International Journal of Computer Science Engineering & Technology
APPROVED BY
NATIONAL SCIENCE LIBRARY (NSL)
NATIONAL INSTITUTE OF SCIENCE COMMUNICATION AND INFORMATION RESOURCES (NISCAIR)
COUNCIL OF SCIENTIFIC AND INDUSTRIAL RESEARCH (CSIR)- NEW DELHI INDIA .
ISSN :2455-9091

A SURVEY ON PRIVACY PRESERVING IN DATA MINING

[1] SUDHA. S, [2] Dr. P. LOGESWARI
[1] Research Scholar, [2] Assistant Professor,
[2] Department of Computer Science,
[1, 2] Sri Krishna Arts & Science College,
[1, 2] Coimbatore, Tamilnadu, India.

ABSTRACT - Privacy has gotten significant in information-based applications. Legitimate reconciliation of individual protection is fundamental for data mining activities. This protection based data mining is significant for areas like Healthcare, Pharmaceuticals, Research, and Security Service Providers, to give some examples. The fundamental arrangement of Privacy Preserving Data Mining (PPDM) strategies falls into Perturbation, Secure Sum Computations and Cryptographic based methods. There exist tradeoffs between security conservation and data misfortune for summed up arrangements. In this paper we survived 15 Literature survey for Privacy Preserving in Data Mining. Ongoing advances in the Internet, in data mining, and in security advances have led to another surge of examination, known as preserving data mining (PPDM). PPDM advances permit us to extricate pertinent information from a lot of information, while conceal touchy information or data from exposure.

Keywords: [Data Mining, Cryptography, Clustering, Decision Tree, Association Rule Mining]

1. INTRODUCTION

Hazardous advancement in systems administration, stockpiling, and processor innovation has prompted the formation of extremely enormous data sets that record uncommon measure of conditional data. The principle issue is that with the accessibility of non-touchy data or unclassified information, one can derive delicate data that should be revealed. Notwithstanding its advantages in different regions like advertising, business, clinical examination, bioinformatics and others, information mining can likewise represent a danger to protection in data set security if not done or utilized appropriately. Protection safeguarding information mining is a clever examination bearing in information mining and factual data sets, where information digging calculations are dissected for the incidental effects, they cause in information security. Privacy preserving data mining (PPDM) has arisen to resolve this issue. The majority of the methods for PPDM utilize changed form of standard information mining calculations, where the alterations as a rule utilizing notable cryptographic strategies guarantee the necessary protection for the application for which the procedure was planned. By and large, the requirements for PPDM are saving precision of the information and the produced models and the exhibition of the mining cycle while keeping up with the protection imperatives. The few methodologies utilized by PPDM can be summed up as beneath:

The data is altered before delivering it to the data miner.

The data is distributed between two or more sites, which cooperate using a semi-honest protocol to learn global data mining results without revealing any information about the data at their individual sites.

While using a model to classify data, the classification results are only revealed to the designated party, who does not learn anything else other that the classification results, but can check for presence of certain rules without revealing the rules.

Many methodologies arose for security safeguarding information mining. The primary methodology included bothering the contribution prior to mining. However it has the advantage of effortlessness it doesn't give a conventional system to demonstrating how much protection is ensured. Secure Computation procedure enjoys the benefit of giving a distinct model to protection utilizing cryptographic strategies and is additionally precise.

