**Chapter 1**

**INTRODUCTION**

Privacy preservation in data management and publishing has grown to be a vital research area in the era of big data. Efficiently protecting individual privacy in data publishing is especially critical due to variation in personal preference and sensitivity. The consequences of private data getting published is causing psychological issues and disturbances in individual’s personal life. This has triggered the requirement to develop various approaches for privacy preservation in data publishing. Game theory is one of the approach adopted for privacy preservation in data publishing. The study of situation involving computing interest, modeled in terms of the strategies, probabilities, actions, gains, and losses of opposing players in a game is game theory. A comparison of PPDP with PPDM has been done to explore the utility. Most current studies only manage to achieve personalized privacy preserving in a statistical sense, though some researchers have investigated the issue of data anonymization using K anonymity algorithms for personalized privacy protection. K-anonymization techniques have been the focus of intense research in the last few years. In order to ensure anonymization of data while at the same time minimizing the information loss resulting from data modifications At present, personalized privacy preserving in data publishing (PPPDP) is still in the premature stage of development. To form a basis of development, a classification method has been formulated to collect sensitive and non-sensitive data separately, where sensitive data has to be strongly protected. A survey was conducted on the sensitivity of data based on individual’s social , economical, personal and psychological factors. A conceptual approach has been designed to achieve personalized privacy preservation in data publishing (PPPDP) based on the classification of sensitivity of individual. Based on the sensitivity classification, Game theory has been proposed for achieving personalized privacy preservation applied in the field of data publishing in the financial and banking sector. This approach can be extended to other sectors of data publishing like social media networks, matrimony data , reviews etc. as future research avenue.

**Chapter 2**

**LITERATURE SURVEY**

As a part of my research I have gone through various journals which would lead to the problem definition of my research, scope of the work and probable solution for the same. I give below the various topics reviewed by me.

Bin Zhou et al (2007) worked on many Web 2.0 applications, more and more social network data has been made publicly available and analyzed in one way or another. Privacy preserving publishing of social network data has become a more and more important concern. In this experiment, authors presented a brief yet systematic review of the existing anonymization techniques for privacy preserving publishing of social network data. In this research, they surveyed a few recent studies on anonymization techniques for privacy preserving publishing of social network data. Although privacy preserving data publishing and analysis techniques in relational data have been well explored, the research and development of anonymization techniques on social network data is still in its infancy. The new challenges in privacy preservation in social network data comparing was discussed to the extensively studied relational case, and examined the possible problem formulation in three important dimensions: privacy, background knowledge, and data utility. Authors reviewed the anonymization methods for privacy preservation in two categories: clustering-based approaches and graph modification approaches.

Rosa Karimi Adl et al (2012) worked on Privacy compatible with animating systems via game theory. Various anonymous methods have been proposed to address privacy issues in individual databases. They described a general approach to solving such games and a model anonymously using the k-code to extend the process. The simulations of the game reveals how the simulations of the game depend on the values of the Q Khan equilibrium, the number of wave-identifiers, the maximum number of recurring recording, the annexation cost and public privacy behavior.

Srinivasa L et al (2012) worked on A Coalitional Game Theoretic Mechanism for Privacy Preserving Publishing Based on k-Anonymity. K-Anonymity is one of the most popular technologies for protecting privacy of one's individual, thus avoiding awareness of serious information. If this is a violation of the limits made to a limited loss, and if the loss of information is found to be too low for the loss, the loss of information can be traced after alienation. Each game must return the entire process. This approach ensured the privacy level of the gateway given to the loss of information. In this process they have good relations between the number of k and the number of Tuples published in the table.

Y. Xu et al (2013) worked on a personalized k-anonymity privacy preserving method. User-based based methodology has been reviewed on data mining related privacy issues. They distinguish between four different user roles commonly used in data mining applications, such as data issuer, data collector, data maker, and decision maker. It has established formal privacy models to measure the potential loss of privacy under various attacks, and use anonymization techniques for data.

