Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 05, May - 2017

Privacy Preservation in Data Centric Environment: Analysis and Segregation

Jyotir Moy Chatterjee
Assistant Professor, CSE department
GD RCET, Bhilai

***Abstract*-The electronic advances are picking up ubiquity every day and the same number of associations keeping up an extensive volume of information about people or different organizations under cloud computing environment that may conveys touchy data that can't be uncovered to general society or specialists. Accumulation of advanced information data by different governments, corporates, and people has made expansive extension of the data and learning based on choice making. The focus should be on to release the aggregate information about the data, however without discharging singular information about the individuals. Therefore, data privacy preservation is an important consideration. This work includes the study of some dangers and privacy preserving procedures.**

Keyword: Cloud computing environment, choice making, data privacy preservation, Dangers and privacy preserving procedures.

I. INTRODUCTION

Privacy is the capacity of an individual to isolate themselves and consequently convey what needs be specifically. The limits and substance of that may be viewed as private contrast among societies and people, yet share regular topics. At the point when something is private to a man, it more often than not implies that something is naturally extraordinary or delicate to them. Privacy preservation means that one is free from all impedance i.e. privacy protection control permits a man to keep up a level of closeness. Privacy is insurance for the honest utilization of one's close to home information of cloud client. Privacy

ruptures may make genuine inconveniences to the cloud clients. According to the AICPA (American Institute of Certified Public Accountants) and CICA (Canadian Institute of Charted Accountants) privacy is defined as, "It may be termed as the right and obligation of individuals as well as organizations with due respect to the collection, use, retention, and disclosure of someone's personal data"] [1]. Parties of computerized data for different corporates, governments, and people have shaped enormous degrees for learning based on decision making. Driven by regular preferences, or

regulations that require definite data to be mined and appropriated, there is an enthusiasm for the exchange and generation of data among distinctive parties. The privacy preservation is used in many fields say example cloud environment [1], distributed environments [2] and web services [3], etc. By far most of the organizations in the cloud environment, e.g., human services administrations, web advancing, back and saving money and web organizing generally depend on upon the usage of individual information. Cloud computing nowadays a standout amongst the most sizzling themes in data development industry (IT). Since out sourcing of all principal data is open with untouchable, there's consistently having stress of cloud organization supplier's trust-esteem. In light of the data security insurance, it is major for customers to scramble their unstable or private information before securing the information into the cloud. There subsist a couple flaws on the condition of the routine encipher. Right where a puzzle key proprietor needs hunt down some information that is secured on the circulated stockpiling, he might be relied upon to download the encoded information from the Cloud-Servers, then after that unscrambles them and endeavors the data. If mixed information are generous or client is a compact customer, so it was amazingly ineffective and is no too supportive. Else it may send his key to Cloud Servers who perform the unscrambling for request frameworks. It may cause an extraordinary detriment that Cloud Server gets riddle key such a substantial number of models exist to guarantee the genuineness of data record.

This highlights issues on privacy protection in the cloud like how the security of the customer is guaranteed. For creating privacy protection concerns, various new advances like HIPAA (Health Insurance Portability and Accountability Act), HHS (Human and Health Services), PHI (Protected Health Information), etc. have been suggested and diverse government on this planet were arranging lawful structures for guarantee of Privacy assurance and Security. Numerous associations' similar to healing facilities, land organizations, Visa organizations, web indexes, money and business informatics gathers and holds substantial measure of information likewise because of development of different web and its related advances, different online applications are picking up prevalence every day. Information excavators utilize this information for the investigation and mining purposes which help these associations for increasing valuable learning. These Information may contain delicate or important data of one and person. For e.g., associations like doctor's facilities contains numerous therapeutic records of the patients; they give these database or records to the exploration group or information diggers to increase help learning with the end goal of the examination. Information diggers of mine and in addition examines these medicinal records to increase value worldwide wellbeing insights so security is a vital matter when information contains delicate data. The fundamental point of privacy is to ensure data; in the meantime, the information must create outside learning. Privacy safeguarding is keeping the revelation of work force

