©APR 2019 | IRE Journals |CCCA - 2019 | Volume 2 Issue 10 | ISSN: 2456-8880

Secret Key Generation for Privacy Preservation Using Advanced Least Lion Optimisation Algorithm

G.SATHISH KUMAR[1], S.K.VISHAL FELIX[2],S.SATHEESH KUMAR[3], S.MANOJ KUMAR[4]
Department of Computer Science and Engineering, Bannari Amman Institute of Technology Sathyamangalam, Tamil Nadu

Abstract - Nowadays, Data Mining is a popular tool for extracting hidden knowledge from huge amount of data. To discover hidden knowledge in the information without revealing sensitive data is one of the significant difficulties in data mining. There are numerous techniques have been proposed to hide private data. Association rule mining is one of the data mining procedures used to mine secret data from large datasets. Privacy Preserving Data Mining (PPDM) plays a major role in Data Mining. Privacy Preserving Data Mining (PPDM) methods are utilized to protect such classified data from un-approved data set. These investigations propose a procedure for secret key generation using Advanced Least Lion Optimization Algorithm (ALLOA). The proposed calculation includes two phases: rule mining and secret key generation. Initially, whale optimisation algorithm mines the association rules for the input database and validates the rules with the newly formulated fitness function. An algorithm, ALLOA is developed by modifying the lion optimisation algorithm (LOA) with the inclusion of least mean square (LMS) which generates a secret key to provide privacy in mining. With the secret key, ALLOA converts the original database into the sanitized database. Then, the algorithm optimally selects a secret key such that the sanitised database hides sensitive information by the utilisation of two factors, namely, privacy factor and utility factor, in its objective function.

Index Terms -Advanced Least Lion Optimisation Algorithm, Association rule mining, Data Mining, Privacy Preserving Data Mining (PPDM), Secret key generation.

I. INTRODUCTION

Data belongs to a person or an organization may have dissimilar sensitive levels. These data are prepared available only for authorized persons [1]. Thus, ensuring the protection of sensitive data by access restriction is not a complete method. This may affect the utility of the data mining result and with help of the knowledge the user may re-identify sensitive data items from non-sensitive data is known as Inference Problem. The privacy preserving data mining is to provide a resolution for protecting sensitive information by developing a data mining technique which could be applied on databases without affecting the accurateness of data mining result and without violating the privacy of individuals is the motivation for this research

Information mining is the technique for deciding examples in extensive informational collections with man-made reasoning, AI, insights and database frameworks [2]. The point of information mining process is to coerce data from an enormous volume of informational collection to have sensible auxiliary portrayal of the information thing in the value-based database. It is use to mine noteworthy and valuable data or information from substantial database. Secured or secret data separated by information mining strategies prompts the danger of dangers to protection. Affiliation rule mining is a strategy in information mining to recognize the regularities made in extensive volume of information. The method is cooperated by allowing third party to recognize and disclose hidden private information for an individual or an organization.

Privacy-preserving data mining with association rule denotes the area of data mining that looks to save susceptible information from preventable or illegal disclosure. Privacy information comprises private or confidential information in business like social security numbers, home address, credit card numbers, credit ratings, purchasing behaviour, medical records and best-selling services [3]. The privacy preservation data mining requires assurance for hiding of sensitive information in efficient manner. The association rule hiding technique protects the sensitive data ultimately under the scanner. Also, it fails to conceal data items which are not sensitive. It affects the confidentiality of rules and the utility of the data mining results.

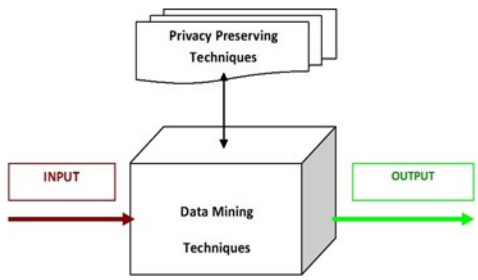

Figure 1: Privacy preserving data mining

The proposed method uses a Advanced Least Lion Optimization Algorithm for association rule hiding. Proposed approach involves two stages: rule mining and secret key generation for the sanitisation. Whale optimisation algorithm (WOA) mines the association rules for the input database provided. It validates the rules mined from the input and the sanitised databases with the newly formulated fitness function.