LEI XU et al (2014) worked on Information Security in Big Data: Privacy and Data Mining. The growing reputation and growth of data mining technologies brings a serious threat to protecting the person's sensitive information. This paper analyzes various approaches to help protect critical information, taking into account the privacy issues associated with the broader perspective associated with data information. In particular, they have identified four different types of users engaged in data mining applications, such as data provider, data collector, data maker, decision maker. In each case, we discuss the user's privacy and concerns that can be followed to protect awareness information. This reviewed the game's theoretical approaches proposed to analyze the interaction between different users in a data mining environment. Each of them has important information. By distinguishing the responsibilities of various users in protecting important information, they provide some useful insights for PPDM's research.

ARUNUDAYA R et al (2015) worked on energy efficient data sharing using information centric approach with information security in big data. Big data strongly demands a network infrastructure having the capability to efficiently collect, process, cache, share, and deliver the data, instead of simple transmissions. The growing popularity and development of data mining technologies bring serious threat to the security of individual's sensitive information. In this paper, they viewed the privacy issues related to data mining from a wider perspective and investigate various approaches that can help to protect sensitive information.. Current studies of PPDM mainly focus on how to reduce the privacy risk brought by data mining operations, while in fact, unwanted disclosure of sensitive information may also happen in the process of data collecting, data is publishing, and information delivering.

Suma Reddy et al (2015) worked on the Privacy Preserving Publishing of Social Network Data Privacy and Big Data Mining. Privacy is a major concern in big data mining. One main characteristic of social networks is that they keep evolving over time. They reviewed the privacy issues related to data mining by using a user-role based methodology. We distinguish four different user roles that are commonly involved in data mining applications, i.e. data provider, data collector, data miner and decision maker.

Jinling Song et al (2015) researched in k-anonymity model which can prevent publishing data from disclosing privacy effectively and efficiently, due to the uneven distribution of the sensitive data, ordinary k-anonymization method cannot guarantee each tuple satisfying the personalized privacy requirement of its data owner although the publishing table has been satisfied k-anonymity constraint. An anonymization method satisfying personalized privacy requirements was presented, in which a utility-driven adaptive clustering method is proposed to partition tuples with similar best data quality. In this experiment, author analyzed the disclosure caused by k-anonymization even under uniform privacy requirements because the slope of sensitive values..

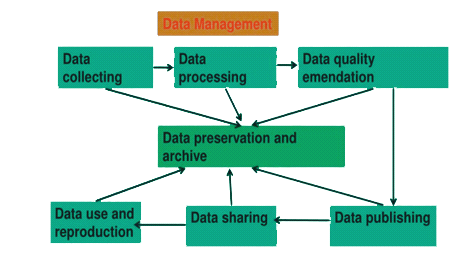
Prof. Harish Barapatre et al (2016) worked on Information Security in Large Amount of Data: Privacy and Data Mining. To avoid access to one’s sensitive information, Privacy Preserving Data Mining (PPDM) is used. In this technique, data gets modified in order to secure one’s sensitive information. PPDM technique is mainly focused on how privacy maintained at Data Mining. However, Sensitive information can be retrieved at Data Collection, Data Publishing and Information Delivering processes. In this Paper, they briefly discussed the Privacy Preserving Data Mining with respect to user such as Data Provider, Data Collector, Data Miner and Decision Maker. They discussed privacy concerns related to each user. The data collector, the data miner also faces the privacy utility trade-off problem. By using different algorithm techniques such as Decision tree, Support Vector Machine, Naive Bayesian Techniques Data collector modifies data and sends it to the Decision maker.

It is often necessary to publish personal information for statistical and research purposes. (Xia oui Xiao et. al) The process of privacy preservation in data publishing works in a certain pattern. Data owner offers the data to data publisher, which would have their individual data containing sensitive information. The data publisher treats the collected data uniformly to meet the privacy model. Finally data publisher publishes the data satisfying privacy requirement to the data receiver for statistical analysis. Although the data owner has authorized the data publisher to deal with his individual data for protecting privacy, the uniform treatment of data may not meet privacy requirement of each individual. This leads to problem of personalized privacy preservation (Jinling Song et. al.). This calls for a privacy preservation technique to reduce the possibility of identifying sensitive information about individuals, which is also called as disclosure control problem (Raymond Chi-Wing Wong et. al.).