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 05, May - 2017

and delicate data. Making privacy issues spoke to as truly imperative and as a need for all. Information assurance is not against advancement and development; unexpectedly information privacy adds to certainty. It is a key variable in the computerized environment. The privacy-protection is utilized as a part of different fields as takes after like in Medical Databases, Bioterrorism Applications, Cyber Security, Banking, Credit card, Genomic Privacy, Location-based Mobiles, etc. Popular privacy preservation techniques include PPDM (Privacy Preservation Data Mining) [4], etc.
The straggling leftovers of the paper is sorted out as takes after. The PPDM (Privacy Preservation Data Mining) is addressed in Section II. The considered privacy breaches in the Cloud environment and protection challenges in disseminated figuring is presented in Section III. Section IV provides the literature survey about, the large obtainable Privacy and Reliability frameworks and additionally requests the present assurance sparing methods of insight. Section V provides comparative study of the various techniques with their advantages and limitations. Finally, Section VI concludes up the paper.

II. PRIVACY PRESERVATION DATA MINING

Data mining insinuates the techniques of isolating rules and samples from data. It is also ordinarily known as KDD (Knowledge Discovery from Data). Data mining advancement has ascended to perceive the distinctive illustrations, ranges of the data of gigantic measure of the database. Data mining's purpose is to accumulate data from diverse affiliations, and party into the data digger's database for making another excellent calculation, model the data, count and appraisal for the best correct outcome for that model. From the above thought, its will be attempted by one get-together data's besides different source affiliation data. The course of action of the perceptive showing task that predicts the estimation of an univariable considering the overt estimations of diverse variables. These estimation of value is in like manner predicting to the estimation of others credit to make the best class for the data model. The particular individual itemized information in its unique frame frequently contains different delicate data about people, and distributed such information quickly abuses singular's security. An errand absolutely critical is to create techniques and devices for mining information in a more unfriendly environment, so that the mined information remains for all intents and purposes valuable while singular protection is saved. This endeavor is called privacy preservation data mining (PPDM).

While the online(web) applications are getting to be expanding pervasive by nature, they additionally exhibit numerous new security and privacy protection challenges. Be that as it may, security dangers additionally impact adversely on touchy information and conceivably prompts the spillage of classified information. Also, PPDM strategies permit us to secure the delicate data before it gets distributed to the general society by changing the first microdata organization and substance. Here we planned to embrace a selective study on a percentage of the dangers to the security and PPDM methods as a bound together answer for ensure against dangers. Privacy is insurance for the honest utilization of individual data of the cloud clients. Security is insurance for the honest utilization of individual data of cloud clients. The fundamental point of privacy assurance of data, in the meantime information must deliver outside learning. The microdata comprises at a table which is known as a micro table. In this micro table the identifiers (e.g., pin code and citizenship) can be utilized exclusively to recognize a table, so they should totally overlook. In view of information worth, it is isolated into four sorts:

1. Identifiers: - An identifier is a name that recognizes either an extraordinary article or a novel class of items. (e.g. pin code)
2. Quasi-Identifiers: - Quasi-identifier are bits of data that are not of themselves one of a kind identifier, but rather are adequately all around connected with an element that they can be joined with other quasi identifiers to make an extraordinary identifier. (e.g. pin code and maturity)
3. Sensitive Attributes: - The attributes which are important (age or salary) or contains sensitive information/data. (e.g. state)
4. Insensitive Attributes: - The attributes which are not very much important.
5. Micro-Data: - Properties which should not be uncovered in the discharged micro data set.
6. Equivalence class: - All arrangement of tables which cannot be recognized from one another as for Quasi Identifier.