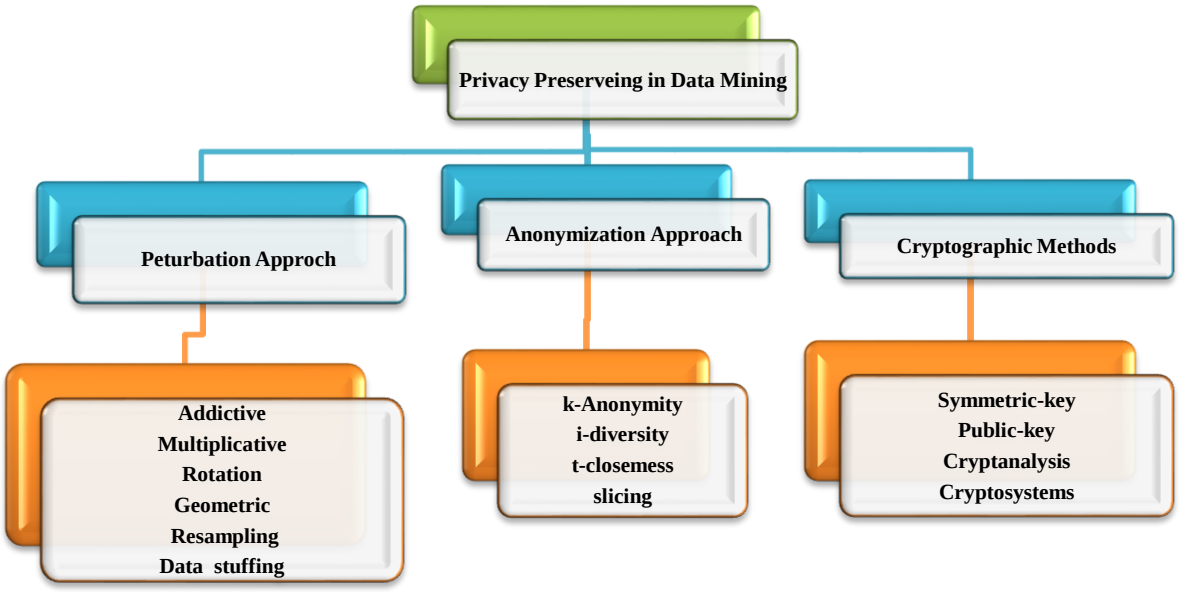

Figure1. Addictive Data Perturbation approach for privacy preserving in data mining

Data privacy and information utility are generally considered as a couple of clashing necessities in security saving information mining frameworks and applications. Multiplicative perturbation algorithm target further developing information protection while keeping up with the ideal degree of information utility by specifically safeguarding the mining errand and model explicit data during the information annoyance measure. By safeguarding the errand and model explicit data, a bunch of "change invariant information mining models" can be applied to the bothered information straightforwardly, accomplishing the necessary model precision. Frequently a multiplicative annoyance calculation might discover various information changes that protect the necessary information utility.

2. LITERATURE SURVEY

1. Kaur and S. Sofat (2016) et.al proposed A Proposed Hybrid Approach for Privacy Preserving Data Mining. Every single organization needs to gather the data of its clients or users for either reason. This data can be put away on a unified server or on the cloud. In the event that the data is being put away on the brought together storehouse then the control of the data is with the archive controllers of the organization. With the

developing technology and measure of data, the utilization of cloud and brought together servers is expanding. At the point when the data of the people is put away by the outsider it prompts the frailty of abuse of their information in the people. Data Mining is carried on this data put away over cloud to get valuable information, examination and decision making purposes. In any case, alongside the way toward mining, privacy preservation is the key concern. Because of expansion in the data stockpiling over the cloud, the need of utilizing the data mining techniques for Privacy Preservation is expanding quickly. A great deal of exploration has been done in the field of anonymization and cryptographic techniques. Anonymization and perturbation techniques can be viewed as better when contrasted with cryptographic techniques as far as complexity and proficiency for huge number of users. At the point when looked at based on information misfortune and privacy accomplished anonymization experiences an essentially high information misfortune. The proposed hybrid strategy can successfully accomplish the objective of privacy preservation with no information misfortune as the utilizing the algorithm the distorted values can be returned to its unique values successfully.

2. A. W. Putri and L. Hira (2016) et.al, proposed Hybrid Transformation In Privacy-Preserving Data Mining. The data gave wherever even without us knowing. In organization, data from any source will be put away in data warehouse and information can be acquired from it straightforwardly or can be processed utilizing data mining. Data mining in some cases may inadvertently divulge such information since certain pieces of data are utilized. To keep up with the privacy of information (privacy-preserving) of the data, it's necessary unique techniques to conceal the real value of data so that can restrict acquiring information from it. However, in some cases concealing real value can cause misfortune some information along these lines make the acquired information isn't altogether right. Procedure to concealing real value while keeping same information called privacy preserving data mining (PPDM). In this paper, we proposed another change method in privacy-preserving data mining by joining two past methods: entropy-based parcel and consolidated data bending called hybrid change. Through tests, we exhibited that the proposed method can work better compared to past research, consequently can adequately give a balance data utilities and data privacy better. This method can be utilized when data will be utilized in outer organization asset, so the privacy of the data will be kept from outcast. This method additionally can be utilized in collaborative examination inside organization utilizing similar data for different uses.

3. S. Liu, Q. Qu, L. Chen and L. M. Ni (2015) et.al proposed SMC: A Practical Schema for Privacy-Preserved Data Sharing over Distributed Data Streams. Gathering data from appropriated data providers is a provoking errand to enormous data related examination networks and businesses. In a conveyed environment, data are put away at every nearby site. A data demander frequently requires the total data from every one of the conveyed locales to lead significant data examination and mining. Because of the privacy worries on the data at every nearby site, more often than not it is amazingly hard for the data demander to secure total and exact data from every one of the individual locales. Data assortment is needed to be protected and efficient thinking about the two data privacy and framework execution. In this paper, we study another issue: dispersed data imparting to privacy-preserving prerequisites. Given a data demander mentioning data from numerous disseminated data providers, the goal is to empower the data demander to get to the dispersed data without knowing the privacy of any individual provider. The issue is tested by two inquiries: how to communicate the data securely and precisely. In this paper, a privacy-protected data sharing issue is concentrated with regards to real-life

conveyed mobile phone organizations. We figure the issue as a conveyed data sharing issue, and we propose a shadow coding method with shadow matrix computation, which is privacy-preserving, efficient, and data-recoverable. A generalization of this method is additionally examined in the paper. To assess the proposed methods, we lead the examinations with an enormous scope real-life datasets. The proposed schema is additionally carried out as a pilot framework in a city to gather appropriated mobile phone data.