From the above reviews, it is seen that very less work has been done on achieving Personalized Privacy Preservation in Data Publishing (PPPDP). More researches can be seen in respect of privacy security in data mining than that of data publishing.

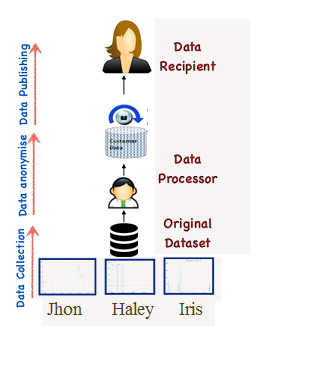
**Privacy-Preservation in Data Publishing**

A typical scenario for data accumulation and publishing is described in Figure 1. In the data amassment phase, the data publisher accumulates data from record owners (e.g., Alice, Bob and peter). In the data publishing phase, the data publisher releases the accumulated data to a data miner or to the public, called the data recipient, who will then conduct data mining on the published data.



**Figure 2.1: Data Management**

In order to solve the privacy preserving problem of association rule in centralized database, before publishing database we should hide the privacy or the sensitive information pattern.

. 

**Figure 2.2: An Overview of Privacy-Preservation in Data Publishing**

Bee-Chung Chen et.al (2009) has worked and explained that given a data set, privacy-preservation in data publishing can be intuitively thought of as a game among four parties:

* **Data user** could be member of library, who is interested to use the data or information in library.
* **Adversary** is person who wants to derive private information from the data.
* **Data publisher**, who collects the data and wants to release the data in a way that satisfies the data user’s need but also prevents the adversary from obtaining private information about the individuals in the data.
* **Individuals**, who collect the data from the data publisher.

In some cases, the individuals agree with the data publisher’s privacy policy, trust the data publisher and give the data publisher all the requested information. In these cases, it is the data publisher’s responsibility to ensure privacy preservation. In other cases, the individuals do not trust the data publisher and want to make sure that the data publisher cannot precisely identify their sensitive information. There is a fundamental tradeoff between privacy and utility. At one extreme, the data publisher may release nothing so that privacy is perfectly preserved; however, no one is able to use the data. At the other extreme, the data publisher may release the data set without any modification so that data utility can be maximized. the following three components need to be defined.

**Sanitization Mechanism**: A sanitization mechanism sanitizes the data set by making the data less precise. This mechanism defines the space of possible “snapshots” of the original data set that are considered as candidates for release. We call such a snapshot a release candidate. Generalization is an example sanitization mechanism.

**Privacy criterion**: Given a release candidate, the privacy criterion defines whether the release candidate is safe for release or not. K-Anonymity is an example privacy criterion.

**Utility metric**: Given a release candidate, the utility metric quantifies the utility of the release candidate (equivalently, the information loss due to the sanitization process).

Given the above three components, one approach to privacy preserving data publishing is to publish the most useful release candidate that satisfies the privacy criterion.

**Chapter 3**

**OBJECTIVES OF THE WORK**

This dissertation addresses the following problem. Here a distributed computing environment is considered having number of nodes (parties) connected via an established communication infrastructure. Each node has some data which is the Knowledge known only to itself. The nodes exchange messages, data with any other node in the network. This research aims at answering the following questions:-

1. How can we define and explain the concept of personalisation with respect to data publishing.
2. How user can specify their own privacy requirements without having to adhere to a monolithic privacy definition.
3. How can we ensure that the required privacy is actually achieved without having to rely on unrealistic assumptions regarding the behaviour of other parties and
4. How can we analyse the interaction among different users in a game theoretical approach.
5. Compare the game theoretical approach with the classification model with reference to data publishing.

The main objective of personalized privacy preservation is to ensure that private data remains protected, while processing or releasing sensitive information of each individual during each phase of the knowledge data discovery process. The objective of data anonymization is to prevent the potential adversary from discovering information about a certain individual (i.e. the target). The adversary can utilize various kinds of knowledge to dig up the target's information from the published data. The anonymization operations will reduce the utility of data. The reduction of data utility is usually represented by information loss: higher information loss means lower utility of the anonymized data.