Table 1. Non-Sensitive and Sensitive Attributes

Sl. No.	Non-Sensitive			Sensitive
	Pin code	Maturity	Citizenship	State
1	24164	19	India	AIDS
2	24179	30	USA	AIDS
3	24179	11	South Korea	Fever
4	24164	14	USA	Fever
5	25964	40	Nepal	Tumor
6	25964	44	India	AIDS
7	25961	38	USA	Fever
8	25961	50	USA	Fever
9	24164	22	USA	Tumor
10	24164	28	Nepal	Tumor
11	24179	27	South Korea	Tumor
12	24179	26	USA	Tumor

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 05, May - 2017

III. PRIVACY BREACHES

While considering privacy dangers in the cloud environment, it is essential as privacy dangers contrast as indicated by the kind of cloud situation. Some of those issues in privacy are like revelation of break and plate security, absence of client control, etc. A novel strategy of privacy safeguarding information mining might ensure against all the divulgence dangers while keeping up the exchange off between information utility, data misfortune and danger of revelation. Numerous creator referred to protection safeguarding algorithms with a specific end goal to smooth keep running of business information, however the privacy conservation algorithms are not free from difficulties. A percentage of the difficulties includes Privacy of Information, Data Ruptures, Malicious Insiders, Data Misfortune, International Information Privacy Laws, etc. At the point, when distributed microdata scale information (small scale information table comprises of sets of records of people or monetary substances), there are four sorts of protection revelation dangers as takes after:

1. Openness of Identity Threat: - It occurs when an individual is
   linked to a particular record in the released table. i.e., from the released table adversary can find a particular person. [5]

2. Participation Uncovering Threat: - In a specific database (such as a dataset containing cancer patients) if someone is unable to decide whether the record of any individual is present on the dataset or not, then dataset is free from membership disclosure. In some case, it will be better to use identity disclosure control techniques when an adversary is unknown about the membership of an individual, and in cases when the adversary knows the individual's record then the membership disclosure techniques is not sufficient enough. [5]

3. Analytical Revealing Threat: - If adversaries are able to estimate the confidential data from the released data then it is said that statistical disclosure has been taken place. For disclosure control either released data is modified (perturbation of data) or reduced (broad branding of data) to an acceptable level. The method chosen fully depends on the data to be released. [5]

4. Trait Uncovering Threat: - This happens when any additional information about an individual is revealed from the released dataset. Identity disclosure leads to attribute disclosure. Attribute disclosure can occur with or without identity disclosure (if someone knows the identity of a person in the dataset, he can easily find the sensitive information or for all matching tuples there are the same sensitive attributes). [5]

IV. Review of Some Exclusive Techniques For Privacy Preservation

In 1998 P. Samarati and L. Sweeney [6] and in 2002 L. Sweeney [7] suggested k-anonymity to restrain the disclosure hazard, they presented the k-Anonymity (privacy) protection procedure, that entails every record in an Anonymized (table) structure to be same with at most (k-1) different record's inside of the same datasets, concerning the arrangement of semi identifier(quasi-identifier) traits. To accomplish this k-Anonymity, they utilized the pair suppression and generalization for information Anonymization. Not at all like customary privacy safeguarding methods, for example, swapping data and noise addision, in k-Anonymous table through suppression as well as generalization stays honest. In particular, a table is k-Anonymous if the QI estimations of each tuple are indistinguishable to those of in any event (k-1) different tuples. All in all, k-anonymity ensures that an individual can be connected with his genuine tuple with a likelihood of at most 1/k. While k-anonymity ensures against personality revelation, it doesn't give adequate assurance against attribute divulgence. There are two assaults in k-anonymity: the background knowledge attack and the homogeneity attack. On account of the impediments of the k-anonymous model stem from two suspicions [8]. To start with, it might be hard for the database proprietor to figure out which of the traits are not accessible in the outer tables. The second confinement is that the k-anonymity model accepts a sure strategy for assault, while in genuine situation there is no motivation behind why the assailant ought not endeavor different strategies.

In 2008 R. Agrawal and R. Srikant[9] builds up a conveyance based information (data) digging computation for characterization obstacle, though the systems.

