ISSN 1816-6075 (Print), 1818-0523 (Online)
Journal of System and Management Sciences
Vol. 12 (2022) No. 1, pp. 498-520
DOI:10.33168/JSMS.2022.0132

Neural Network-based Game Theory Approach for Personalized Privacy Preservation in Data Publishing

Manoj Kumar D. P.[1], Ananda Babu J. [2]

[1] Department of Computer Science & Engineering, Kalpataru Institute of Technology, Tiptur-572 201, India

[2] Department of Information Science & Engineering, Malnad College of Engineering, Hassan-573 202, India

manojkumardp@gmail.com; abj@mcehassan.ac

Abstract. There is no doubt that information sharing has become one of the prime necessities of today's society and hence it calls forth the discussions related to effective methodologies associated with privacy preserving and data publishing that assures data security and user's privacy. It's very well learnt that majority of the individuals and functional units such as research, hospital etc.… are concerned and regard Personalized Privacy as the most significant parameter. The present work emphasizes on publishing matrimonial data keeping the user's data safe and secure along with maintaining data utility. There have been various privacy models and methodologies constructed that guard against identification of user information through structure information of matrimonial data. Though they don't focus upon the user's privacy concerns and neither employ overall benefits of distributed attributes. This encourages towards building a Neural network-based game technique that attends the concern related to personalized privacy preservation and data publishing. Experiments are carried over certain real-world matrimonial datasets for analyzing the performance of the suggested technique. The results reveal algorithm's efficiency and better data utility. Moreover, privacy has been ensured without compromising on data utility.

Keywords: Personalized, Publishing, Neural network, Game Theory, privacy

1. Introduction

According to the pre-requisite of today's society, diverse platforms have been launched that caters to individual oriented data. Basically, such platforms fetch raw data from the data owners and thereafter provides value-added data services to data consumers. Mostly, the users find it unsafe to share their private information with any personalized service since such services are prone to security related threats as mentioned in (Xiao and Tao, 2006). Any information management system aims towards maintaining secrecy that safeguards the resources and data from any illegal access, maintaining integrity by preventing inappropriate and unauthorized modifications, and no occurrence of DoS (denial-of-service) by committing availability to the users. Such type of protection enforcement demands full control over the system and its resources accessibility as specified in (Elgendy et al, 2017). That is without any information leakage or breach, these systems must carry on the information sharing. With the increase in data size and rate at which the data is generating, Privacy-preserving data sharing and privacy-preserving data mining mechanisms are confronting serious issues that is mentioned in (Chamikara et al, 2019).

To resolve the above concern, data publishing with differential privacy can be achieved by agitating the data before it is published and hiding the individual's private info during statistical analysis/data mining as per (Wang et al, 2016). Around the world, people, researchers and service providers are racking their brains towards maintaining data privacy prior to publishing and towards this, the current research put forth the concept of PPDP (Privacy Preserving Data Publishing) as put forth in (Sattar et al, 2013). Various facts reveal that data publishing can hamper the individual's privacy and such threat is referred to as re-identification attack. The proposed PPDP ensures protection against such re-identification attack thereby safeguarding the essential information in the published data. Privacy preserving data publishing emphasizes on the manner in which data publishing must happen rather than doing the actual data mining task. The resultant data can then be utilized by data mining.

There prevail numerous privacy preserving data publishing approaches which are primarily of three types: 1.) differential privacy, 2.) clustering and perturbation, and 3.) generalization and suppression which being highlighted in (Lin, 2020). It's highly expected that these techniques have the potential for evaluating the prior and later adversarial belief regarding the individual's attribute values as well as the sensitivity of any identifier in privacy characterization as put forth in (Afifi et al, 2018).

The PPDP comprises of a major process known as Data anonymization that helps in hiding the user's private data along with preserving the effectiveness of data to carry out subsequent data analysis and utilization tasks. Some of the significant data anonymization techniques include k-anonymity, l-diversity, and t-

closeness that preserves privacy while data publishing. Albeit, for voluminous data, the above methodologies stand infeasible as reported in (Ouazzania and Bakkali, 2018). Attribute such as EI (Explicit Identifier) in private information helps inidentification of social security number or name of the concerned user.QIDs (Quasi Identifier attributes) depicts the background knowledge which in association with EI helps in determining user attributes like pin code, date-of-birth, and gender. Remaining sensitive information such as occupation, salary etc….is stored in Sensitive Attributes (SAs)that should remain entirely secure from unauthorized personnel as put forth in (Abdelhameed et al, 2019).

For maintaining at par data utility and protecting the information from the data receiver, the dataset is anonymized prior it is published. This enables hiding the structure attributes and append noise so that the process of data anonymization is successful. Any substitution better than this is infeasible to get amidst data utility and privacy as stated in (Fang et al, 2019).

The special purpose information metrics preserves data effectiveness for the process of data mining. Apparently, huge information is achieved by anonymization process,resulting in loss of privacy and prohibition of further such processes. The trade-off metrics takes into consideration the privacy parameters and information pre-requisite during each anonymization process which helps in determining an optimal trade-off between them according to (Fung et al, 2010). Technical tools are a great support for privacy-preserving data publishing where voluminous resource, legal regulation and surplus conventional security approaches are involved. This also emphasizes a set tradeoff amidst data privacy and utility which has been frequently mentioned in the work (Chen et al, 2009).

PPDP needs to attend the issue of trade-off amidst privacy and utility of the single and independent data as per (Shah et al, 2019). Towards this, the technique of game theory has been adopted so that PPDP can work effectively.

The mathematical structure of game theory significantly aids in reforming social dilemmas. Research pertaining to implementation of game theory of privacy protection, much emphasizes is towards information sharing, anonymity, cryptography, integrity and many more as cited in (Mengibaev et al, 2020). Though the facts reveal that in case multiple data sets are connected, the advanced privacy mechanism of differential privacy is prone to vulnerabilities. Here, the concern related to trade-off amidst privacy and utility is transformed to a game problem, wherein every player's payoff relies on its privacy parameters as well as its neighbors too, as per (Wu et al, 2020).

The Game theory model targets towards acquiring the Nash Equilibrium among all the players involved. It comprises of user's rating strategy in order to restrict the user for reforming its privacy by deviating from that point. User's strategies converge to the Nash Equilibrium point prior to the update of iterative best-response

strategy as cited in (Halkidi and Koutsopoulos, 2011).