4. V. Baby and N. S. Chandra (2016) et.al proposed Distributed Threshold K-Means Clustering for Privacy Preserving Data Mining. Privacy preserving is significant in wherein data mining transforms into a helpful task among members. In data mining, a champion among the most proficient and frequently used frameworks is k-means clustering. In this paper, we propose an efficient circulated threshold privacy-preserving k means clustering algorithm that utilization the code based threshold secret sharing as a privacy-preserving component. Development includes code based methodology which permits the data to be partitioned into numerous offers and processed independently at various servers. Our convention takes less number of emphases contrast and existing protocols and it don't need any trust among the servers or users. We additionally give explore results correlation and security investigation of the proposed plot. In this work, we propose a circulated threshold privacy preserving k-means clustering algorithm that utilization the code based threshold secret sharing plan and secure expansion and correlation protocols. We permit gatherings to collaboratively perform clustering and accordingly keeping away from trusted outsider. We contrast our convention and CRT based clustering proposed in. Our algorithm doesn't need any trust among the servers or users and it give amazing privacy preserving of client data.

5. R. S. Mohammed, E. M. Hussien and J. R. Mutter (2016) et.al, proposed A Novel Technique of Privacy Preserving Association Rule Mining. Privacy Preserving Association Rule Mining (PPAM) turns into a significant issue lately. Since data mining alone isn't sufficient to divide data among organizations without privacy preserving. In this paper, another procedure has been proposed to keep up with the classification of the data by creating of association rule utilizing a stochastic standard guide without getting back to mining sensitive data once more. The framework reproduction utilizing Mat lab and tried that shows the effective contrast between the first data and fabricated. And furthermore been accomplished fast and less memory prerequisites. Since a past works need a further report no time like the present efficient, data size, data form, really concealing association rule. This paper proposes an original method to defeat the issues of PPDM and PPAM techniques in a specialist. The proposed strategy utilizes a stochastic Standard guide to fabricate association rules straightforwardly without a re-visitation of mining again of a unique sensitive dataset. So this strategy has great properties in a period efficient. Likewise, it can attempt to any size and sort of data. It permits the individual select their security level in simple. And furthermore fabricate a sensitive data and association rule in simple.

6. P. S. Wang, F. Lai, H. Hsiao and J. Wu (2016) et.al proposed Insider Collusion Attack On Privacy-Preserving Kernel-Based Data Mining Systems. In this paper, we consider another insider danger for the privacy preserving work of circulated kernel-based data mining (DKBDM, for example, dispersed help vector machine. In this paper, we propose an insider collusion attack that is a danger to most data mining systems that work on kernels and talk about the number of insiders are adequate to dispatch this sort of attack. We additionally present two privacy-preserving methods to protect against the attack. At last, exploratory outcomes are given to demonstrate the viability of the proposed attack and guard plans. Note that our proposed

attack conspire isn't simply material to the in an upward direction partitioned data yet in addition relevant to horizontally partitioned data and discretionarily partitioned data; as long as each kernel value is made out of two data vectors and put away in a kernel matrix, our proposed method can turn around those kernel values back to the first data. Indeed, most data mining systems working on kernel computation particularly those in a circulated environment are potential victims of the proposed attack. Later on work, we will talk about whether the privacy break rule depicted can be loose, to such an extent that despite the fact that the specific recuperation is unimaginable, however the attacker can distinguish the subspace of the private information (Corresponding an excessive number of answers for the arrangement of linear equations). Under the present circumstance, how to assess the security level of the framework the extent of this paper is restricted to non-homomorphic encryption methods. The primary justification this is that, as of not long ago, homomorphic encryption systems have been too delayed to ever be functional. In any case, in the event that we consider homomorphic encryption based systems such; we accept that the proposed insider dangers could prompt a known-plaintext attack.