In 2000 Y. Lindell & B. Pinkas[10] tended to the issue of privacy-protection information mining. In specific, they considered situation where two parties having secret database wish to undergo an information mining algorithm on the unision of the databases, without revealing part of non-essential data. Their work is influenced by the need to both secure uncommon information and enable its usage for examination or distinctive purposes. The above issue is a specific specimen of secure multi-party computation and in light of current circumstances, can be handled using known non particular traditions. In any case, information mining calculations are regularly mind boggling and, moreover, the info generally comprises of gigantic information sets. Non-specific conventions in such a case are of no down to earth use and in this way more productive conventions are required. They focused on the issue of decision tree learning with the standard ID3 estimation. Their tradition is broadly more beneficial than nonexclusive game plans and demands both not a lot of rounds of correspondence and sensible information transmission.

In 2002 Agrawal et al. [11] proposed techniques of the Privacy Safeguarding Association Rule Mining. There are shred methodologies produced for conveyance based digging of information for specific issues, for example, association rule mining and grouping, it is clear that utilizing dispersions rather than unique records limit the scope of calculation that can be utilized on the information [12]. In the perturbation method, the circulation of each of information measurement is reproduced freely. This implies any dispersion based information mining calculation works under a verifiable assumption to treat all measurements autonomously. In numerous cases, a ton of pertinent data for data mining

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 05, May - 2017

calculations. Different methodologies, for example, multivariate decision tree calculations can't be as requirements acclimated to work with the information irritation strategy. This is an immediate aftereffect for autonomous treatment for distinctive qualities by irritation technique. this proposes transport established information extraction calculations has a sure insult of loss of comprehended data accessible in multidimensional records.

In 2003 M. Kantarcıo˘glu and C. Clifton [13] addressed secure mining of association rule over on a horizontally level plane divided information. The strategies consolidate cryptographic methods to minimise the data sharing, although adding minimal mining overhead to the undertaking. Cryptographic appliances can empower information mining that would somehow be forestalled because of security concerns. The strategies to mine circulated affiliation rules on a level plane divided information are suggested. The requirement for mining of information where access is limited by privacy concerns will increment. Indeed, even inside of a solitary multi-national organization, privacy laws in various purviews might anticipate sharing individual information. Computations for mining association rules over numerous databases have been portrayed for evenly and vertically parceled information separately. In the safe multi-party calculation approach, the objective is to construct an information mining model over numerous databases without uncovering the individual records in every database to alternate databases and summed up ways to deal with decreasing the quantity of "on-line" parties required for computation.Information(data) mining as a multidisciplinary joint exertion from databases, machine learning, and insights, is championing in transforming piles of information into chunks. The term is a misnomer, in light of the fact that the objective is the extraction of examples and information from a lot of information, not the extraction (mining) of information itself [14].

In 2005 Wang et al. [15] has pointed that cryptography does not spare the yield of a figuring data. Possibly, it thwarts protection spills amid the time spent the count. As needs be, it comes up short concerning giving a complete response for the issue of security sparing information(data) mining.

In 2006 Luar et al. [16] showed that the Cryptography framework ended up being immensely standard for fundamentally two reasons: Firstly, cryptography shows an inside and out described model for protection, which joins both methodologies for exhibiting and assessing it. Additionally, there wins an enormous toolset of cryptographic estimations and assembles to execute security assurance sparing data mining figuring.

In 2010 Kumar et al. [17] gave a novel convention to figure the aggregate of individual information inputs with zero likelihood of information spillage when two neighboring parties plot to know the information of a center gathering. By breaking the information piece of every gathering into number of sections and redistribute the fragments among gatherings before the computation. These whole steps make a situation in which it gets to be outlandish for semi fair gatherings to know the private information of some other gathering. The suggested dk-Secure Sum protocol for computation of whole of individual gatherings saving privacy protection of their inputs. The convention permits gatherings to break their information inputs into fragments and appropriating these portions among gatherings before computation. It gives zero likelihood to two conniving neighbors when they need to assault information of a center gathering. This is a considerable change over past conventions accessible in the writing. Further endeavors can be made to lessen the calculation and the correspondence intricacy safeguarding the property of zero hacking.