PPDP process comprises of data collection and data publishing phases. These phases refers to the following roles:

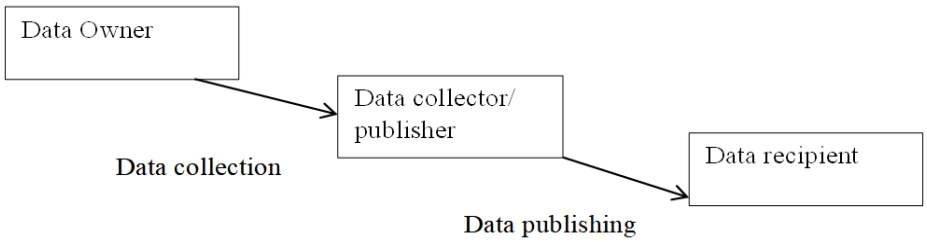


Fig. 1: Two phases of PPDP

● Data owner: they own the data which is then forwarded to the data collectors.

● Data Collector: They acquire the data from the data owners which then undergoes the process of PPDP and the resultant modified data is transmitted to the data user.

● Data User: the modified data then undergoes the process of data mining when it's received by the data user. Moreover, the data collector is offered incentives for obtaining the modified data. Once received, the data user carries out the task of data mining on the modified data fetched from the data collector. Mining the database results in quiet significant information.

On the other hand, the data publisher must ensure and is allowed to publish the data if it complies with the given criteria. That is, there should be no identifying attribute in the published dataset as it can hamper and destroy the individual's privacy.

Even the neuroscience domain has witnessed the approach of game theory. On the other hand, neuroeconomics, has experimented by integrating human and nonhuman players for comprehending the approach of human decision-making, as cited in (Schuster and Yamaguchi, 2010). These security techniques has helped in building personalized information services, enabling various business prospective without any threats or illegal intrusion to one's personal information.

The present research put forth the idea of neural network model in accordance with the innovative game theory approach that involves neurons behaving optimally according to the given payoff matrix.

2. Related Works

The work aims towards establishing database privacy through the execution of a game in which two function sets are being balanced: first is the hiding of the "private" functions and second is the "information" functions that needs to be

revealed. Publishing data in a hostile region is quiet challenging and requires effective and robust techniques which can protect the individual's privacy as well as publish useful data effectively. There are diverse techniques which have been recommended by the researches lately which ensures in attending the above issue. An individual's identities are composed in form of raw data referred to as micro data which is kept hidden and unpublished for ensuring privacy.

(Xiao and Tao, 2006) Has put forth a generalization framework that is founded on the personalized anonymity approach. The framework carry out minimum generalization that fulfills all possible requirements, thereby retaining maximum amount of information from the microdata. The standard concept of Generalization helps in transforming the identifying values into less specific forms through which the individual's identity is not exhibited precisely as stated in (Wang et al, 2018). There is also emphasis on the privacy-preserving personalized recommendation services through which the present work is classified based on the fundamental methodologies for protecting privacy and personalized recommendation. This is followed by in-depth discussion and comparison of their advantages and disadvantages.

(Elgendy et al, 2017) It is been proposed the PPDP (privacy preserving data publishing) model by integrating PBFW, CPBAC, MD-TRBAC thereby safeguarding database administrator technique stimulated by the oracle vault technique and including merits of anonymization technique for safeguarding published data through k-anonymity. (Chamikara et al, 2019) introduces SEAL (Secure and Efficient data perturbation Algorithm utilizing Local differential privacy) that depicts a data perturbation algorithm based upon Chebyshev interpolation and Laplacian noise which effectively balances privacy and utility.

(Dinur and Nissim, 2003) Suggests that for implementing reconstruction algorithm to statistical databases there should be appending of perturbation magnitude. As smaller the perturbation, higher will be the privacy violation. This query restriction approach enforces the queries to comply with a special structure, for avoiding querying adversary in obtaining large information pertaining to specific database entries. The concept of query auditing involves, storing the query log, verifying the new query for any compromise, and based on that allowing/disallowing the query as mentioned in (Wang et al, 2016). There is discussion and analysis regarding the issue of real-time spatio-temporal data publishing in OSNs involving privacy preservation in accord with continuous publication of population statistics. Moreover, there is proposal of a design Rescue DP, which being an online aggregate monitoring framework across infinite streams with assurance of w-event privacy.

(Kavianpour et al, 2019) Strongly asserts the occurrence of privacy threats over the OSN and towards confronting the same recommends a privacy-preserving model for carrying out social engagements amidst the users and third-parties. The

aim being to strengthen users' privacy by granting data access to the relevant third-parties. (Sattar et al, 2013) Has put forth a novel approach of PPDP (privacy preserving data publishing) that regards the adversary inferring a sensitive value from published dataset as high compared to the inference according to public knowledge. (Lin, 2020) Has reviewed the concerns associated with protecting privacy in trajectory datasets from adversaries that can misuse this partial knowledge for getting unknown locations. Towards this, a tree-based indexing structure has been recommended that stores complete trajectory data. In addition pruning strategies have been built along with two algorithms for identifying secure option against the actual trajectory dataset.

Data anonymization as yet another privacy preserving approach suggested by the researches as mentioned in (Afifi et al, 2018). Though its quiet evident from recent work that the privacy of anonymized data is prone to vulnerabilities. (Ouazzania and Bakkali, 2018) Has put forth a technique to imbibe k-anonymity for quasi-identifier attributes using the algorithm referred to as "k-anonymity without prior value of the threshold k".

(Bayardo and Agrawal, 2005) Introduces an optimization algorithm called "k-anonymization" towards the robust process of de-identification. It discovers space of possible anonymizations to overcome the combinatorial issue, and built data-management techniques for minimizing costly operations of sorting. There exist major computational concerns even with basic constraints of optimized k-anonymity that being NP-hard. (Mehta and Rao, 2019) Has put forth the approach of Improved Scalable l-Diversity (ImSLD) oriented on scalable k-anonymization and an extended version of Improved Scalable k-Anonymity. It employs MapReduce as a programming model. The novel approach of Restricted Sensitive Attributes-based Sequential Anonymization (RSASA) has been suggested by (Abdelhameed et al, 2019) towards privacy-preserving data stream publishing. In addition, there is suggestion of two new privacy restrictions namely, Semantic-diversity and Sensitivity-diversity for prohibiting the published Sensitive Attributes values.