**Chapter 4**

**WORK CARRIED OUT SO FAR**

**4.1 Work carried out so is Survey on Privacy preservation in Data Publishing**

Data Publishing is used by many companies with a strong consumer focus such as retail, financial and marketing organizations. The word data privacy[1] refers to the relationship between collection and diffusion of data ability, the public privacy and the legitimate and opinionated issues. Data privacy concerns can arise in angry reply to information from a very wide range of resources, such as: health care records, residence and geographic records, ethnicity, criminal investigation and privacy breach.

We consider microdata is stored in a table and each record(row) corresponds to one identity. Each record has the form:

D(Identifier, Quasi Identifier, Sensitive Attributes, Non-Sensitive Attributes).

1. Identifier: Attributes that clearly identify the individuals. Examples include Social Security Number and Name.
2. Quasi Identifier: Attributes whose values when taken together can identify an individual. Examples include birthdate and Zip code. An adversary may already know( possibly from other publicly available databases)e.g., a voter registration list that include both explicit identifier and quasi identifiers
3. Sensitive Attributes: Attributes, which are unknown to the adversary and are considered sensitive such as disease, salary and disability status.
4. Non Sensitive Attributes: Attributes, that donot fall into the above three categories.
   * 1. **Privacy Preserving Techniques**

* **Generalization**

Generalization is one of the anonymized approaches, which replaces quasi identifier values with the values that are less specific but semantically consistent. Here the values are generalized to a range in order to reduce the granularity of representation. All QI values in the group would be generalized to the entire group extent in QID space.[3] If atleast 2 transactions in a group have distinct values in a certain column, then all information about that item in the current group is lost. Due to the high dimensionality of the QI, it is likely that any generalization methods would incur extreme high information loss, rendering the data useless[4].To perform data analysis on the generalized table, the data analyst has to make the uniform distribution assuming that every value in the set is equally possible, as no other distribution assumption be justified. This significantly reduces the data utility of the generalized data and also because, each attribute is generalized separately, correlations between different attributes are lost. assumption be justified. This significantly reduces the data utility of the generalized data and also because, each attribute is generalized separately, correlations between different attributes are lost.

* **Bucketization**

It is to partition the tuples in table ‘T’ into buckets and then to separate sensitive attributes from non-sensitive attributes by randomly permuting SA values within each bucket. The sanitized data then consists of the buckets with permuted sensitive values. It first partitions tuples in the table into buckets and then separate QI values with the sensitive attributes by randomly permuting the SA values in each bucket. The Anonymized data consist of a set of buckets with permuted SA values. In particular, bucketization has been used for anonymizing high dimensional data. However, their approach assumes a clear separation between Quasi Identifiers and Sensitive attributes.

* **K-Anonymity**

K-Anonymity guarantees that the data released is accurate. It focuses on two techniques in particular: Generalization and Suppression.[6] To protect respondents identity when releasing micro data, data holders often remove explicit identifiers. This k-anonymity technique, captures the protection of a micro data with respect to possible re-identification of respondents to which the data refer. One of the interesting aspect of k-anonymity is its association with protection techniques preserve truthfulness of the data. The first approach towards privacy protection was to perturb the data before it is mined. Similarly, second approach was developed using cryptographic techniques.

* **Game Theory Model**

Game theory provides a formal approach to model situations where a group of records(agents) have to choose optimum actions considering the mutual effects of other agents decisions.

The essential elements of a game are players, actions, payoffs and information[9]. Players have actions that they can perform at designated times in the game. As a result of the performed actions, players receive payoffs. The payoff to each player depends on both the player’s action and other player’s actions also. Information is modelled using the concept of information set which represents a player’s knowledge about the values of different variables in the game.

The outcome of the game is a set of elements picked from the values of actions, payoffs and other variables after the game is played out. A player is called “rationale” if he acts in such a way as to maximise his payoff. A player’s strategy is a rule that tells him which action to choose at each instant of the game, given his information set. A strategy profile is an ordered set consisting of one strategy for each of the player’s in the game. An equilibrium is a strategy profile consisting of a best strategy for each of the player in the game. The most important equilibrium concept for the majority of game is nash equilibrium. A strategy profile is a nash equilibrium if no player has the incentive to deviate from his strategy, given that other players donot deviate.