In 2011 Bhuyan et al. [18] attempted to add a protected computation model for the security conservation in conveyed information by executing multiparty correspondence over the distributed system. The distributed data are considered for computation purpose in decentralized and centralized manner. So, the data generated is mined in a centralized manner for pattern recognition or efficient knowledge discovery through distributed manner or for collaborative computing. So, it raised the seriousness of the privacy issues. For solving this challenge, various data mining communities responded by suggesting various different algorithms for preserving the privacy. The annoyance methodology is less expensive and additionally the shield an expansive number of clients in zones that are disconnected. The diversion hypothesis is utilized as a part of numerous conveyed information mining methods in the protected multiparty system [19]. The suggested approach is about the privacy issue in coursed information mining, and learning assurance strategies utilizing unsettling influence structure. The unsettling influence systems are successful in guaranteeing the privacy of the information reasonably and shield the exactness of the first dataset also keep up the consistency of client information. This methodology, gives more protection of privacy via twofold figuring at focus as well as switch. To exhibits a concurrent/non-concurrent system of checking as well as ensuring the isolation for information. Structure's able and definite as it once includes shuts every middle point the system gets a general sensible result. Because of solid correspondence adaptable quality and locally synchronous nature of the system, it is particularly flexible.

In 2012 Bhuyan et al. [20] suggested a features selection scheme for the privacy-preservation of data in a centralised networks. In a centralized network data assessment, the classification of data and featurse selection for mining data makes the structural model. For better performance, the gain ratio method of feature selection is performed in the centralized task. As all features don't have confidential information so it does not need to preserve-privacy of all the feature. For classification purpose of data, the Chi-Square test in the centralized model of data mining is used having its own processing unit. Without violating the privacy of the individuals to the alias name for preserving privacy in data-mining has considered to develop methods of data-mining for making the finest model, the task of Data Miner which is a Trusted Third Party is responsible for choosing the best features. For privacy preservation of data Chi-Square test and gain ratio is applied.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 05, May - 2017

In 2012 Sugumar et al. [21] depicted the usage of cryptography in that information digging for privacy protection safeguarding. The suggested secure multiparty computation for protection saving information mining has accomplished striking results nonspecific developments can be utilized to register any capacity safely and a few capacities can be figured considerably all the more productively utilizing particular developments. Still, a safe convention for registering a specific capacity will dependably be costlier than an innocent convention which doesn't give much security. Utilization of cryptograph systems to store touchy information as well giving access to, the put away information in view of an individual's part, guarantees that the information is sheltered from privacy breaks. The essential thoughts from a huge collection of cryptographic exploration on secure appropriated calculation, and their applications to information mining is illustrated. The primary parameter that influences the plausibility of actualizing a protected convention in view of the nonspecific developments is the extent of the best combinatorial circuit that figures the capacity that is assessed. The further research around there is pivotal for the improvement of secure and effective conventions in this field.

In 2012 SathiyaPriya et al. [22] suggested the method of hiding fuzzy association rule, where the fuzzified information is mined utilizing a changed Apriori algorithm as a part of request to remove leads and recognize delicate tenets. By diminishing bolster estimations of Right Hand Side (RHS) of the principle the sensitive rules can be hidden. To safeguard security and privacy of the database and to keep the utility and uncertainty of the mined rules at the peak Genetic algorithm is used.