(Li et al, 2020) A new prefix tree structure along with the models of incremental privacy budget allocation and spatial-temporal dimensionality reduction for improvising the sanitized data utility along with accelerating the runtime efficiency. Besides, there is examining of the publication of non-interactive sanitized trajectory data based upon Differential Privacy. A graph data anonymization approach has been put forth by (Fang et al, 2019) via GDAGAN (Generative Adversarial Network) that broadly comprehends the data attributes thereby generating anonymous graph data possessing equivalent feature distribution of the actual. Differential privacy noise is appended to the generated process for building anonymous graph data.

(Li et al, 2013) Positive membership privacy that prohibits the adversary from making increased attempts in deciding if the entity belongs to the input dataset. It also introduces negative membership privacy that prohibits leaking of non-membership. So, it's essential that the privacy measure thoroughly safeguards everyone in the anonymized dataset towards membership disclosure. This kind of privacy definition must also specify the adversary's prior knowledge regarding what the dataset contains.

(Skarkala et al, 2012) Have suggested an anonymization approach for weighted graphs (social networks) which place high significance on the link's strength unlike the prior studies which placed importance on just the unweighted graphs. The proposed method offers k-anonymity of nodes against threats, with the adversaries' possessing information regarding the network's structure and its edge weights. (Fung et al, 2010), (Chen et al, 2009), and (Shah et al, 2019) have well elaborated and analyzed diverse PPDP techniques, overviewed the hurdles confronted in actual data publishing, resolving dissimilarities and pre-requisites against other relevant issues and forecasting future research options.

The overall approach of Game theory stands formal against model circumstances wherein the agents need to select ideal actions in accord with the mutual effects of remaining agents' decisions. For handling privacy concern in in data-driven applications, diverse game theoretical techniques are recommended. (Shah et al, 2019) Has put forth varied scope of game theoretical techniques in the field of privacy preservation, network security, intrusion detection and optimization of resources. (Mengibaev et al, 2020) Proposes the evolutionary game theory approach for exploring the concern associated with privacy protection in OSNs along with the effect of heterogeneous interaction dependency strength on privacy protection.

(Wu et al, 2020) Has recommended a differential privacy approach that confronts vulnerability in case multiple data sets are interlinked. Here, the concern related to trade-off amidst privacy and utility is transformed to a game problem, wherein every player's payoff relies on its privacy parameters as well as its neighbors too. Unlike encryption, differential privacy tends to be a rather simplistic privacy preservation approach. (Xu et al, 2015) Has represented the association amidst different players namely the data user, data providers and data collectors in form of a game thereby determining the Nash equilibriums by considering k-anonymity as the anonymization method.

(Halkidi and Koutsopoulos, 2011) Has made use of the game theory for overviewing the users' interaction so as to determine conditions and expressions for NEP (Nash Equilibrium Point). It includes user's rating strategy so that there is no reformation in its privacy by drifting from the NEP. The rating strategies of every user converge to the Nash Equilibrium point prior to the update of iterative best-

response strategy. (Arpitha, 2018) Has suggested a system based upon the Nash equilibrium approach of game theory that ensures data privacy protection by transforming model to other modules by computing the best payoff.

For overcoming the constraints of k-anonymity (Chakravarthya et al, 2012) has put forth an approach based on the Coalitional Game Theory. As a result, the privacy levels can be reformed as per the given threshold for the information loss. For obtaining anonymity, the agent represents a tuple and is allocated payoff through CHT (concept hierarchy tree) of QID (quasi-identifiers) for the available data; and coalitions are built among the tuples as per their payoffs.

(Schuster and Yamaguchi, 2010) put forth the idea of neural network model in accordance with the innovative game theory approach that involves neurons behaving optimally according to the given payoff matrix. Also, it evaluates a paired neuron system and acclaims that the value game theory can possess an organizing standard for the said system. (Zapechnikov, 2020) Has studied the essential components pertaining to the Zero knowledge proofs, secure multi-party calculations, and holomorphic encryption are examples of cryptographic data processing protection.The differential privacy mechanism, which is the cornerstone for privacy-preserving machine learning and certain cryptographic techniques, has also been reviewed. (Andrew et al, 2019) Has proposed the Mondrian-based k-anonymity strategy, which allows for a trade-off between user privacy and data utility, as well as the DNN (Deep Neural Network) framework for protecting the privacy of high-dimensional data has suggested using a deep learning model. (Rasim et al, 2019) for privacy preservation analysis with a two- stage architecture. A modified sparse deionizing auto coder has been employed by the proposed model that enables data transformation and a CNN classifier for classification of data.

(Stier et al, 2018) Has looked at game theoretic measurements like the Shapley value for separating relevant from irrelevant ANN segments (artificial neural network). Towards this, a coalitional game has been built for ANN, wherein coalitions are formed by the neurons and the Shapley value is the average contribution of neurons to coalitions. By removing the low-contributing neurons and assessing the effect on network performance, the contribution of individual neurons can be measured by Shapley value.

3. Proposed Work

3.1. Overview

The research proposes a neural network model based on game theory for matrimonial data publishing through the Advanced SVM based personalized privacy preservation. Game theory for matrimonial datasets comprise of key parameters such as name, DOB, Gender, etc. Also, there are modeling strategies and results based on certain rules, involving 2 or more players in a game. The players in

the game theory are referred to as client of matrimonial site and the actions depicts game which basically searches to identify their pair. The idea of effectiveness reward and/or punishment incentives tends to be quiet popular for elevating cooperative behavior between the participants.