Game theory has been successfully applied to various application areas such as economics, political science, computer science etc. this game theory approach has been employed to deal with the privacy issues related to data publishing.

* + 1. **Applications**

The problem of privacy preservation in data publishing has numerous applications in homeland security, medical database mining and customer transaction analysis. Some of these applications such as those involving bio-terrorism and medical database mining are described here.

1. Bio-Terrorism Applications

Here, we like to analyze medical data for privacy preservation in data publishing. Often a biological agent such as anthrax produces symptoms are similar to other respiratory diseases such as the cough, cold and the flu. In the absence of prior knowledge of such a attack, health care providers may diagnose a patient affected by an anthrax attack of having symptoms from one of the respiratory diseases. The key is to quickly identify a true anthrax attack from a normal outbreak of a common respiratory disease. In many cases, an unusual number of such cases in a given locality may indicate a bio-terrorism attack.

1. Security Applications

Number of applications for homeland security are inherently intrusive, because of a very nature of surveillance. Various privacy preserving techniques may be used in order to deploy these applications effectively without violating user privacy. Some examples of such applications are:

1. Identity Theft

A Technology[10] is to use a more active approach to avoid identity theft. The identity angel system, crawls through cyberspace and determines people who are at risk from identity theft. This information can be used to notify the appropriate parties. We note that both the above approaches to prevention of identity theft are relatively non-invasive and therefore donot violate privacy.

1. Video Surveillance

In the context of sharing video data, a major threat is the use of facial recognition software, can match the facial images in videos to the facial images in a driver license database, while a straight forward solution is to completely black out each face, the result is of limited now, since all the facial information has been wiped out.

1. Web Camera Surveillance

One possible method for surveillance is with the use of publicly available webcam[11], which can be used to detect unusual activity. It is much more invasive approach than the previously discussed technique, because of person specific information being captured in the webcams.

* 1. **Development of Classification Model for Personalized Privacy Preservation in Data Publishing**
     1. **Introduction**

Privacy protection becomes an important issue when one wants to make use of data that involves individuals' sensitive information. Research on protecting the privacy of individuals and the confidentiality of data has become one of the major criteria of data security. With the rapid development of information technology and the wide application of networks, large-scale of digital information is stored and published. Knowledge discovery and data mining applications in information retrieval are playing an active role and has greatly contributed to the various departments right from data mining of useful information needs. In addition to the benefits of the digital information, many problems regarding the privacy is envisaged. The disclosure of sensitive information has become prominent now a days , and privacy preservation has become a research hotspot in the field of data security. The association among the illegal records in public security system, the customer's credit card transactions, telecommunications users' personal information, housing information, and soon are some data publishing entities. It is of great significance for government and business organization not to destroy the citizens' personal privacy. A reasonable and effective method of protection, which can protect the user’s privacy and keep the data available at the same time, is the trend of developments in information security.

* + 1. **Methodology**

Given a data set, privacy-preservation in data publishing can be intuitively thought of as a game among four parties:

* **Data user**, could be member of library , who is interested to use the data or information in library.
* **Adversary**, is person who wants to derive private information from the data.
* **Data publisher**, who collects the data and wants to release the data in a way that satisfies the data user’s need but also prevents the adversary from obtaining private information about the individuals in the data.
* **Individuals**, who collect the data from the data publisher.

****

**Figure 4.1: Game Theory Techniques**

* + 1. **Results**

**Figure 4.2 : Rating of Sensitivity for the Raw Data**

* 1. **A Nash Equilibrium Approach To User Based Data Masking and Retrieval**

Privacy protection becomes a key issue when you want to use data that contains important information from individuals. One of the key criteria for research data protection is to protect the privacy of individuals and to protect data confidentiality. The theory of the game is the study of the context that is based on strategies, probabilities, actions, benefits and losses to combat players in a match. A comparison of PPDP with PPDM has been done to explore the utility. Privacy concerns related to each user with the approach of game theory. In fact, quality data processing techniques are expected to be used in published data. In contrast, data holders in PDTM need to rectify data to retrieve data.