7. M. Chaudhari and J. Varmora (2016) et.al proposed Advance Privacy Preserving in Association Rule Mining. Data mining is the way toward getting information from enormous database. In data mining, vital themes in research community is Privacy preserving data mining (PPDM). Privacy preserving in data mining has become more mainstream since it permits privacy of sharing sensitive data to others for investigation purposes. It is fundamental to keep a proportion between privacy insurance and information revelation. Taking care of such issues there are a few algorithms which are introduced by different creators around the world. In this paper, we attempt to portray what privacy is preserving in association rule mining and ADSRRC algorithm. Privacy preserving Association rule mining guarantees that mining interaction won't unveil the sensitive information of any including party. We proposed an algorithm named ADSRRC which conceals sensitive association rules with fewer changes on database to keep up with data quality and to diminish the result of database. We are attempting to carry out it. In future, ADSRRC algorithm can be stretched out to expand the proficiency and lessen the incidental effects by limiting the changes on database.

8. R. K. Dhandhania, P. K. Baruah and R. Mukkamala (2014) et.al proposed Privacy-Preserving Mining of Decision Trees Using Data Negation Approach. Preserving privacy of data has now gotten extremely essential for organizations that gather and keep up with the data. Such data incorporates client data, patient health records, personal duty records, and so on Likewise, with the high value inferred through data mining exercises, for example, arrangement and association rule mining, particularly over data kept up with by various organizations, sharing of data across organizations has additionally become a reality. With the consistently expanding need to share data across organizations, the interest for its privacy is additionally turning out to be progressively significant. Previously, a few techniques have been proposed for privacy-preserving data mining. In this paper, we propose a privacy-preserving strategy to fabricate decision trees. We allude to it as invalidation strategy. Since our strategy doesn't utilize any cryptographic techniques, it is computationally efficient. In any case, such sharing ought to be refined remembering the data privacy. Privacy is the core of numerous applications, particularly in the current web environment. Indeed, even with cryptography there is risk of information spillage through right outcomes. Along these lines data disinfection turns out to be vital to keep the data hidden while keeping up with its utility. The techniques introduced in this proposition won't work if all the datasets are spilled on the

grounds that the data recreation algorithm is nonexclusive. Subsequently, further examination should be taken up alongside cryptographic models, to take out this impediment. Cryptographic apparatuses alongside the strategy introduced in this proposal can be substantially more privacy preserving and computationally feasible.

9. Sharma, S. and Shukla, D. (2016), et.al proposedefficient multi-party privacy preserving data mining for vertically partitioned data. The data in computational space put away in advanced arrangement. This organization of data devours less exertion and capacity. Along these lines various association and foundations are preserving their information in this arrangement. In the proposed work an association is viewed as where the decisions are made with the diverse office based data and their characteristics. Furthermore to make decisions the traits of the relative multitude of divisions are required. Be that as it may, the offices can't unveil the privacy of data proprietor. In this paper they show how the various divisions of same association join their data without hurting the privacy of the customer for making powerful decisions in productive and exact way. Along these lines the strategy in an upward direction data blend, cryptography and decision mining is illustrated. To mine the decisions from the data a C4.5 decision tree is utilized. The execution of the proposed privacy preserving data mining and decision making technique is performed utilizing JAVA innovation. Moreover the presentation of the framework is figured as far as precision, blender rate, memory utilization and time utilization. At long last to legitimize the results of the proposed data mining technique the conventional J4.5 tree utilizing WEKA device is utilized with same data for relative execution study. The proposed work is planned to discover the answer for privacy preserving data mining technique in effective and precise way. The worker end creates the irregular cryptographic key. This key is utilized to scramble or encode the data before the mining of data. The scrambled data is utilized as the images for an unknown dataset and utilizing the pre-arranged new encoded data the mining is performed for decision making. The mining of encoded data is performed with the assistance of C4.5 decision tree algorithm.

10. Vinay, M. G. and Ravi Kumar, V. G. (2017), et.al proposed A New Model for Privacy Preserving Multiparty Collaborative Data Mining. The data in computational area put away in computerized design. This configuration of data burns-through less exertion and capacity. Hence various association and establishments are preserving their information in this arrangement. In the proposed work an association is viewed as where the decisions are made with the diverse office based data and their traits. Moreover to make decisions the traits of the multitude of divisions are required. However, the divisions can't uncover the privacy of data proprietor. In this paper they show how the various branches of same association consolidate their data without hurting the privacy of the customer for making viable decisions in productive and exact way. Consequently the technique in an upward direction data blend, cryptography and decision mining is illustrated. To mine the decisions from the data a C4.5 decision tree is utilized. The execution of the proposed privacy preserving data mining and decision making technique is performed utilizing JAVA innovation. Furthermore the exhibition of the framework is figured as far as exactness, mistake rate, memory utilization and time utilization. At last to legitimize the results of the proposed data mining technique the customary J4.5 tree utilizing WEKA device is utilized with same data for near execution study. The proposed work is expected to discover the answer for privacy preserving data mining technique in productive and exact way. The worker end creates the arbitrary cryptographic key. This key is utilized to encrypt or encode the data before the mining of data. The Encrypted data is utilized as the images for an unknown

dataset and utilizing the pre-arranged new encoded data the mining is performed for decision making. The mining of encoded data is performed with the assistance of C4.5 decision tree algorithm.