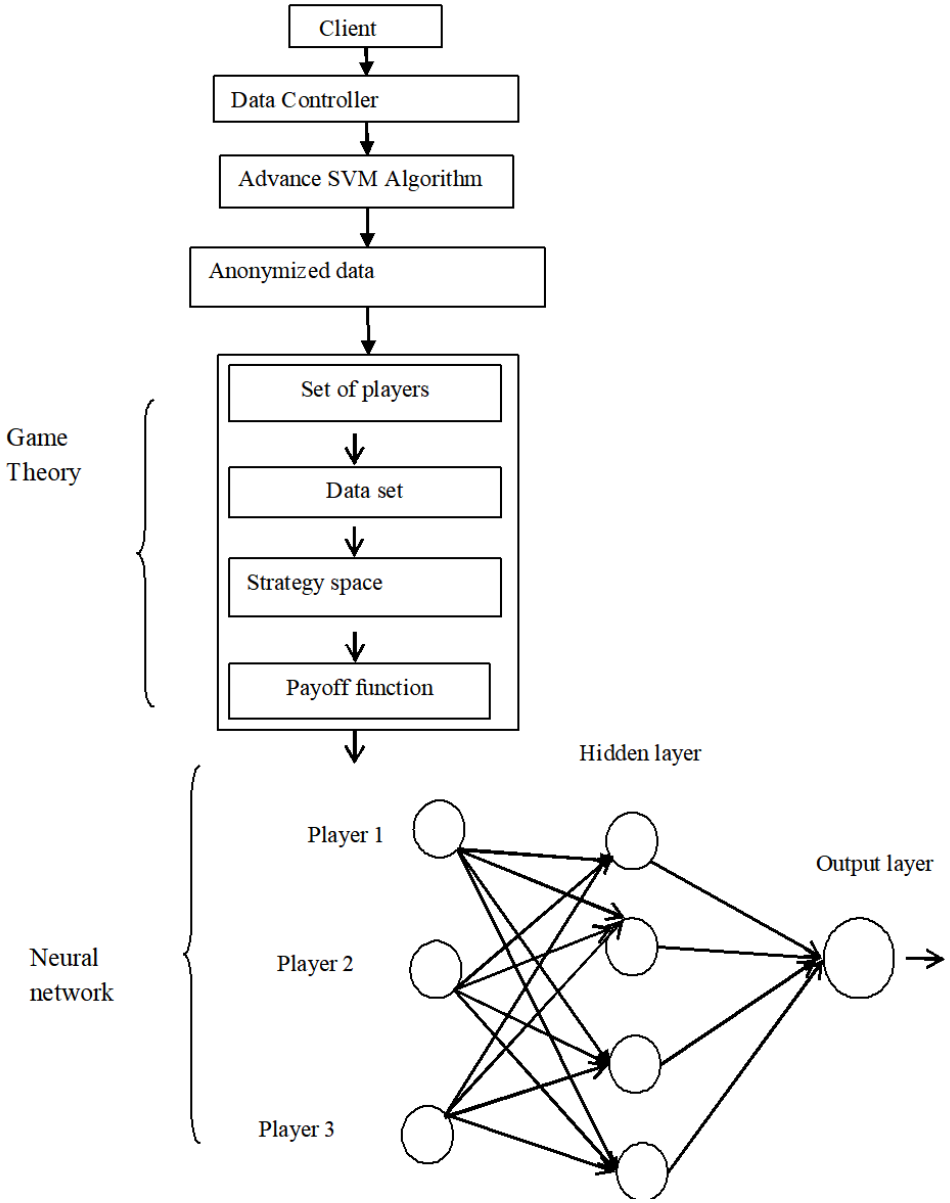


Fig. 2: Overall Architecture

A game model involves many players wherein the users can publish their personal data sets which are sanitized by anonymization approaches. The data

controller gathers user's data and classify it via Advance SVM (Support vector Machine) which aids in resolving the issue of classification. By the means of anonymized approach user's private data remains secure which is then published to the third party. Game analysis-based model depicts adequate criteria for the availability and distinctness of the pure NE (Nash Equilibrium) in the game. Matrimonial datasets that comprises of independent data, privacy of user's data relies on the user's privacy parameter as well as its neighbor's too. Also, the privacy choice is switched from the issue of tradeoff to game. In the game theory, the matrimonial site's users represent the players and their personalized privacy parameters as well as the private settings represent strategies. By using the payoff functions, a balance is maintained amidst the data utility. The proposed privacy-preserving approach in accord with AVSMclassification of data is successful enough in numerous real-world applications. Moreover, the neural network model is trained and tested to build diverse approaches for privacy preserving of the matrimonial dataset via learning. Also, it could be very well extended in the domain of privacy preserving data publishing.

3.2. Game Theory

The Game Theory considers all the individuals as "rational". In other words, there are person-specific aims of each individual towards which each individual would employ the most ideal technique to achieve them. The Game Theory comprises of:

- Player Identification and evaluating the payoff for each player.
- Employing equilibrium strategies for making descriptive or prescriptive predictions.

Above all, the Game Theory aims towards deriving solutions that fits ideal for all the players.

Following are the standard terminologies in a game theory:

- Players: they take decisions and performs action when Engaging with others while adhering to specified techniques that are bound by payoffs specific to each player
- Payoffs: indicates how satisfied each player is after making a decision. Depending on the player's decision, this measure can be either negative or positive.
- Strategy: depicts the player's strategy throughout the duration of the game. The player chooses the next action based on the strategy, which is frequently based on the opponent's previous acts.
- Nash Equilibrium: Nash Equilibrium depicts an action profile, adhering to which no player can perform a different action in order to gain better results. Alternatively, it denotes all the players' common and best responses.

Minimax and Maximini principle: Mini-max algorithm resembles a recursive algorithm employed in decision-making and game theory. It enables the player to

take an optimal move considering that the opponent's move is optimal too. In involves two players, one called MAX and other called MIN.

Figure 3 projects the general flow of game theory, wherein the NEP (Nash Equilibrium Point) comply with the optimal strategy of the players.