* + 1. **Methodology**

Most of the proposed approaches follow the following research precedent:

* Define the elements of the game, namely the players, the actions and the payoffs.
* Determine the type of the game: static or dynamic, complete information or incomplete information.
* Solve the game to find equilibriums.
* Analysis the apocalypse to odysema soy implications warfare.

The privacy protection has been implemented for the following steps

* Data entry
* Personalized Data masking
* Personalized retrieval of Data

This has been achieved using game theory approach –Nash Equilibrium technique with the programming language of Java.

**Chapter 5**

**EXPECTED OUTCOMES**

To achieve the privacy-preserving goals of different user’s roles, various methods from different research fields are required. We have reviewed recent progress in related studies, and discussed problems waiting to be further investigated. We hope that the review presented in this paper can offer researchers different insights into the issue of privacy-preserving data mining, and promote the exploration of new solutions to the security of sensitive information.

In this paper, classification of data based on sensitivity is done. This forms the basis for a theoretical approach of personalized privacy preservation in data publishing. The goal of this work is to implement a practical personalized privacy preserving framework to keep privacy of an individual as per the personalized preference of the individual. The core benefit of this work is to promote data sharing with personalized privacy protection. Based on the classification of sensitivity, a practical implementation is planned using Game Theory with concept of Nash Equilibrium for achieving personalized privacy preservation in data publishing. This is planned to be implemented on specific area like financial sector where personalized privacy preservations is considered to be important and is under the classification of restricted sensitivity player. The equivalent class in the anonymous table is created by creating different tuples based on their payoffs.

**Chapter 6**

**FUTURE WORK**

The implementation of Game Theory with concept of Nash Equilibrium for achieving personalized privacy preservation in data publishing is further improved for banking process and matrimony services.

**6.1 Know Your Customer (KYC)**

Finance is one of the rudimentary sources of every act in the world today. Institutes that bank or trade financial instruments or, are into any other pertinent business the institute does have a set of responsibilities in knowing their customers. This set of responsibilities predicated on desideratum of a country led to the birth of KYC. KYC will be a set of procedure that an individual, a group or any organization will undergo while trading with any financial institute. Identity proofs, proof of address, ilicit status, verification of signature are few of the processes that a person will undergo in KYC. Reserve Bank of India, the regulatory body of financial institutes in India made KYC compulsory to all banks. This order was passed in the year 2004 and according to it the objective of KYC is to obviate financial institutes and germane companies from being utilized by terrorists and launderers in a malefactor act.

KYC is an indispensable requisite to financial institutions in India. Failure to engender obligatory information or indispensable documents to KYC to any financial institution will lead to Refusal to open an account or discontinuation of subsisting account – for a bank. Refusal to sanction for any investment for mutual funds. By taking into account above fact in this paper a conceptual research has been developed with regard to personalized privacy preservation in data publishing. PPPDP has been applied to the KYC (know your customer) form of a bank and applied to KYC form of the Bank using Game Theory to achieve personalized privacy preservation.

**REFERENCES**

1. X. Xiao, Y. Tao, ``Personalized privacy preservation'', in *Proc. ACM SIGMOD Int. Conf. Manage. Data*, 2006, pp. 229\_240.

2. K. Qing-Jiang,W. Xiao-Hao, and Z. Jun, ``The (p, α, k) anonymity model for privacy protection of personal information in the social networks,'' in*Proc. 6th IEEE Joint Int. Inf. Technol. Artif. Intell. Conf. (ITAIC)*, vol. 2. Aug. 2011, pp. 420\_423.

3. B. Wang and J. Yang, ``Personalized (α, k)-anonymity algorithm based on entropy classi\_cation'',*J. Comput. Inf. Syst.*, vol. 8, no. 1, pp. 259\_266, 2012.

4. Xua, X. Qin, Z. Yang, Y. Yang, and K. Li, ``A personalized k-anonymity privacy preserving method'',*J. Inf. Comput. Sci.*, vol. 10, no. 1, pp. 139\_155, 2013.