11. Kaur, A., and Sofat, S. (2016), et.al proposed a proposed hybrid approach for privacy preserving data mining. With the developing innovation and measure of data, the utilization of cloud and concentrated workers is expanding. At the point when the data of the people is put away by the outsider it prompts the uncertainty of abuse of their information in the people. Data Mining is carried on this data put away over cloud to get helpful information, examination and decision making purposes. However, alongside the way toward mining, privacy safeguarding is the key concern. While performing data mining over huge arrangement of data there are odds of privacy infringement, loss of information, loss of privacy of people. While focusing just on preserving privacy it can expand shortcoming and intricacy of data mining techniques. Techniques utilized for accomplishing privacy lead to a huge information loss. The strategies worried about information loss for the most part lack behind in accomplishing a decent degree of privacy. Accordingly track down the proficient data mining technique which can save privacy without causing high loss of information. This paper investigations different privacy preserving data mining techniques for their advantages and disadvantages and afterward gives a proposed idea to accomplish privacy with least information loss. Anonymization and perturbation techniques can be viewed as better when contrasted with cryptographic techniques as far as intricacy and productivity for enormous number of clients. At the point when thought about based on information loss and privacy accomplished anonymization experiences an essentially high information loss. In randomization, records in the data are randomized and rearranged so that the mean, middle of the data stays unaltered. Perturbation is utilized for PPDM by adding or increasing some commotion to twist the upsides of data. The proposed technique centers around concealing the information one can get from semi identifiers like postal district, age, ethnicity, gender, conjugal status, and so forth and preserving the character of the people. Semi identifiers itself don't uncover the character of the people however can give sufficient information about the individual to link for certain different data and know the personality of the individual.

12. Baby, V., and Chandra, N. S. (2016), et.al proposed Distributed threshold k-means clustering for privacy preserving data mining. Privacy preserving is significant in wherein data mining transforms into an agreeable task among individuals. In data mining, a champion among the most competent and regularly used frameworks is k-means clustering. In this paper, they proposed a productive disseminated threshold privacy-preserving k-means clustering algorithm that utilization the code based threshold secret sharing as a privacy-preserving instrument. Development includes code based methodology which permits the data to be isolated into numerous offers and prepared independently at various workers. The convention takes less number of emphases contrast and existing conventions and it don't need any trust among the workers or clients. The examination results with correlation and security investigation of the proposed plot. In this work, they propose a disseminated threshold privacy preserving k-means clustering algorithm that utilization the code based threshold secret sharing plan and secure expansion and correlation conventions. They permitted gatherings to collaboratively perform clustering and hence keeping away from confided in outsider. They contrast the convention and CRT based clustering. The algorithm doesn't need any trust among the workers or clients and it gives wonderful privacy preserving of client data. The privacy of data is reasonably characterized as the fitting utilization of data. Getting delicate data is normally known as data security and

generally alluded to as the accessibility, privacy and respectability of data. Data security ensures that the data is right, trustworthy and open when those with allowed admittance require it. Associations need to support an approach of data security for the single motivation behind guarantying data privacy or the privacy of their purchasers 'data, especially when it is being used. One procedure for securing the privacy of the individual records is to annoy the first data. Data perturbation systems are genuinely based techniques that attempt to guarantee restricted information by adding arbitrary commotion to private, mathematical traits, along these lines protecting the first data.

13. Mohammed, R. S., Hussien, E. M., and Mutter, J. R. (2016), et.al proposed a novel technique of privacy preserving association rule mining. Privacy Preserving Association Rule Mining (PPAM) turns into a significant issue as of late. Since data mining alone isn't sufficient to divide data among organizations without privacy preserving. In this paper, another technique has been proposed to keep up with the privacy of the data by manufacturing of association rule utilizing a stochastic standard map without getting back to mining touchy data once more. The framework reproduction utilizing MatLab and tried that shows the effective contrast between the first data and manufactured. And furthermore been accomplished high velocity and less memory prerequisites. This paper proposes an original technique to conquer the issues of PPDM and PPAM techniques in a trained professional. The proposed technique utilizes a stochastic Standard map to manufacture association rules straightforwardly without a re-visitation of mining again of a unique delicate dataset. So this technique has great properties in a period productive. Likewise, it can attempt to any measure and kind of data. It permits the individual select their security level in simple. And furthermore manufacture a touchy data and association rule in simple. This paper proposed another technique for PPAM by straightforwardly treating with association rules to give a quick and can handle any size and type of data dependent on altering of a standard map for privacy preserving applications. Heuristic, Reconstruction and Cryptography are generally recently proposed techniques give the objective of privacy protection; however it isn't adaptable and restricted to a specific structure and size data. By and large, these techniques can not treat with association rules straightforwardly however it got back to a data set to alter touchy information to get Association rule covering up. For additional, since they do data mining to get association rules and afterward it is stowing away by mining again of touchy information dependent on the acquired delicate rules to get a manufacture association rules. So this technique has not adaptability execution with the quick, data size, and data design.

