysis is performed. With respect to first cloud server, data frequency is present in the encrypted format and the database present in this server is also in encrypted format. As the proposed Elgamal encryption algorithm provides different encrypted data for the same plain text data after every encryption process, there is guarantee of no leakage of data. With respect to second cloud server, as the original data doesn't contain the front of cipher text, there are no chances of data exposure from second cloud server. This shows that the proposed algorithm maintains the partial honest standard of cloud computing infrastructure.

3.2. Performance evaluation

For evaluation of performance analysis of the proposed secure ARM is carried out. For the evaluation purpose, windows operating system is used. For avoiding any CPU overhead and deadlock issue 32 GB of DDR 4 RAM is configured. Intel Xeon E3-1220 processor is used. For the generation of big integer GMP (GNU Multiple Precision Arithmetic Library) library with 6.2.1 version is used. GMP is a free arithmetic library that works with signed integers, rational numbers, and floating-point numbers with variable precision. Xun also suggested a data privacy ARM, rely on the k anonymity which focuses on the encrypted dataset as well as query encryption [11]. Hence the proposed algorithm is the comparison with the algorithm that has been proposed by Xun. For performance analysis a database which is selected for the is based on retail data [27]. Table 1 show system information.

Operating system	Windows 10
RAM	32 GB DDR 4
Processor	Intel Xeon E3-1220
Library	GMP
Library version	6.2.1

In this analysis the number of data is also varied to get the broader performance overview. The minimum support is also varied for carrying performance evaluation at different minimum support value ranging from 5% to 30%. For the evaluation with respect to data size, count of data varies from 2,000; 4,000;

6,000; 8,000 and 10,000. To determine the impact of fake transactions, fake transaction ratio is differentiated from 50% and 100%. Key size for advanced encryption algorithm is 1,024. Table 2 explained details of the analysis evaluation of performance for the secure association rule mining.

Table 2. Analysis evaluation of performance for the secure association rule mining

Minimum support (data size)	Count of data (impact of fake transactions)	Fake transaction ratio	Key size (advanced encryption algorithm)
5% \longrightarrow 30%	2,000; 4,000; 6,000; 8,000; and 10,000	50% and 100%	1,024

Figure 5 illustrates the comparison of the proposed algorithm with the data privacy association rule mining. With fake transaction ratio of 50% and minimum support of 10%, the proposed algorithm shows 205% improvement as compared to data privacy association rule mining. With fake transaction ratio of 100% and minimum support of 10%, the proposed algorithm shows 405% improvement as compared to data privacy association rule mining. This result shows that the proposed algorithm does not require fake transactions like data privacy association rule mining. It also shows that the proposed algorithm doesn't require binary operation as secure plain-text equality test is implemented. Where the result in Table 3 explained details of the improvement proposed algorithm as compared to data privacy association rule mining (ARM).

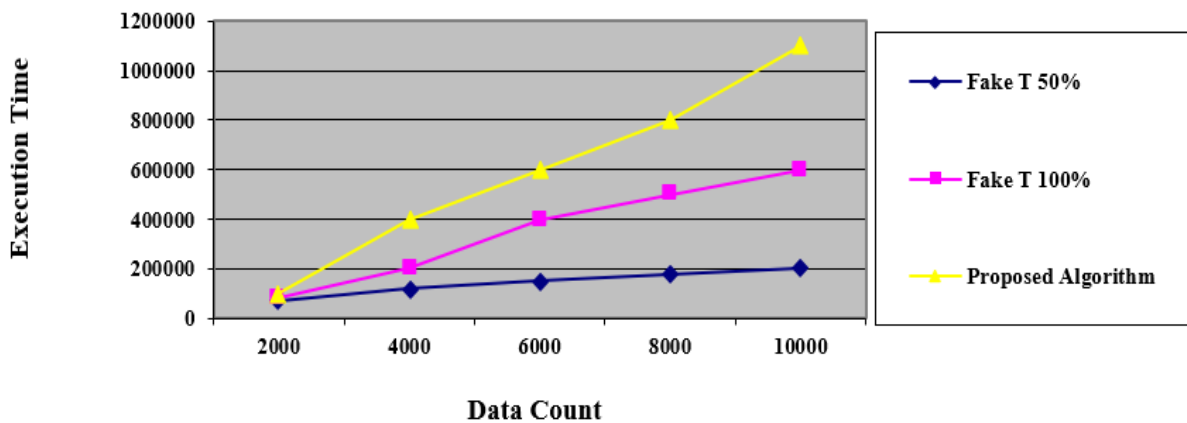


Figure 5. Evaluation using different count of data

Table 3. Improvement privacy association rule mining (ARM)