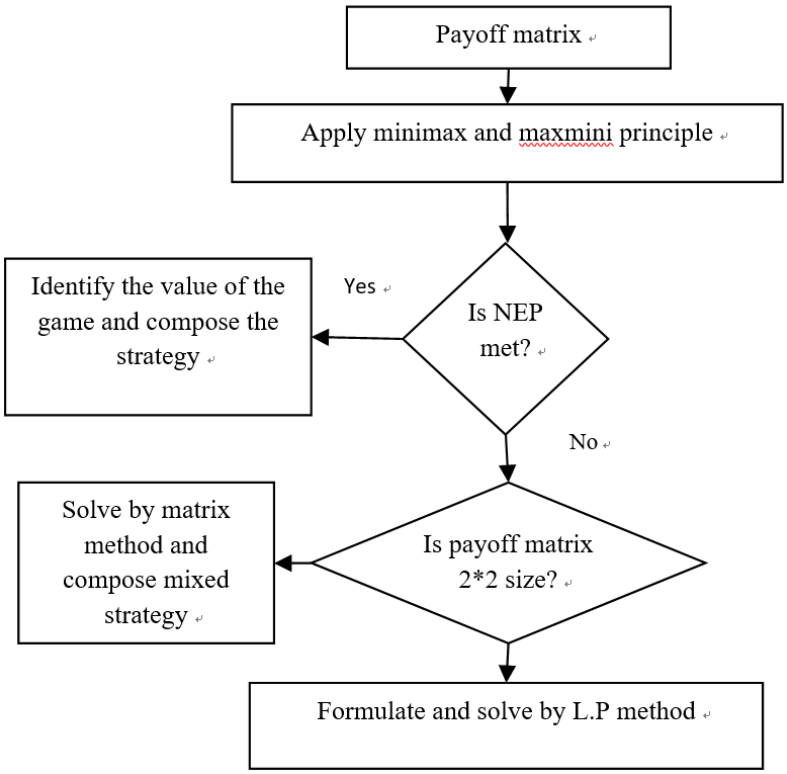


Fig. 3: Game Theory

In game theory, the playoff matrix depicts a table wherein a single player's strategies are recorded in lines and the remaining others are recorded in sections. The cells reveal playoff to every player with the conclusion that the result of the column player is recorded first. Subsequently the minimax and maximini principle is applied thereby verifying criteria with Nash Equilibrium. The playoff parameter is of utmost importance as it aggregates the fundamental data and enables to take a decision whether there prevails a predominant procedure or potentially Nash equilibrium. Post NEP verification, the strategy value game is found. In case of a 2*2 payoff matrix, the matrix method and compose mixed strategy is employed for computation. Figure 3 depicts the overall formulated and computed value.

3.3. Game Model for Privacy Data Publishing

A finite game *GM = (P, S, A)* consist of,

- A finite set of players (data publishers) P = {1, ..., n};
- Each player has a finite strategy space Fi.$i \in P$;
- Payoff is$a_i(f) : F \rightarrow X^+$ for each and every result outcomes.$f \in F = F_1 \times \cdots \times F_n$ and $A = \{a_1, \cdots, a_n\}$.

3.4. Pure Nash Equilibrium

In a finite game *GM = (P, S, A)*, the *strategy* $s* = (s*1, \cdots, s* n)$ is a pure Nash Equilibrium if and only if, for each player i and each feasible strategy s_i in S_i, $(s*i, s*-i) \gtrsim (s_i, s*-i)$. Using the Nash Equilibrium one can precisely assume what will happen if there is interaction among multiple. If *D* is the data set then for each player *i* then $D_i = 1...., n$. The strategy space Si of player i is the set of privacy parameter $e_i \in R^+$. And, some strategy $s_i \in Si$ of player *i*, is equal to e_i respectively.

A set of players - *N* and the payoff function - $a_i(s)$ allocates a value to each player stating the payoff of the players that can be moved for training the neurons within a network.

For each player *i*, its payoff includes two components, 1.) Utility of the anonymized dataset -$U_i(s)$ and 2.) Loss due to privacy leakage - $L_i(s)$. Represented as following:

$ai(s) = Ui(s) - Li(s).$

For building a game model, the below mentioned factors must be precisely considered.

(i) Association between different data sets under consideration with a different privacy parameter for each data set owner,

(ii) utility measures - $Ui(\cdot)$ of anonymized data and

(iii) Loss measures $Li(\cdot)$ in case of a privacy leakage.

The prevailing issue is framed in form of a sequential game wherein the data is anonymize by the data collector for securing the privacy of data providers'. The game model helps in determining the Nash equilibrium based on which the amount of anonymization to be applied on the data can be determined so as an equilibrium is maintained amidst the data collector and the data user.

Let's consider a game comprising of 3 player/person. Each of them is given a set number of pure actions. That is, 1st, 2nd, 3rd players have pure actions denoted by m, n, and q, respectively. The payoffs are represented by 3-dimensional matrices. 3rd Player selects a tri-matrix, 1st Player selects row, and 2nd Player selects a column, which is followed by the payoffs. For instance, let's assume a 3-person game with each one of them allotted two pure actions and payoffs given by the trimatrices.

A=l A=2 B = 1 [(0, 0,1) (0,1,0) 1 B=2 (1,0,0) (0,1,0,) X=l

A=l A=2 B=l [(0,0,2) (1,0,0) 1 B=2 (0,1,0) (0,1,0) X=2

The above one depicts 3 player, each of whom can hold one or two coins. If equal no: of coins is taken by all the players, then each player has an output of payoff 0. The person who has a different number of coins than the other players receives a payment equivalent to the number of coins in his hand, while the other players earn nothing.The Nash equilibrium of this game's mixed extension has also been calculated. [0,1] x [0,1] x [0,1] x [0,1] x [0,1] x [0,1] x [0,1] x [0,1] x [0, R is assigned to player i.

x (l - V) (l - z) + 2(1 - x) vz K1x, z, v = x (l - V) (l - z) + 2(1 - x) vz

2 - (1 + V) (l + z)x + 2vz = K2 x, z, v = (2 - (1 + V) (l + z)x + 2vz

V (l - x) (l - z) + 2 K3 x, z, v = V (l - x) (l - z) + 2 (1 - V) (2 - (1 + x) (l + z) = x z z (l – x) (l – V) + 2(1 – z) V + 2xz = z (l – x) (l – V) + 2(1 – z) XV = z + 2xV = (2 - (1 + x) (l + v)

I an xv-plane piece limited by (l+x) (l+ V) = 2 and the lines V = 1, x = 1, and z=1

One interior point (J2 - 1, J2 - 1, J2 - 1) and six line segments [(0,0,1), (1,0,1)] make up the intersection of these graphs. [(0,1,0), (1, 1,0)] [(0,1,0), (1, 1,0)] [(0,1,0), (1, 1,0 '[(0,0,1), (0, 1, 1)], [(1,0,0), (1, 1,0)], [(1,0,0), (1,0,1)], [(1,0,0), (1,0,1)], [(1,0,0), (1,0,1)] '(0, 1, 0), (0, 1, 1) The set of Nash equilibrium coincides with this intersection. Every line segment has one participant who is unconcerned because the reward is always zero.In the symmetric Nash equilibrium, the payoff for each player is 6 - 4J2.