This paper is sorted out as pursues: Section II contains the related works of essential writing papers, Section III demonstrates the proposed ALLOA technique, Section IV contains the Experimental results and Section V deserves the conclusion. The execution of ALLOA is contrasted and three existing techniques, such as PSO, COA, Firefly, and LOA, to gauge its performance. From the examination, it very well may be demonstrated that the proposed ALLOA technique picks up a most extreme protection.

II. RELATED WORK

Umesh Kumar Sahuet al.[4] proposed a association rule hiding algorithm to preserve sensitive association rules. Heuristic-Based approach and Reconstruction-Based approaches along with cryptographic method were used in association rule hiding. Heuristic-Based approach is employed to find the suitable dataset for transaction. The Reconstruction-based approach is used to rebuild the data from perturbed data.

SaimaKanwalet al.[5] proposed a confident address configuration based authentication protocol to overcome the intermediate attack. These authentication protocols have three characteristics, such as a) Anonymous authentication b) Privacy c) Efficiency. In Anonymous authentication the authentication is provided to the message generator, Privacy deals with the communication content and Efficiency is measured for storage requirement and verification of message. Address Configuration algorithm is used to collect address structure from the network. Address forgery attack and address exhaustion attack were analysed and this proposed algorithm seems to be much efficient.

ShubhraRana, Dr. P. anthiThilagam proposed a novel mechanism for performing PPDARM [6] on horizontally distributed databases. Pattern Count tree structure has been used to improve the scalability of the DARM algorithm as PC tree requires only one scan for construction and provides a compact and complete representation of the database. Paillier cryptosystem used for additive homomorphic properties leaks negligible information about the private data. The HHE scheme enhances the scalability of the PPDARM approach by using a tree aggregation structure which minimizes the number of messages exchanged. The proposed scheme can be extended to be secure under a malicious adversarial assumption. Key generation mechanism can be made more robust by including Zero Knowledge Proof mechanisms and allowing distributed key generation.

Chun-Wei Lin et al.[7] proposed a Genetic Algorithm (GA) centred framework along with two optimization algorithms to preserve the user's privacy. Initial population is encoded with chromosome for performing crossover and mutation to evaluate the fitness function.GA centred framework is established for hiding the individual's private data. These algorithms are used to decrease the rescanning time of actual dataset. Mushroom and BSMWebview data sets were used to perform the research based on the proposed algorithm. This work results in disguising the sensitive data through transaction deletion.

Cheng, P., Roddick, J.F., Chu, S.C., et al. another twisting based technique [8] is proposed which shrouds touchy guidelines by expelling a few things in a database to decrease the keep up or certainty of delicate standards underneath indicated limits. So as to lessen symptoms on data, the data on non-delicate item sets contained by every exchange is utilized to sort the supporting exchanges. The applicants that contain littler amount non-touchy item sets are chosen for change ideally. So as to diminish the twisting degree on information, the base number of exchanges that should be adjusted to hide a delicate guideline is determined. Similar tests on certifiable datasets demonstrated that the new technique can accomplish palatable outcomes with less reactions and information misfortune.

D. Menaga et al. [9] proposed a LLOA for generating the secret key. The secret key is used to preserve the individual's private information. Least mean square method is used for secret key generation in LLOA for transforming the actual database to sanitize database.

III. PROPOSED METHODOLOGY

3.1 Advanced Least Lion Optimisation Algorithm (ALLOA).

This method works in two stages, specifically, association rule mining, and secret key creation for the sanitisation. At first, the proposed technique uses WOA which mines the association rules from the original database using a well-formulated fitness function. In the second phase, LLOA is developed by modifying the update rule of LOA with the utilization of weight update of LMS. This algorithm selects the secret key optimally using two factors, such as privacy factor and utility factor, as two objectives in the fitness function. Then, LLOA converts the original database into the sanitised database by the secret key. After that, the algorithm optimally selects a secret key such that the sanitised database hides sensitive information by the utilization of two factors, namely, privacy factor and utility factor, in its objective function. Thus, this privacy preserving technique improves the search process by the optimal selection of secret key to design the sanitised database and thereby, provide PPDM. The performance of LLOA is compared with three existing techniques, such as PSO, COA, Firefly, and LOA, to estimate its performance. From the analysis, it can be shown that the proposed LLOA technique gains a maximum privacy. The main contributions of the proposed privacy preserving rule hiding algorithm are as follows: • Association rule mining using WOA with the fitness satisfying the support and the confidence thresholds for privacy preservation.