Fake transaction ratio	Minimum support	Improvement proposed algorithm
50%	10%	205%
100%	10%	405%

Figure 6 illustrates the comparison of the proposed algorithm with the data privacy association rule mining with respect to various minimum support values. With count of data as 10,000 and fake transaction ratio of 50%, the proposed algorithm shows 216% improvement compared to data privacy association rule mining. With count of data as 10,000 and fake transaction ratio of 100%, the proposed algorithm shows 429% improvement compared to data privacy association rule mining. This result shows that the proposed algorithm does not require fake transaction like data privacy association rule mining.

Where the result in Table 4 explained details of the improvement proposed algorithm as compared to data privacy ARM. Table 5 shows algorithms and methods that were utilized in the previous works. In terms of comparison, the performance analysis is noticed of privacy preservation in cloud computing utilizing apriori algorithm (PA) more effective.

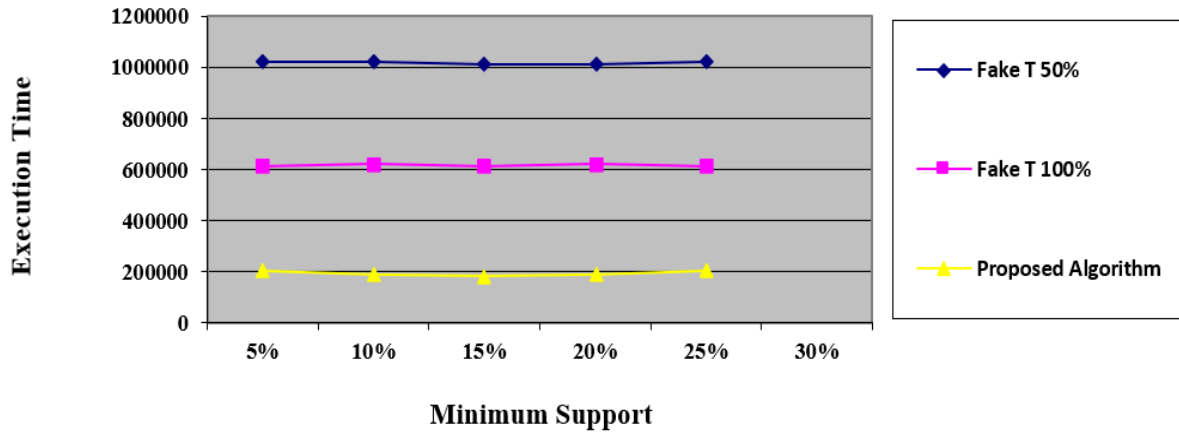


Figure 6. Evaluation using different minimum support value

Table 4. Improvement privacy association rule mining (ARM)

Count of Data (impact of fake transactions)	Fake transaction ratio	Improvement proposed algorithm
10000	50%	216%
10000	100%	429%