3.5. Artificial Neural Network

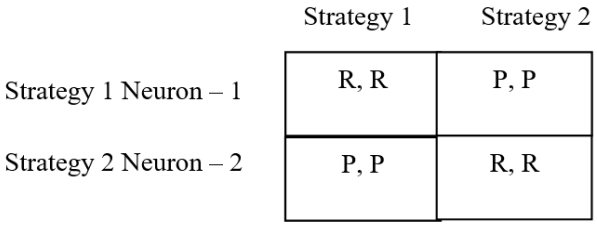
The purpose of ANN (Artificial Neural Networks) is to comprehend from data, develop new information based on learning, and make use of innumerous variables. The concept behind ANN model is to imitate brain on computers, emphasizing on the mathematical modeling of biological neurons thereby simulating the working of human brain in an easy manner. System formulated in ANN model employs the functions such as sigmoid, threshold, hyperbolic tangent and linear activation. Based on the dataset distribution, one of the above activation functions is chosen. There is automatic updating of the weight values till the target output values are obtained as per the learning rules. Once the training process gets over, the network classifies the test dataset along with the final weight value.

3.6. Game Theoretic and Neural Network Interpretations

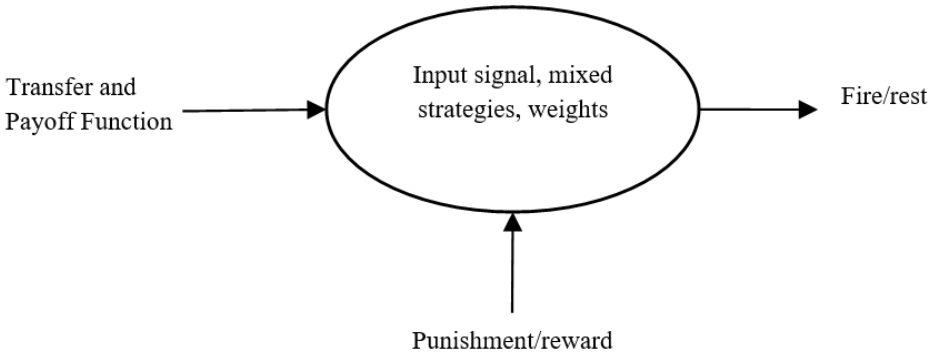
The game theory comprises of the players, a set of rules for playing the game, and a result or the payoff matrix which can be either a reward or a punishment for the player, depicting the dynamic performance of the game. In a neural network, the neurons are of competing and collaborating nature towards which the Game Theory offer feasible theoretical background for precisely choosing the paramount “player” in a game. Figure 4 portrays the implementation of these important approaches to

neuron system in which the neurons aids in computing their strategies as per their individual payoff matrix.

Figure 4(A) depicts the behavior patterns of the payoff matrix, in which the payoff for each player is allocated to each neuron as either a reward- R or a punishment- P for diversified strategies.Figure 4(B) portrays the model comprising of elements from game theory. In a derivation process, the neural network structure is built from the top down using a top-down approach and an initial root model. This results in the creation of a trained ANN model with the fewest number of players (i.e., neurons) when compared to the initial root model.



(A) Artificial Neurons



(B) Game Theory

Fig. 4: Interpretation between

3.7. Advanced Support Vector Machines (ASVM)

Usually, Game-theoretical computations are often intrinsically distributed with close interaction amidst players without the need of a focal position. In addition, these computations mostly iterative rather than involving broad retraining. It can be deduced that the game theoretical application offers an innovative and elective approach for handling concerns related to voluminous information. The present research exhibits the ASVM (Advanced Support Vector Machine) computation that helps to resolve the amplified classification issues in networked systems by collecting and categorizing. Its implementation could also help in depicting

registering stages thus revealing how the game-theoretical variation of ASVM can be further imbibed. The ASVM Game depicts a two-player iterated game in which the players refers to the data patterns. While dealing with critical data and the information of a trained model, the data privacy must be preserved to prevent against various malicious attack. Towards this, the research recommends a security-preserving learning framework via support vector machine model ASVM Game. The framework aspires towards achieving best possible output by defining the players' actions and its subsequent effects. The player don't have the liberty to choose the desired actions (pass or hold), in fact their actions are ruled by their closeness to the reference point of the opposite player. That is, the action of the players indicate or refers to the distinctive properties of the players.

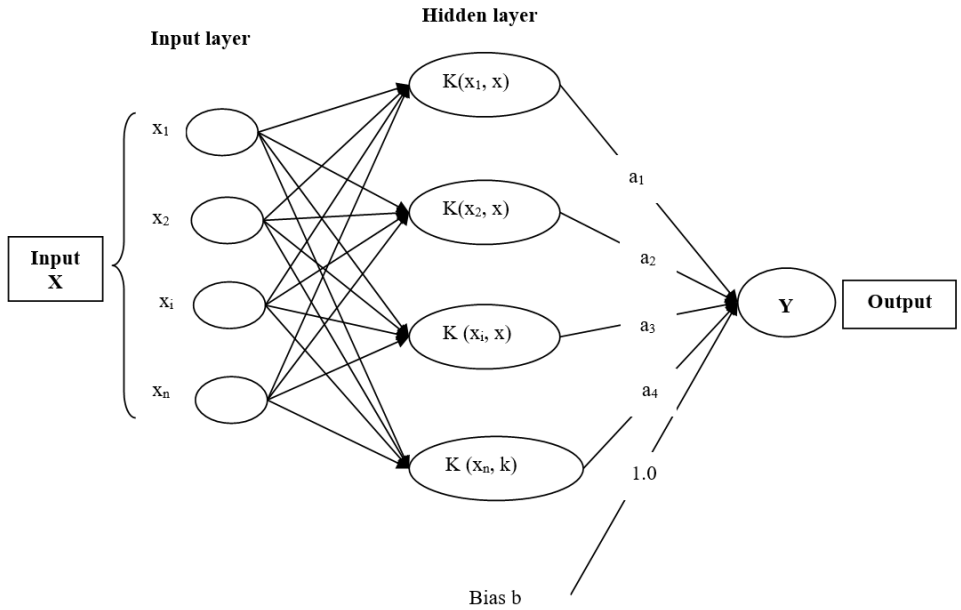


Fig. 6: Architecture of ASVM

For achieving paired characterization, the technique of ASVM (Advanced Support Vector Machine) prove significant and outstanding. It mainly emphasizes towards determining the best hyper plane that segregates information into its two classes in a proficient manner. Through combining multiple parallel ASVM, multiclass grouping has been achieved. Figure 6 presents the ASVM design. Leave 1 alone the preparation occasions $\{X_i, Y_j\}$, i=1,…,l each example holds an information xi and a class label Yi {-1,1}. For defining a hyper plane, weight vector (w) and a bias (b) is utilized, wherein N depicts the number of preparing occasions, xi depicts the contribution of preparing occurrence and y_i is its relating class mark, b is a bias, and $K(x_i, x)$ represents the pre-owned bit work that maps the information vectors into an extended element space.