The aim of PPDM techniques is to hide and preserve the confidential data by sanitising the original database. Sanitisation refers to the process of modifying the database so that the third party who receives the database can acquire only the required information. Although, there are [14] numerous methods to sanitise the data, the method of generating key offers better performance than the common sanitization processes of addition and deletion. The proposed approach of LLOA for sanitisation creates the secret key through which it can build the sanitised database along with the utilisation of details regarding the original database, and the mined association rules.

3.1.1 Secret key generation:

It is important that the database provided to the third party about the user must have the sensitive information hidden. This requires a database that should not reveal the confidential information but, does not differ much from the original database. Therefore, the process needs a technique to create a sanitised database by the generation of the secret key in a random manner. LLOA is an effective algorithm that modifies LOA by the integration of LMS into the update equation of LOA.

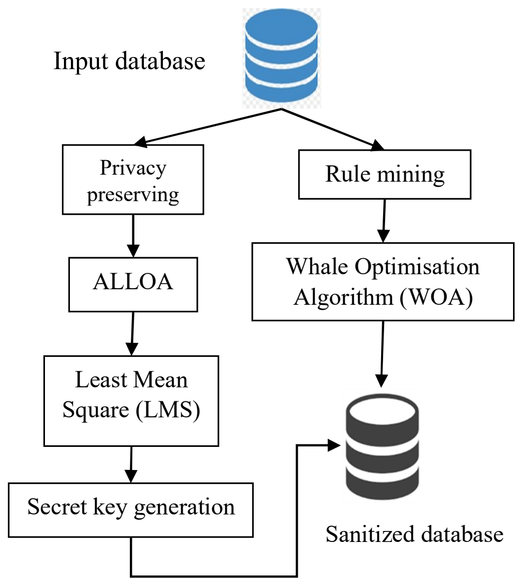

Figure 2: System Model

3.1.2 WOA-based rule mining algorithm:

WOA is an optimisation algorithm inspired by the social behaviour of whales. For the input database, this algorithm performs rule mining such that the privacy is preserved. It creates the model with the utilization of whale behaviour, such as encircling the prey, attacking the prey, and search for prey. With the newly formulated fitness that uses the support and confidence metric, the algorithm executes the rule mining process enhancing the search process.

IV. EXPERIMENTAL RESULTS

The algorithm is employed in PHP and JAVA. The proposed method has the maximum privacy than the existing methods. The proposed method considers both the privacy factor and utility factor in its objective function and improves the search process by the optimal selection of secret key to design the sanitised database and thereby, provide PPDM. The proposed method has the maximum utility than the existing methods by taking the advantages of the WOA and LMS. Also, the proposed algorithm generates the new objective function by considering the privacy factor and the utility factor.

Data

Sno	PatientID	Name	Gender	Age	Disease	Mobile	Address	City
1	PatientID	Name	Gender	Age	Disease	Mobile	Address	City
2	P0001	Priya.S	Female	41	Typhoid	9098311234	11,EC Road	Trichy
3	P0002	Sujan	Male	49	Skin Alergy	9098261254	431/2,MK street	Chennai
4	P0003	Kishnan	Male	37	Typhoid	7098355342	32, NG Nagar	Madurai
5	P0004	Aruna.D	Female	33	Typhoid	8098311256	11A, SS Road	Salem
6	P0005	Logeswaran	Male	27	Angiography	9828311288	MMG Colony	Chennai
7	P0006	Saran	Male	32	Typhoid	9998217236	123, Kk nagar	Covai
8	P0007	Divya.K	Female	59	Fever	9098311234	11,EC Road	Erode
9	P0008	Vinith	Male	30	Eye Surgery	7098355342	21/2,MK street	Karur
10	P0009	Vinoth	Male	38	Bone Fracture	8098311251	56, NG Nagar	Trichy
11	P0010	Thamaraiselvan	Male	36	Eye Surgery	9828311288	23C, FT Road	Covai
12	P0011	Selvan	Male	35	Eye Surgery	9998217234	RT Colony	Madurai
13	P0012	Kokila	Female	29	Skin Alergy	7098355342	123, Kk nagar	Salem
14	P0013	Govind	Male	31	Eye Surgery	8098311256	11,EC Road	Tanjore