Table 5. Illustrates a comparison of previous systems

Ref	Author	year	Algorithm/ method
[2]	X.Yi, <i>et al.</i>	2015	associate rule mining based on k anonymity and elagmal cryptography technique
[11]	F. Giannotti, <i>et al.</i>	2018	k anonymity-based association rule mining algorithm
[12]	M.Saravanan	2018	association rule mining (ARM), Rob frugal algorithm
[13]	T. N. Jabeen <i>et al.</i>	2016	ISARM (Improved Secured Association Rule Mining)
[14]	V. Kumar <i>et al.</i>	2021	DM techniques in cloud computing and data security
[15]	Dr.S.Priyadarsini, et al	2016	Apriori formula (FP Growth)
[16]	T. H. Duc	2021	association rule mining to classification to clustering,
[17]	M. N. Lakshmi <i>et al.</i>	2012	uses a sign-rely secure sum cryptography approach to find global association rules (GAR)
[18]	H. Khurana <i>et al.</i>	2014	Apriori algorithm (FP Growth)
[19]	V. R. Redekar <i>et al.</i>	2014	association rule mining
[20]	K. Gai, <i>et al.</i>	2017	Dynamic Data Encryption Strategy (D2ES) for privacy importance of big data (BD) and take up cloud computing (CC) implementations
[21]	J. Park, <i>et al.</i>	2020	SCC-PPRL algorithm using FHE scheme
[22]	A.S. Rajput <i>et al.</i>	2019	logistic-tent system and ElGamal cryptosystem
[23]	Z. Zhou, <i>et al.</i>	2014	association among the keywords and data files rely Galois connection (GC), variable threshold fuzzy concept lattice
[24]	H. Rong, <i>et al.</i>	2016	k-nearest neighbor (kNN)

4. CONCLUSION

Cloud computing is emerging as in primary data storage for individual as well as organisations, privacy preservation is an also important aspect of cloud computing. In this research an apriori algorithm is proposed to enhance the privacy of encrypted data. The proposed algorithm is integrated with elgamal cryptography and it does not require fake transactions. The proposed secure plain-text equality test-based algorithm improves the data protection as well as query privacy and it hides the data frequency. It was observed that the proposed algorithm shows 3 to 5% improvement in performance when compared to other existing algorithms. This performance analysis with varying count of data and fake transactions shows that proposed algorithm doesn't require fake transaction like data privacy association rule mining. It also shows that the proposed algorithm doesn't require binary operation as secure plain-text equality test is implemented.

ACKNOWLEDGEMENTS

The authors would like to express their thanks to Mustansiriyah University (www.uomustansiriyah.edu.iq) Baghdad, Iraq for its support in the present study.