In 2014 Pattnaik et al. [23] tried to contrast the consequence of privacy-protection with and without a trusted third party for horizontal apportioned information. What's more, to give high security to the information Gatherings with zero rate of information spillage utilizing the Apriori algorithm. As the description and prediction are the two basic fundamental objectives of data mining. The fundamental point in numerous scattered routines for privacy protecting information mining is to consent to helpful total calculations of total information set by safeguarding Privacy protection of individual Parties information or data. Especially in distributed information mining, security safeguarding is individual critical element. Secure multi party calculation is a valuable way to deal with recovery from the privacy protection in circulated information mining. Privacy Protection safeguarding information mining uses a data mining algorithmic rule for getting commonly gainful Global information extraction targets without accommodating personnel information. In this manner, in numerous information mining applications protection safeguarding has turned into a critical subject. The diversion hypothesis is utilized as a part of numerous conveyed information mining methods in the protected multiparty system [19].

V. COMPARATIVE STUDY

Here we are providing a comparative study of the various techniques with their advantages and limitations.

Author	Technique and Parameters	Advantages	Limitations
P. Samarati and L. Sweeney [6]	K-anonymization technique Prevent Data distortion (preserve integrity), preserve data quality, generalization, data suppression	A record from a dataset can't be recognized from at any rate k-1 records whose information is additionally in the dataset.	There are two assaults: the homogeneity assault and the background learning assault. k-anonymity model accepts a specific technique for assault, while in genuine situations there is no motivation behind why the aggressor ought not attempt different strategies.
L. Sweeney [7]	K-anonymity technique Information protection, re-identification of attacks	A record from a dataset can't be recognized from in any event k-1 other records present in the dataset.	Two attacks are possible: homogeneity and background knowledge attack
R. Agrawal [9]	Perturbation technique Data accuracy, data reconstruction	Free treatment of the diverse qualities by data perturbation approach	The technique does not remake the first information(data) values, but rather just appropriations to complete mining of the information accessible.
B. Pinkas [10]	Cryptographic technique secure multi party computation	A procedure through which touchy information can be scrambled. There is likewise an appropriate toolset for algorithm of cryptography.	This methodology is particularly hard proportional when more than a couple parties are included. Likewise, it doesn't hold useful for substantial databases.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 05, May - 2017