Figure3: Original Data

Encrypted Data

Sno	277dea0530861d1be6	397df309	3079f008309a	367bfb	3375ed09349b0c	3a73fc05398d	3678fa1e309b1a	347
1	277dea0530861d1be6	397df309	3079f008309a	367bfb	3375ed09349b0c	3a73fc05398d	3678fa1e309b1a	3475ea
2	272cae5c64	276ef71534c63a	3179f30d398d	432d	2365ee043a810d	4e2ca75466d9586091e5	462db22916c83b3dc3b5	236ef7
3	272cae5c67	2469f40d3b	3a7df209	4325	2477f70275a90537d0b667	4e2ca75467de586097e5	432faf4367c4241982a26a239d9d30	3474fb
4	272cae5c66	3c75ed043b8907	3a7df209	442b	2365ee043a810d	402ca75466dd5c6196e3	442eb24c1baf491cc3b67f23	3a7dfa
5	272cae5c61	366eeb0234c62d	3179f30d398d	442f	2365ee043a810d	4f2ca75466d9586097e7	462ddf4075bb3a72f0be7f35	247df2
6	272cae5c60	3b73f909269f0820c3bf	3a7df209	452b	3672f9053a8f1b33d2b967	4e24ac5466d958609ae9	3a51d94c1687053dcca8	3474fb
7	272cae5c63	247dec0d3b	3a7df209	442e	2365ee043a810d	4e25a75467d95e6091e7	462ead4075a30272ccb079308a	3473e8
8	272cae5c62	3375e81534c622	3179f30d398d	4225	3179e80927	4e2ca75466d9586091e5	462db22916c83b3dc3b5	326ef1
9	272cae5c6d	2175f0052180	3a7df209	442c	3265fb4c069d1b35c7a367	402ca75466dd5c6196e3	452db15e79a52272d1a56c349d8c	3c7dec
10	272cae5c6c	2175f0032180	3a7df209	4424	3573f00975ae1b33c1a56b239d	4f2ca75466d9586097e0	422ab24c1baf491cc3b67f23	236ef7
11	272cae5d65	2374ff01349a083bd1b472279996	3a7df209	442a	3265fb4c069d1b35c7a367	4e24ac5466d958609ae9	452fdd4075ae3d72f0be7f35	3473e8
12	272cae5d64	2479f21a3486	3a7df209	4429	3265fb4c069d1b35c7a367	4e25a75467d95e6091e5	2548be2f3a84063cdb	3a7dfa
13	272cae5d67	3c73f5053989	3179f30d398d	4525	2477f70275a90537d0b667	402ca75466dd5c6196e3	462ead4075a30272ccb079308a	247df2
14	272cae5d66	3073e8053b8c	3a7df209	442d	3265fb4c069d1b35c7a367	4f2ca75466d9586097e7	462db22916c83b3dc3b5	237df0
15	272cae5d61	276efb0134	3a7df209	4428	3265fb4c069d1b35c7a367	4e24ac5466d958609ae9	432faf4367c4241982a26a239d9d30	237df0
16	272cae5d60	3c7df30d39	3a7df209	4524	3265fb4c069d1b35c7a367	4e25a75467d95e6091e5	442eb24c1baf491cc3b67f23	237df0
17	272cae5d63	257df31534	3179f30d398d	4525	2365ee043a810d	4e2ca75466d9586091e5	462ddf4075bb3a72f0be7f35	236ef7
18	272cae5d62	2469ec092680	3a7df209	442e	3d7deb0231810a37	402ca75466dd5c6196e3	3c4cd34c1687053dcca8	3474fb