14. Kalaiselvi, K., and Sara, V. J. B. (2017), et.al proposed Privacy preserving in data mining classification and visualization. In present Era, knowledge procurement and separating expected information in any field is exceptionally unpredictable and voluminous. Information rich Data mining is data preparing utilizing complex data search capacities and factual algorithms to find patterns and relationship in enormous prior database. The frequent pattern upholds planning building block and to discover associations rules and brief connection between databases. Most significant difficulties in the framework of the questionable database are the way to deal with the security and doled out likelihood an incentive for every item. Association Rule Mining (ARM) is an imperative data mining procedure that targets removing fascinating relationships, frequent patterns, association rules. It lessens the intricacy of moving forward measure. To satisfy the hole between existing methodology and advancing application prerequisite, an itemized logical examination was led. This investigation assesses the general work which is identified with the privacy of data classification and

visualization and frequent thing pattern. This paper gives a close to examination of different security based data mining algorithms for mining the frequent pattern from vague information with various algorithms. The key hardships in this setting of a problematic database are the means by which to security and likelihood designated worth for every item. It keeps away from intricacy applicant age measure. The paper assesses the general all work which is identified with the privacy of data classification and visualization and frequent thing pattern in privacy-preserving the information modified by changing their state. The design of yield will be same as of uncertain data in the circumstance. The essential objective of this paper is to plan successful framework for recovering and mining the data securely and proficiently. Here, the exact investigation needs to concentrate to carry productive way to deal with satisfy the hole between current methodologies and data mining application prerequisite.

15. Wang, S., Sinnott, R., and Nepal, S. (2017) et.al proposed Privacy-protected place of activity mining on big location data. Individuals consistently invest their energy at a couple of significant locations for different exercises in bunches during explicit time allotments, called spot of movement (POA), e.g., resting at home among relatives during night and working at office among partners during work time. Deducing such places is huge for not just the exact publicizing on the business perspective however the distinguishing rallies or gatherings among a gathering of individuals and tracking of the objective people on the part of public security, e.g., finding and tracking presumed fear mongers for hostile to psychological militant work. In any case, it is a test to map from big location data to spots of movement because of the volume and intricacy while leading to privacy concerns, e.g., actually significant spot mining. In the paper, a technique for POA mining on big location data is proposed, named PPAM, focusing on big data examination and privacy concerns. Further, powerful privacy-preserving instruments under differential privacy are implanted into clustering results and location entropy assessment that admittance to crude location data. They exhibit the utility of the proposed approach with huge scope location datasets got from geo-referred to online media. The test results recommend that the POA mining approach can effectively scale to big data situations while preserving singular client privacy. A whenever Omega-Cluster clustering is executed to deal with the time intricacy and boundary affectability issues of existing location clustering approaches while the fleeting elements include being thought of. A clustering result spatial perturbation is applied to secure the privacy of clustering brings about a differential privacy preserving way. Further, a private LEbased POA recognizing algorithm is proposed to distinguish POA dependent on location entropy with privacy preserving under differential privacy. To decrease the contortion of normal private LE esteem computation, a compelling truncation component is execute to diminish the affectability bound and works on the general utility. C2 means to distinguish POA from the significant locales from C1. To assess the meeting variety of each significant area for ubiquity assessment, the location entropy esteem is determined.

3. PROPOSED METHODS, MERITS AND DEMERITS

Authors Name	Proposed Methods	Merits	Demerits
A. Kaur and S. Sofat (2016)	Hybrid Approach for Privacy Preserving Data Mining	The proposed hybrid method can successfully accomplish the objective	Cryptographic techniques as far as complexity and not