4. Results and Discussion

The research put forth a novel approach that highlights individual neurons within the game theoretical framework in accordance with ASVM. While preserving the matrimonial data of any user, only handpicked information is published and the sensitive information is kept hidden or preserved. Every user is offered with a personalized privacy metric using which the user can define its privacy level as per its priority.

The game problem attends the trade-off concern amidst privacy and utility wherein each player depicts a user who has built its profile as per a certain privacy level. The game theory helps in achieving the payoff functions, payoff matrix and the mixed strategies. By the means of an ANN (Artificial Neural Network) model and output gained through the game model, multiple approaches can be built for privacy via learning. Apparently, the number of neurons is same as the number of players and depending on the activation function, the neural network update its weights values until its obtains the desired result as per the learning rules. Post completion of the training process, the test data can be classified by the network along with the final weight. Minigames catching the pith of Public Goods depicts that in the absence of discernment suppositions, punishment and reward will avoid obtaining the prosaically conduct.

Every player in a game problem is allocated a payoff function in form of either reward or punishment for the strategy they adopt. Also, reward of any player relies on their neighbors. Each user's privacy relies upon the other user's privacy level. That is, if any user prevents revealing its family details publically, then that particular user can't see family details of other users too.

Table 1: Data set

Explicit identifiers		**Quasi identifiers**		**Sensitive attributes**		
Name	**Gender**	**Age**	**D.O.B**	**Address**	**Education**	**Occupation**
ABC	M	29	1992	A	S	K
PQR	F	27	1994	B	T	L
XYZ	F	25	1992	D	U	M

Table 1 presents the Matrimonial dataset comprising of explicit identifiers, quasi-identifiers and sensitive attributes. The data holder is responsible for publishing huge amount of information pertaining to various participants. It's made certain that the relationships amidst the quasi-identifiers and sensitive attributes

such as address, education and occupation remains secure and private. Although, every user or player's privacy parameter vary for keeping its sensitive attributes hidden. The existing dataset has date of birth, address and occupation as sensitive attributes. Table 3 represents the two variance anonymized data of quasi-identifier having similar sensitive attributes.

Table 2: Anonymized data

Explicit identifiers		**Quasi identifiers**		**Sensitive attributes**		
Name	**Gender**	**Age**	**D.O. B**	**Address**	**Education**	**Occupation**
PQR	M	[29-30]	[1991-1992]	B	SS	M
XYZ	F	[27-28]	[1993-1992]	D	TU	K
ABC	F	[25-26]	[1991-1992]	A	US	L

Table 3: Variance anonymized data

Name	**Gender**	**Age**	**Education**
PQR	M	[29-30]	SS
XYZ	F	[27-28]	TU
ABC	F	[25-26]	US

These sensitive attributes are kept hidden prior to the dataset being published. Table 3 depicts the resultant details after the attributes are suppressed.

The security level of collected possible output corresponds to the value of un-good result. This implies that the approximation of any result is in accordance with the security level. In a game play framework, any user's security level depicts or appears to others players as the choice of that particular user. Such a loosened scenario of the game involves players trying to hamper each other, though in actuality they may not hold such nature. Consider the case where Z depicts a by b result grid, Sl depict level of security (1^{st} player) I column, at that point.

$Sl := \min \{Z(x,y): x=1,2,3...,b\}$, $y = 1,2,3...,a$.

Payoff value is approximately:

The security levels for Player1

$S_{1st} := \min \{4.5 , -2.5, 8 \} = -1.5$

$S_{2nd} := \min \{10 , 3.8, 7\} = 4.8$

Similarly, Player 2's security levels are described as follows:

Sl:= min {Z(x,y): x=1,2,3...,b} , y=1,2,3...,a.

S_{1st} := min {25 , 8, 15 } =8

S_{2nd} := min {-2.5 , 3, 7 } = 3

S_{3rd} =min{6, 4, 3}=3

Therefore, player 2 has the ideal security level as 4, which becomes terrible all the more in case segment 2 is selected. By annexing the levels of security to the payoff matrix, the analysis can be summarized:

	2nd Player			SP1
1st Player	4.5	- 2.5	8	- 1.5
	9	4	3	3
SP2	3.5	- 1.5	3	

Let's consider that every single player attempts towards optimizing its own security level (SP1, SP2): that is Player 1 tries to increase its security level and so does Player 2. As per this example, Player 1 opts for 2ndrow and Player 2 opts for 2nd column.

Table 4: Security data of Player

No.	1	2	3
Player 1	9	4	3
Player 2	1.5	4	1.5

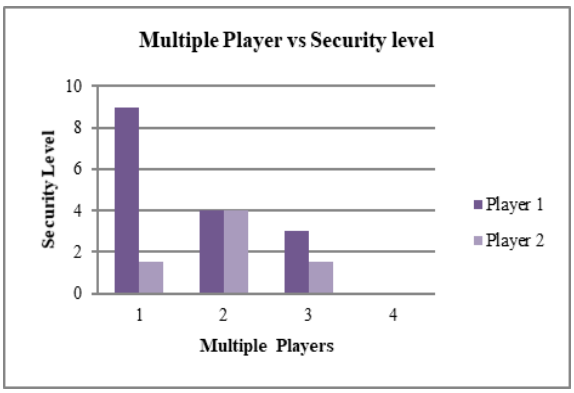


Fig. 7: Comparison of Player Security

Player 1 increases its security level, while Player 2 increases its own security level, as seen in Fig 7. That is, the privacy level of the players is displayed using the values of Sp1 and Sp2 of optimal security in line with the above-mentioned security data.