REFERENCES

- [1] H. J. Kim, J. H. Shin, Y. H. Song, and J. W. Chang, "Privacy-preserving association rule mining algorithm for encrypted data in cloud computing," in *IEEE International Conference on Cloud Computing, CLOUD*, Jul. 2019, vol. 2019-July, pp. 487–489, doi: 10.1109/CLOUD.2019.00086.
- [2] X. Yi, F. Y. Rao, E. Bertino, and A. Bouguettaya, "Privacy-preserving association rule mining in cloud computing," in *ASIACCS 2015 - Proceedings of the 10th ACM Symposium on Information, Computer and Communications Security*, Apr. 2015, pp. 439–450, doi: 10.1145/2714576.2714603.
- [3] N. Domadiya and U. P. Rao, "Privacy preserving distributed association rule mining approach on vertically partitioned healthcare data," *Procedia Computer Science*, vol. 148, pp. 303–312, 2019, doi: 10.1016/j.procs.2019.01.023.
- [4] L. Liu *et al.*, "Privacy-preserving mining of association rule on outsourced cloud data from multiple parties," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10946, 2018, pp. 431–451.
- [5] G. A. Afzali and S. Mohammadi, "Privacy preserving big data mining: association rule hiding," *Journal of Information Systems And Telecommunication (JIST)*, vol. 4, no. 2, pp. 70–77, 2016.
- [6] H. S. Chyad, R. A. Mustafa, and D. N. George, "Cloud resources modelling using smart cloud management," *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 2, pp. 1134–1142, Apr. 2022, doi: 10.11591/eei.v11i2.3286.
- [7] W. K. Wong, D. W. Cheung, E. Hung, B. Kao, and N. Mamoulis, "Security in outsourcing of association rule mining," in *33rd International Conference on Very Large Data Bases, VLDB 2007 - Conference Proceedings*, 2007, pp. 111–122.
- [8] S. R. M. Oliveira, O. R. Zaiane, A. Tosello, B. Geraldo, and O. R. Zaane, "Privacy preserving frequent itemset mining," in *Proceedings of the IEEE international conference on Privacy, security and data mining*, 2002, vol. 14, pp. 43–54, doi: 10.5555/850782.850789.
- [9] K. Bhuvanewari and K. Saravanan, "Privacy preserving association rule mining from highly secured outsourced databases," *International Journal of Computer Engineering & Technology*, vol. 8, no. 4, pp. 98–107, 2017.
- [10] A. Chawla and K. Singh Dhindsa, "Implementation of association rule mining using reverse apriori algorithmic approach," *International Journal of Computer Applications*, vol. 93, no. 8, pp. 24–28, May 2014, doi: 10.5120/16236-5759.
- [11] F. Giannotti, L. V. S. Lakshmanan, A. Monreale, D. Pedreschi, and H. Wang, "Privacy-preserving mining of association rules from outsourced transaction databases," *IEEE Systems Journal*, vol. 7, no. 3, pp. 385–395, Sep. 2013, doi: 10.1109/JSYST.2012.2221854.
- [12] M. Saravanan, "Privacy preserving and enhancing security in association rule mining using robfrugal algorithm," *International Journal of Advance Research and Innovative Ideas in Education*, vol. 2, no. 4, pp. 238–243, 2018.
- [13] T. N. Jabeen, M. Chidambaram, and G. Suseendran, "Security and privacy concerned association rule mining technique for the accurate frequent pattern identification," *International Journal of Engineering and Technology (UAE)*, vol. 7, no. 1.1, pp. 19–24, Dec. 2018, doi: 10.14419/ijet.v7i1.1.8908.
- [14] V. Kumar, M. Bala, and P. Choudhary, "Implementation of secure data mining approach in cloud using image encryption," *Communications on Applied Electronics*, vol. 5, no. 1, pp. 17–21, May 2016, doi: 10.5120/cae2016652186.
- [15] D. B. D. M. S. A. Prasanth, "An efficient privacy preserving in frequent item set for cloud environment using apriori," *Annals of the Romanian Society for Cell Biology*, vol. 25, no. 6, pp. 2934–2946, 2021.
- [16] H. D. Tran, "Speeding up privacy preserving data mining techniques," Doctoral thesis, Nanyang Technological University, Singapore, 2016, [Online]. Available: <https://hdl.handle.net/10356/68814>.
- [17] N. V. MuthuLakshmi and K. S. Rani, "Privacy preserving association rule mining in horizontally partitioned databases using cryptography techniques," *International Journal of Computer Applications*, vol. 39, no. 13, pp. 29–35, 2012, doi: 10.5120/4883-7321.
- [18] H. Khurana, K. Bahl, and D. Academics, "An approach to mine frequent itemsets in cloud using apriori and FP-tree approach," *IJCAT International Journal of Computing and Technology*, no. 7, pp. 1–3, 2014, Accessed: Apr. 19, 2022. [Online]. Available: www.IJCAT.org.
- [19] D. K. N. H. V. R. Redekar, "Privacy-preserving mining of association rules in cloud," *International Journal of Science and Research (IJSR)*, vol. 3, no. 11, 2014.
- [20] K. Gai, M. Qiu, and H. Zhao, "Privacy-preserving data encryption strategy for big data in mobile cloud computing," *IEEE Transactions on Big Data*, vol. 7, no. 4, pp. 678–688, 2021, doi: 10.1109/TBDATA.2017.2705807.
- [21] J. Park, D. S. Kim, and H. Lim, "Privacy-preserving reinforcement learning using homomorphic encryption in cloud computing infrastructures," *IEEE Access*, vol. 8, pp. 203564–203579, 2020, doi: 10.1109/ACCESS.2020.3036899.
- [22] A. S. Rajput and B. Raman, "Privacy-preserving smart surveillance using local color correction and optimized elgama1 cryptosystem over cloud," in *IEEE International Conference on Cloud Computing, CLOUD*, Jul. 2019, vol. 2019-July, pp. 73–80, doi: 10.1109/CLOUD.2019.00024.
- [23] Z. Zhou, H. Zhang, Q. Zhang, Y. Xu, and P. Li, "Privacy-preserving granular data retrieval indexes for outsourced cloud data," in *2014 IEEE Global Communications Conference, GLOBECOM 2014*, Dec. 2014, pp. 601–606, doi: 10.1109/GLOCOM.2014.7036873.
- [24] H. Rong, H. M. Wang, J. Liu, and M. Xian, "Privacy-preserving k-nearest neighbor computation in multiple cloud environments," *IEEE Access*, vol. 4, pp. 9589–9603, 2016, doi: 10.1109/ACCESS.2016.2633544.
- [25] H. J. Kim, H. Il Kim, and J. W. Chang, "A privacy-preserving kNN classification algorithm using Yao's garbled circuit on cloud computing," in *IEEE International Conference on Cloud Computing, CLOUD*, Jun. 2017, vol. 2017-June, pp. 766–769, doi: 10.1109/CLOUD.2017.110.
- [26] M. Nateghizad, T. Veugen, Z. Erkin, and R. L. Lagendijk, "Secure equality testing protocols in the two-party setting," in *Proceedings of the 13th International Conference on Availability, Reliability and Security*, Aug. 2018, doi: 10.1145/3230833.3230866.
- [27] H. K. Algebri, Z. Husin, A. M. Abdulhussin, and N. Yaakob, "Why Move toward the Smart Government," in *Proceedings - 2017 International Symposium on Computer Science and Intelligent Controls, ISCSIC 2017*, Oct. 2018, vol. 2018-Febru, pp. 167–171, doi: 10.1109/ISCSIC.2017.34.
- [28] H. S. Chyad, H. S. Chyad, R. A. Mustafa, and K. T. Saleh, "Study and implementation of resource allocation algorithms in cloud computing," *International Journal of Engineering & Technology*, vol. 7, no. 4.28, pp. 591–594, Nov. 2018, doi: 10.14419/ijet.v7i4.28.25394.
- [29] M. H. Kadhim, A. Bindal, and A. Professor, "An effective protocol and algorithmic approach for disaster management using wireless sensor networks," *Journal of Advanced Research in Dynamical and Control Systems*, vol. 10, no. 13, 2018.