		of privacy preservation with no information misfortune as the utilizing the algorithm the distorted values can be returned to its unique values successfully.	effectiveness for huge number of users.
A. W. Putri and L. Hira (2016)	Hybrid Transformation In Privacy-Preserving Data Mining	1. This method likewise can be utilized in collaborative investigation inside organization utilizing similar data for different uses. 2. Effectively give a balance data utilities and data privacy better.	Consider to add different sensitive attributes, randomize the changed data and apply data quality evaluation to try not to uncover the first value of data.
S. Liu, Q. Qu, L. Chen and L. M. Ni (2015)	SMC: A Practical Schema for Privacy-Preserved Data Sharing over Distributed Data Streams	Implemented as a pilot framework in a city to gather conveyed mobile phone data.	Attack models and address the conveyed data sharing issue in an asynchronous circulated environment
V. Baby and N. S. Chandra (2016)	Distributed Threshold K-Means Clustering for Privacy Preserving Data Mining	Our algorithm doesn't need any trust among the servers or users and it give wonderful privacy preserving of client data.	The privacy issues in distributed data-mining and privacy-preserving data.
R. S. Mohammed, E. M. Hussien and J. R. Mutter (2016)	A Novel Technique of Privacy Preserving Association Rule Mining	Technique has good properties in a time efficient. Also, it can try to any size and type of data. Selected Item can be suppressed.	Privacy of all specified item is not sufficient, as some information remain even after perturbation.
P. S. Wang, F. Lai, H. Hsiao and J. Wu (2016)	Insider Collusion Attack On Privacy-Preserving Kernel-Based Data Mining Systems	Most data mining systems working on kernel computation particularly those in a circulated environment are potential victims of the proposed attack.	The privacy break rule can be loose, with the end goal that despite the fact that the specific recuperation is beyond the realm of imagination.
M. Chaudhari and J. Varmora (2016)	Advance Privacy Preserving in Association Rule Mining	ADSRRC which hides sensitive association rules with fewer changes on database to maintain data quality and to reduce the	ADSRRC algorithm cannot be extended to increase the efficiency and reduce the side

		side effect of database.	effects by minimizing the modifications on database.
R. K. Dhandhania, P. K. Baruah and R. Mukkamala (2014)	Privacy-Preserving Mining of Decision Trees Using Data Negation Approach	Cryptographic devices alongside the strategy introduced in this proposal can be considerably more privacy preserving and computationally feasible.	Privacy of all specified item is not sufficient, as some information remains.
Sharma, S. and Shukla, D. (2016)	Efficient multi-party privacy preserving data mining for vertically partitioned data	1. The effective performance and security in the given privacy preserving technique is improved. 2. The technique is able to combine multiparty vertically partitioned data securely.	This technique is not extendible for hierarchical data mining.
Vinay, M. G. and Ravi Kumar, V. G. (2017)	A New Model for Privacy Preserving Multiparty Collaborative Data Mining	This approach has the benefit of being exact when an ideal solution can be identified. This technique overcomes the security issues.	Thus, Apriori algorithm is used the major drawback with Apriori algorithm is of time and space.
Kaur, A., and Sofat, S. (2016)	Hybrid approach for privacy preserving data mining	The cryptography based approaches achieve the privacy preservation goal and Homomorphic encryption overcomes the information loss drawback.	The working of this model is complex and time consuming.
Baby, V., and Chandra, N. S. (2016)	Distributed threshold k-means clustering for privacy preserving data mining	This algorithm does not require any trust among the servers or users and it provide perfect privacy preserving of user data.	As this used clustering algorithm, accuracy depends on the quality of the data.
Mohammed, R. S., Hussien, E. M., and Mutter, J. R. (2016)	A novel technique of privacy preserving association rule mining	1. This technique has a good property in time efficient 2. It can try to any size and type of data.	Due to frequent updates, the steps taken towards the minima are very noisy
Kalaiselvi, K., and Sara, V. J. B. (2017)	Privacy preserving in data mining classification and visualization	This provides Cross Platform Performance Comparisons	1. It cannot visualize a large amount of data in the web environment.
Wang, S.,	Privacy-protected	This work lays the	We cannot extend P-

Sinnott, R., and Nepal, S. (2017)	place of activity mining on big location data	foundation for robust privacy models used for big location data analytics.	PAM into more comprehensive and practical cases.
Samanthula, B. K. (2017)	Privacy-preserving outsourced collaborative frequent itemset mining in the cloud	This protocol ensures that the entire frequent itemset mining task is performed on the cloud-side, thereby fully utilizing the cloud computing capabilities to handle big data needs and also eliminating the user's need for staying online during the mining process.	However, additional modifications are needed to our protocol to achieve the scalability needs of big data.

CONCLUSION

Privacy preserving is an exceptionally gigantic field. In any case, there is no any single strategy that is reliable in all spaces. Every method has a few constraints and drawbacks. All strategies act in an alternate manner relying upon the sort of information just as the kind of utilization or area. Alongside detriments, every procedure enjoys a few benefits. This survey assists us with understanding the ideas and procedures in security safeguarding in information mining. Moreover, it assists us with investigating the upsides and downsides, all things considered. This backing us to have a wide information on the methods and ideas Data mining is vital apparatus utilized by associations for offering better support, accomplishing more prominent benefit, and better dynamic.

REFERENCES

[1]. R. S. Mohammed, E. M. Hussien And J. R. Mutter, "A Novel Technique Of Privacy Preserving Association Rule Mining," 2016 Al-Sadeq International Conference On Multidisciplinary In It And Communication Science And Applications (Aic-Mitcsa), 2016, Pp. 1-6, Doi: 10.1109/Aic-Mitcsa.2016.7759930.