Figure 4: Encrypted data

Data

Sno	**ti****	Na****	Ge****	Ag****	**se****	Mo****	**dr****	Ci****
1	**ti****	Na****	Ge****	Ag****	**se****	Mo****	**dr****	Ci****
2	P0****	**iy****	Fe****	4*****	**ph****	**98****	**,E****	Tr****
3	P0****	Su****	Ma****	4*****	**in****	**98****	**1/****	**en****
4	P0****	**sh****	Ma****	3*****	**ph****	**98****	**, ****	**du****
5	P0****	**un****	Fe****	3*****	**ph****	**98****	**A,****	Sa****
6	P0****	**ge****	Ma****	2*****	**gi****	**28****	**G ****	**en****
7	P0****	Sa****	Ma****	3*****	**ph****	**98****	**3,****	Co****
8	P0****	**vy****	Fe****	5*****	Fe****	**98****	**,E****	Er****
9	P0****	Vi****	Ma****	3*****	**e ****	**98****	**/2****	Ka****
10	P0****	Vi****	Ma****	3*****	**ne****	**98****	**, ****	Tr****
11	P0****	**am****	Ma****	3*****	**e ****	**28****	**C,****	Co****
12	P0****	Se****	Ma****	3*****	**e ****	**98****	** C****	**du****
13	P0****	Ko****	Fe****	2*****	**in****	**98****	**3,****	Sa****
14	P0****	Go****	Ma****	3*****	**e ****	**98****	**,E****	**nj****
15	P0****	Pr****	Ma****	3*****	**e ****	**28****	**1/****	**nj****
16	P0****	Ka****	Ma****	2*****	**e ****	**98****	**, ****	**nj****
17	P0****	Ra****	Fe****	2*****	**ph****	**98****	**A,****	Tr****
18	P0****	Su****	Ma****	3*****	**un****	**98****	**M ****	**en****
19	P0****	Ga****	Ma****	2*****	**ph****	**98****	**3,****	**du****

Figure 5: Rule hiding

V. CONCLUSION

This study proposes a novel optimisation algorithm, called Advanced Least Lion Optimisation Algorithm (ALLOA) for privacy preserving association rule hiding. The proposed approach involves two stages: rule mining and secret key generation for the sanitisation. Whale optimisation algorithm (WOA) mines the association rules for the input database provided. It validates the rules mined from the input and the sanitised databases with the newly formulated fitness function. An algorithm, LLOA, which modifies lion optimisation algorithm (LOA) with the inclusion of LMS, generates a secret key to provide privacy in mining. With the secret key, LLOA converts the original database into the sanitized database. The algorithm optimally selects a secret key such that the sanitised database hides sensitive information by the utilization of two factors, namely, privacy factor and utility factor, in its objective function. Thus, this privacy preserving technique improves the search process by the optimal selection of secret key to design the sanitised database and thereby, provide PPDM.

REFERENCES

[1] ShabnumRehman and Anil Sharma, “Privacy Preserving Data Mining Using Association Rule Based on Apriori Algorithm”. Springer Nature Singapore Pte Ltd, 2017.

[2] Mistry, B.R., Desai, A.: ‘Privacy preserving heuristic approach for association rule mining in distributed database’. Proc. IEEE Int. Conf. on Innovations in Information, Embedded and Communication Systems (ICIIECS), Coimbatore, India, 2015, pp. 1–7.

[3] Rachit V. Adhvaryu, Nikunj H. Domadiya, “Privacy Preserving in Association Rule Mining On Horizontally Partitioned Database”. International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), Vol. 3, Issue 5, May 2014.

[4] Umesh Kumar Sahu, Anju Singh, “Approaches for Privacy Preserving Data Mining by Various AssociationsRule Hiding Algorithms – A Survey”.International Journal of Computer Applications, 2016.

[5] SaimaKanwal, Chunxue Wu, NaixueXiong ‘An effective and secure user authenticated protocol for location services in road networks', International Journal of Engineering Research & Technology, ISSN: 2278-0181, Vol. 6 Issue 04, April-2017.

[6] ShubhraRana, Dr. P. anthiThilagam, “Hierarchical Homomorphic Encryption based Privacy Preserving Distributed Association Rule Mining”. IEEE 13th International Conference on Information Technology, 2014.
[7] Chun-Wei Lin, Hong, T.P., Yang, K.T., et al.: ‘The GA-based algorithms for optimizing hiding sensitive item sets through transaction deletion’, Appl. Intell., 2015, 42, (2), pp. 210–230.
[8] Cheng, P., Roddick, J.F., Chu, S.C., et al.: ‘Privacy preservation through a greedy, distortion-based rule-hiding method’, Appl. Intell., 2016, 44, (2), pp.295–306.
[9] D. Menaga, S. Revathi , ”Least lion optimisation algorithm (LLOA) based secret key generation for privacy preserving association rule hiding”, The Institution of Engineering and Technology, ISSN 1751-8709,pp.1-9.