- [30] H. K. Algabri, Y. A. Taha, S. S. Gaikwad, and R. K. Kamat, "Curriculum technology integration for higher education," *Journal of Advanced Research in Dynamical and Control Systems*, vol. 12, no. 1, pp. 295–300, Feb. 2020, doi: 10.5373/JARDCS/V12I1/20201043.
- [31] T. Brijs, "Retail market basket data set," *Workshop on Frequent Itemset Mining Implementations (FIMI'03)*, 2003.

BIOGRAPHIES OF AUTHORS






Raniah Ali Mustafa    she was born in Baghdad, Iraq, in 1983. She received her M. Sc. Degree in computer science from Mustansiriyah University, Baghdad, Iraq, in 2016. She joined a programmer position in the laboratories of the Department of Computer Science at Al-Mustansiriya University for the period from 2005 to 2013 and after obtaining a master's degree, she became a teacher in the Department of Computer Science. She can be contacted at email: rania83computer@uomustansiriyah.edu.iq



Haitham Salman Chyad    He was born in Baghdad, Iraq, in 1979. He received her M. Sc. Degree in computer system from Osmania University, India, Hyderabad, in 2016. He joined a programmer position in the laboratories of the Department of Computer Science at Al-Mustansiriya University for the period from 2007 to 2013 and manager of the division of electronic calculator of the College of Education at Al-Mustansiriya University and In addition to teaching in the Department of Computer Science. He can be contacted at email: dr.haitham@uomustansiriyah.edu.iq



Jinan Redha Mutar    she was born in Baghdad, Iraq, in 1969. She received her M. Sc. Degree in computer science from Mustansiriyah University, Baghdad, Iraq, in 2005. She joined a programmer position in the laboratories of the Electronic Calculator Center at Al-Mustansiriya University for the period from 1992 to 2005 and after obtaining a master's degree, she became a teacher in the Department of Computer Science. She can be contacted at email: jinan_redha@uomustansiriyah.edu.iq