R. Agrawal [11]	Randomized response limiting privacy breaches, support recovery, randomization, data recovery	The randomized method is a simple technique which can easily implemented at data collection time.	Randomized response technique is not for multiple attribute databases.
M. Kantarcioglu [13]	Cryptographic procedures safely mining of association ruls, over on a level horizontally divided data, minimize data sharing, commutative encryption	In light of Secure Multi-Party Computation (SMC) systems addresses secure mining of association rules over on a level horizontal parceled information.	It involves overhead cost during mining task. It is not free from collision
Ke Wang [15]	K-Anonymization Technique Template based privacy preservation, preserve information, limit useful undesirable touchy inductions got from information, optimization for data suppression, preserve classification value of data	k-anonymization for data owned by multiple parties was considered. To protect the data for a needed classification examination and breaking point the convenience of undesirable delicate derivations that might be gotten from the information. Delicate inductions are determined by an arrangement of privacy-templates.	For an extensive information set, finding an ideal suppression is hard, since it requires enhancement over all suppressions.
S. Laur [16]	Cryptographic technique Privacy Preservation, private classification, private polynomial kernel computation, integrity preservation, private kernel sharing, evaluate polynomial kernels	Here portrayed cryptographically secure conventions for Kernel Perceptron and kernelized Support Vector Machines. Likewise gave cryptographically secure conventions to assessing polynomial portions, for kernelized direct characterization and for accumulation of scrambled classification results.	There are still numerous open issues in private SVM order and learning. Be that as it may, how to safely conceal the merging rate of the Kernel Perceptron and the Kernel Adatron algorithms. What's more, whether there are any iterative private direct classification techniques that need no circuit assessment.
R. Sheikh [17]	Cryptographic technique data leakage, privacy of inputs, communication complexity preserving, zero hacking	Secure sum protocols allow multiple cooperating parties to figure entirety capacity of their personal data without disclosing the data to one another.	It causes problem if parties are increased i.e. more than 4 and in case if some parties are dishonest then it's a disadvantage.
Hemanta Kumar Bhuyan [18]	Monitoring and Privacy of data, knowledge discovery of data using centralized or decentralized collaborative computation, privacy of distributed data, preservation of accuracy of data set	Multiparty computation in peer-to-peer network. The computation is performed by taking the distributed data-set of a particular scenario through centralized and decentralized fashion.	Firstly, like out of chosen 300 instances how to evaluate only 14 features out of the 76 features. How to check the importance of only 14 features out of 76 features? Secondly, each node has to preserve their privacy while sending data to the coordinator. And lastly, after privacy preservation of data by the coordinator how it will be decrypted the data back to original form.
Hemanta Kumar Bhuyan [20]	Privacy preservation in centralized network (perturbation technique, gain ratio, chi square testing, micro aggregation of data, feature preservation	Features selection is finished with privacy-preservation in centralized networks. The utilization of gain ratio proportion system for better execution for selecting feature has taken to perform the incorporated computational errand.	Privacy preservation for large number of multiparty is not possible.
Sugumar [21]	Cryptographic technique data safety, secure distributed computation	By using generic constructions to compute any function securely data mining calculation on the unison of the databases, without uncovering any unnecessary information.	A safe protocol for processing a certain function will dependably be costlier than a gullible protocol that does not give any security. The fundamental parameter that influences the possibility of executing a safe protocol in view of the non-specific developments is the measure of the best combinatorial circuit that figures the capacity. Improvement of secure and productive protocols in this field is must.
K. SathiyaPriya [22]	Association Rule Mining Ensure database security (genetic algorithm), mining of fuzzified data (apriori rule), hiding of sensitive rules (decrease support value of RHS), privacy preservation, minimum modification of data	Here a technique to stow away fuzzy association rule is proposed, in which, the fuzzified information is mined utilizing adjusted apriori algorithm as a part of request to concentrate runs and recognize sensitive rules. Genetic algorithm is utilized to guarantee security of the database as well as keep the utility and assurance of the mined standards at largest amount.	This approach makes minimum modification of data. There must be a method of finding the criteria to choose the best chromosome for crossover in order to reduce the number of generation. It is also required to reduce the side effects when choosing the generation with minimum modification.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 05, May - 2017

Prasant Kumar Pattnaik [23]	Cryptographic technique Privacy preservation with and without trusted parties for horizontal partition of data, zero data leakage, high preservation to data parties	Here contrasted the outcome privay saving system with and without trusted party for flat/horizontal divided information. Also, give high security to information parties with rate of information spillage is zero percent by utilizing the Apriori algorithm.	Centralized network is a bottleneck and all parties have similar type of dataset and features.

VI. CONCLUSION

The motivation behind this work is to complete the study of the diverse privacy safeguarding information digging strategies for enhancing information quality and viability. There are different strategies present for security protection in information mining yet they have some imperfection like data misfortune, and so forth. This work is essentially centered around consolidated strategies for k-Anonymity, association rule mining, cryptographic, and information irritation systems i.e., data perturbation to protect the privacy

of information and lessening data misfortune. This investigation of distinctive methods helps in building up another arrangement of privacy protection conservation in the information mining that may join two or more procedures to secure database against diverse dangers. This paper expects to repeat a privacy saving information(data) mining advancements obviously and after that returns to dissect the benefits and deficiencies of these innovations. In future we will attempt to discover the procedures of privacy safeguarding information(data) mining so it will ensure against most extreme number of dangers with less data misfortune and high utility of information present in the database.

REFERENCES

[1] Bhagyashri S, Gurav YB. A Survey on Privacy-Preserving Techniques for Secure Cloud Storage. International Journal of Computer science and Mobile computing. 2014 Feb.

[2] Kamakshi P, Babu AV. Preserving privacy and sharing the data in distributed environment using cryptographic technique on perturbed data. arXiv preprint arXiv:1004.4477. 2010 Apr 26.