5. S.Yang, L. Lijie, Z. Jianpei, and Y. Jing, ``Method for individualized privacy preservation,'' *Int. J. Secur. Appl.*, vol. 7, no. 6, p. 109, 2013.

6. R. K. Adl, M. Askari, K. Barker, and R. Safavi-Naini, “Privacy consensus in anonymization systems via game theory'', in *Proc. 26th Annu. DataAppl. Security Privacy*, 2012, pp. 74\_89.

7. R. KarimiAdl, K. Barker, and J. Denzinger, “A negotiation game: Establishing stable privacy policies for aggregate reasoning'', Dept. Comput. Sci., Univ. Calgary, Calgary, AB, Canada, Tech. Rep., Oct. 2012.

8. H. Kargupta, K. Das, and K. Liu, “Multi-party, privacy-preserving distributed data mining using a game theoretic framework'', in *Proc.11th Eur. Conf. Principles Pract. Knowl.Discovery Databases (PKDD)*, 2007, pp. 523\_531.

9. A. Miyaji and M. S. Rahman, “Privacy-preserving data mining: A game-theoretic approach'', in *Proc. 25th Data Appl. Security Privacy*, 2011, pp. 186\_200.

10. M. Halkidi and I. Koutsopoulos, “A game theoretic framework for data privacy preservation in recommender systems'', in *Proc. Mach. Learn.Knowl. Discovery Databases*, 2011, pp. 629\_644.

11. S. Ioannidis and P. Loiseau, “Linear regression as a non-cooperative game'', in *Proc. Web Internet Econ.*, 2013, pp. 277\_290.

12. S. L. Chakravarthy, V. V. Kumari, and C. Sarojini, “A coalitionalgame theoretic mechanism for privacy preserving publishingbased on *k*-anonymity'',*Proc. Technol.*, vol. 6, pp. 889\_896,2012.

13. J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*. San Mateo, CA, USA: Morgan Kaufmann, 2006.

14. R. Agrawal and R. Srikant, ``Privacy-preserving data mining'',*ACM SIGMOD Rec.*, vol. 29, no. 2, pp. 439\_450, 2000.

15. E. Rasmusen, *Games and Information: An Introduction to Game Theory*, vol. 2. Cambridge, MA, USA: Blackwell, 1994

16. R. Gibbons, *A Primer in Game Theory*. Hertfordshire, U.K.: Harvester Wheatsheaf, 1992.

17. M. Terrovitis and N. Mamoulis, “Privacy preservation in the publication of trajectories'', in *Proc. 9th Int. Conf. Mobile Data Manage.(MDM)*,2008, pp. 65\_72.

18. G. Poulis, S. Skiadopoulos, G. Loukides, and A. Gkoulalas-Divanis, “Distance-based k\_ m-anonymization of trajectory data'', in *Proc. IEEE14th Int. Conf. Mobile Data Manage. (MDM)*, vol. 2. Jun. 2013, pp. 57\_ 62.

19. V. S. Verykios, “Association rule hiding methods'',*Wiley Interdiscipl. Rev., Data Mining Knowl. Discovery*, vol. 3, no. 1, pp. 28\_36, 2013.

20. A. Miyaji and M. S. Rahman, “Privacy-preserving data mining: A game-theoretic approach'', in *Proc. 25th Data Appl. Security Privacy*,2011, pp. 186\_200.

21. Arik Friedmanetr.al.(2006)k-Anonymous Decision Tree Induction, Appeared in The 10th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD), Berlin, September 2006.

22. Abou-el-elaAbdouHussien et al.(2013)Based Privacy-Preserving: A Survey for Data Mining and DataJournal of Information Security, 2013, 4, 101-112 Published Online April 2013.

23. Bee-Chung Chen et.al.(2009)Privacy-Preserving Data Publishing Foundations and Trends in Databases Vol. 2, Nos. 1–2 (2009) 1–167 c 2009 B.-C. Chen, D. Kifer, K. Lefebvre and A. Machanavajjhala DOI: 10.1561/1900000008.

24. BENJAMIN C. M. FUNG et.al. (2010)Privacy-Preserving Data Publishing: A Survey of Recent Developments. DOI = 10.1145/1749603.1749605.