[2]. P. S. Wang, F. Lai, H. Hsiao And J. Wu, "Insider Collusion Attack On Privacy-Preserving Kernel-Based Data Mining Systems," In IEEE Access, Vol. 4, Pp. 2244-2255, 2016, Doi: 10.1109/Access.2016.2561019.

[3]. M. Chaudhari And J. Varmora, "Advance Privacy Preserving In Association Rule Mining," 2016 International Conference On Electrical, Electronics, And Optimization Techniques (Iceeot), 2016, Pp. 2527-2530, Doi: 10.1109/Iceeot.2016.7755148.

[4]. R. K. Dhandhania, P. K. Baruah and R. Mukkamala, "Privacy-Preserving Mining of Decision Trees Using Data Negation Approach," 2014 International Conference on Contemporary Computing and Informatics (Ic3i), 2014, Pp. 43-48, Doi: 10.1109/Ic3i.2014.7019599.

[5]. P. G. Shynu, H. Md. Shayan. and C. L. Chowdhary, "A Fuzzy based Data Perturbation Technique for Privacy Preserved Data Mining," 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), 2020, pp. 1-4, doi: 10.1109/ic-ETITE47903.2020.244.

[6]. W. Ouyang and Q. Huang, "A Privacy Preserving Algorithm for Mining Rare Association Rules by Homomorphic Encryption," 2019 6th International Conference on Systems and Informatics (ICSAI), 2019, pp. 1403-1407, doi: 10.1109/ICSAI48974.2019.9010219.

[7]. M. M. Hasan, S. Hossain, M. K. Paul and A. H. M. S. Sattar, "A New Hybrid Approach For Privacy Preserving Data Mining Using Matrix Decomposition Technique," 2019 4th

International Conference on Electrical Information and Communication Technology (EICT), 2019, pp. 1-5, doi: 10.1109/EICT48899.2019.9068789.
[8]. S. G. Teo, J. Cao and V. C. S. Lee, "DAG: A General Model for Privacy-Preserving Data Mining: (Extended Abstract)," 2020 IEEE 36th International Conference on Data Engineering (ICDE), 2020, pp. 2018-2019, doi: 10.1109/ICDE48307.2020.00228.
[9]. A. Afrin, M. K. Paul and A. H. M. S. Sattar, "Privacy Preserving Data Mining Using Non-Negative Matrix Factorization and Singular Value Decomposition," 2019 4th International Conference on Electrical Information and Communication Technology (EICT), 2019, pp. 1-6, doi: 10.1109/EICT48899.2019.9068846.
[10]. L. Zhang, W. Wang and Y. Zhang, "Privacy Preserving Association Rule Mining: Taxonomy, Techniques, and Metrics," in IEEE Access, vol. 7, pp. 45032-45047, 2019, doi: 10.1109/ACCESS.2019.2908452.
[11]. S. Qiu, B. Wang, M. Li, J. Liu and Y. Shi, "Toward Practical Privacy-Preserving Frequent Itemset Mining on Encrypted Cloud Data," in IEEE Transactions on Cloud Computing, vol. 8, no. 1, pp. 312-323, 1 Jan.-March 2020, doi: 10.1109/TCC.2017.2739146.
[12]. I. Anikin and R. Gazimov, "Approach to Privacy Preserved Data Mining in Distributed Systems," 2019 International Russian Automation Conference (RusAutoCon), 2019, pp. 1-5, doi: 10.1109/RUSAUTOCON.2019.8867812.
[13]. S. Yaji and B. Neelima, "Optimizing Privacy-Preserving Data Mining Model in Multivariate Datasets," 2019 PhD Colloquium on Ethically Driven Innovation and Technology for Society (PhD EDITS), 2019, pp. 1-3, doi: 10.1109/PhDEDITS47523.2019.8986965.
[14]. Y. Zhou, Y. Tian, F. Liu, J. Liu and Y. Zhu, "Privacy Preserving Distributed Data Mining Based on Secure Multi-party Computation," 2019 IEEE 11th International Conference on Advanced Infocomm Technology (ICAIT), 2019, pp. 173-178, doi: 10.1109/ICAIT.2019.8935900.
[15]. S. G. Teo, J. Cao and V. C. S. Lee, "DAG: A General Model for Privacy-Preserving Data Mining," in IEEE Transactions on Knowledge and Data Engineering, vol. 32, no. 1, pp. 40-53, 1 Jan. 2020, doi: 10.1109/TKDE.2018.2880743.
[16]. M. Rathi and A. Rajavat, "High Dimensional Data Processing in Privacy Preserving Data Mining," 2020 IEEE 9th International Conference on Communication Systems and Network Technologies (CSNT), 2020, pp. 212-217, doi: 10.1109/CSNT48778.2020.9115771.