[3] Rezgui A, Ouzzani M, Bouguettaya A, Medjahed B. Preserving privacy in web services. InProceedings of the 4th international workshop on Web information and data management 2002 Nov 8 (pp. 56-62). ACM.

[4] Aggarwal CC, Philip SY. Privacy-preserving data mining: a survey. InHandbook of database security 2008 (pp. 431-460). Springer US.

[5] Templ M, Meindl B, Kowarik A. Introduction to statistical disclosure control (sdc). Project: Relative to the testing of SDC algorithms and provision of practical SDC, data analysis OG. 2013.

[6] Samarati P, Sweeney L. Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression. Technical report, SRI International; 1998 May.

[7] Sweeney L. k-anonymity: A model for protecting privacy. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems. 2002 Oct;10(05):557-70.

[8] Wong RC, Li J, Fu AW, Wang K. (α, k)-anonymity: an enhanced k-anonymity model for privacy preserving data publishing. InProceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining 2006 Aug 20 (pp. 754-759). ACM.

[9] Agrawal R, Srikant R. Privacy-preserving data mining. InACMSigmod Record 2000 May 16 (Vol. 29, No. 2, pp. 439-450). ACM.

[10] Lindell Y, Pinkas B. Privacy preserving data mining. InAdvances in Cryptology—CRYPTO 2000 2000 Aug 20 (pp. 36-54). Springer Berlin Heidelberg.

[11] Evfimievski A, Srikant R, Agrawal R, Gehrke J. Privacy preserving mining of association rules. Information Systems. 2004 Jun 30;29(4):343-64.

[12] Hong JI, Landay JA. An architecture for privacy-sensitive ubiquitous computing. InProceedings of the 2nd international conference on Mobile systems, applications, and services 2004 Jun 6 (pp. 177-189). ACM.

[13] Kantarcioglu M, Clifton C. Privacy-preserving distributed mining of association rules on horizontally partitioned data. IEEE Transactions on Knowledge & Data Engineering. 2004 Sep 1(9):1026-37.

[14] Han J, Kamber M, Pei J. Data mining: concepts and techniques. Elsevier; 2011 Jun 9.

[15] Wang K, Fung B, Yu PS. Template-based privacy preservation in classification problems. InData Mining, Fifth IEEE International Conference on 2005 Nov 27 (pp. 8-pp). IEEE.

[16] Laur S, Lipmaa H, Mielikäinen T. Cryptographically private support vector machines. InProceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining 2006 Aug 20 (pp. 618-624). ACM.

[17] Sheikh R, Kumar B, Mishra DK. A distributed k-secure sum protocol for secure multi-party computations. arXiv preprint arXiv:1003.4071. 2010 Mar 22.

[18] Bhuyan HK, Kamila NK, Dash SK. An approach for privacy preservation of distributed data in peer-to-peer network using multiparty computation. IJCSI International Journal of Computer Science Issues. 2011;8(4).

[19] Kargupta H, Das K, Liu K. Multi-party, privacy-preserving distributed data mining using a game theoretic framework. InKnowledge Discovery in Databases: PKDD 2007 2007 Sep 17 (pp. 523-531). Springer Berlin Heidelberg.

[20] Bhuyan HK, Mohanty M, Das SR. Privacy Preserving for Feature Selection in Data Mining Using Centralized Network.

[21] Sugumar J. R., Rengarajan, C.: Design a Secure Multi Site Computation System for Privacy Preserving Data Mining. International Journal of Computer Science and Telecommunications. 2012; 3:101-5.

[22] Priya KS, Sadasivam GS, Karthikeyan VB. A New Method for Preserving Privacy in Quantitative Association Rules using Genetic Algorithm. International Journal of Computer Applications. 2012 Jan 1;60(12).

[23] Kumar PD, Raghvendra K, Yogesh S. Privacy Preservation in Distributed Database. European Journal of Academic Essays. 2014;1(2):35-9.