5. Conclusion

The present research work has put forth a novel game theoretic based neural network model that enables publishing of matrimonial data by maintaining data utility and most importantly fulfilling the need of personalized data privacy. To accomplish the same, a game model has been built using the ASVM (Advanced support vector machine) of multiple players, each of who attempt to publish the data set. Thereafter, there is representation of adequate criteria for the availability and uniqueness of the pure NE (Nash Equilibrium). In addition, by employing the anonymization approach, sensitive information have been successfully protected from other participants. By the means of game theory, privacy of the matrimonial data has been achieved. Also, by using the sequences of the game theory process, the ANN (Artificial Neural Network) is being trained. This makes it quiet evident that the strategic game theoretic approach is most ideal for building the behavior of neural network in terms of privacy preservation techniques in the domain of data publishing.

At last, various experiments have been carried out on certain real-world matrimonial datasets for assessing the practical proficiency of the proposed model. It has been well proved from the results obtained the proposed algorithm offers effective and robust personalized data privacy to the users as well as helps in uplifting the data utility. In a nutshell, the information privacy has been significantly achieved without having to compromise on the data utility.

References

Xiaokui Xiao and Yufei Tao. (2006). Personalized Privacy Preservation. *Proceedings of the 2006 ACM SIGMOD international conference on Management of data*, 229-240. DOI: https://doi.org/10.1145/1142473.1142500

Cong Wang, Yifeng Zheng, Jinghua Jiang, Kui Ren. (2018). Toward Privacy-Preserving Personalized Recommendation Services. *Engineering*, 4(1), 21-28.

Rana Elgendy, Amr Morad, Hicham G. Elmongui, Ayman Khalafallah, Mohamed S. Abougabal. (2017). Role-task conditional-purpose policy model for privacy preserving data publishing. *Alexandria Engineering journal*, 56(4), 459-468.

M.A.P. Chamikara, P. Bertok, D. Liu, S. Camtepe, I. Khalil. (2019). An efficient and scalable privacy preserving algorithm for big data and data streams. *Computers & Security*, 87, 101570. DOI: https://doi.org/10.1016/j.cose.2019.101570

Irit Dinur and Kobbi Nissim. (2003). Revealing Information while Preserving Privacy. *Proceedings of the twenty-second ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, 202-210. DOI: https://doi.org/10.1145/773153.773173

Qian Wang, Yan Zhang, Xiao Lu, Zhibo Wang, Zhan Qin, Kui Ren. (2016). Real-time and Spatio-temporal Crowd-sourced Social Network Data Publishing with Differential Privacy. *IEEE Transactions on Dependable and Secure Computing*, 15(4), 591-606.

Sanaz Kavianpour, Ali Tamimi, Bharanidharan Shanmugam. (2019). A privacy-preserving model to control social interaction behaviors in social network sites. *Journal of Information Security and Applications*, 49. DOI: https://doi.org/10.1016/j.jisa.2019.102402

A.H.M. Sarowar Sattar, Jiuyong Li, Xiaofeng Ding, Jixue Liu, Millist Vincent. (2013). A general framework for privacy preserving data publishing. *Knowledge-Based Systems*, 54(C), 276–287.

Chen-Yi Lin. (2020). Suppression techniques for privacy-preserving trajectory data publishing. *Knowledge-Based Systems*, 206. DOI: https://doi.org/10.1016/j.knosys.2020.106354

M. H. Afifi, Kai Zhou, Jian Ren. (2018). Privacy Characterization and Quantification in Data Publishing. *IEEE Transactions on Knowledge and Data Engineering*, 30(9), 1756-1769.

Zakariae El Ouazzania and Hanan El Bakkali. (2018). A new technique ensuring privacy in big data: K-anonymity without prior value of the threshold k. *Procedia Computer Science*, 127, 52–59.

R. J. Bayardo and Rakesh Agrawal. (2005). Data Privacy Through Optimal k-Anonymization. *Proceedings of the 21st International Conference on Data Engineering (ICDE 2005)*, DOI: https://doi.org/10.1109/ICDE.2005.42

B. B. Mehta and U. P. Rao. (2019). Improved l-diversity: Scalable anonymization approach for Privacy Preserving Big Data Publishing. *Journal of King Saud University – Computer and Information Sciences*, In Press. DOI: https://doi.org/10.1016/j.jksuci.2019.08.006

S. A. Abdelhameed, S. M. Moussa and M. E. Khalifa. (2019). Restricted Sensitive Attributes-based Sequential Anonymization (RSA-SA) approach for privacy-preserving data stream publishing, *Knowledge-Based Systems*, 164, 1-20. DOI: https://doi.org/10.1016/j.knosys.2018.08.017

Yang Li, DasenYanga, Xianbiao Hu. (2020). A differential privacy-based privacy-preserving data publishing algorithm for transit smart card data. *Transportation Research Part C: Emerging Technologies*, 115. DOI: https://doi.org/10.1016/j.trc.2020.102634

Junbin Fang, Aiping Li, Qianyue Jiang. (2019). GDAGAN: An Anonymization Method for Graph Data Publishing Using Generative Adversarial Network. *Proceedings of 6th International Conference on Information Science and Control Engineering (ICISCE)*. DOI: https://doi.org/10.1109/ICISCE48695.2019.00068

Ninghui Li, Wahbeh Qardaji, Dong Su, Yi Wu, Weining Yang. (2013). Membership Privacy: A Unifying Framework for Privacy Definitions. *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*, 889-900. DOI: https://doi.org/10.1145/2508859.2516686

Maria E. Skarkala, Manolis Maragoudakis, Stefanos Gritzalis, Lilian Mitrou, Hannu Toivonen, Pirjo Moen. (2012). Privacy Preservation by k-Anonymization of Weighted Social Networks. *Proceedings of 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 423-429. DOI: https://doi.org/10.1109/ASONAM.2012.75

Benjamin C. M. Fung, Ke Wang, Rui Chen, Philip S. Yu. (2010). Privacy-Preserving data publishing: A survey of recent developments. *ACM Computing Surveys*, 42(4), 1-53.

Bee-Chung Chen, Daniel Kifer, Kristen LeFevre, Ashwin Machanavajjhala. (2009). Privacy-Preserving Data Publishing, *Foundation and Trends in Databases*. 2(1-2), 1-167. DOI: https://doi.org/10.1561/1900000008

Hitarth Shah, Vishruti Kakkad, Reema Patel, Nishant Doshi. (2019). A survey on game theoretic approaches for privacy preservation in data mining and network security. *Procedia Computer Science*, 155, 686–691. DOI: https://doi.org/10.1016/j.procs.2019.08.098

Ulugbek Mengibaev, Xiaodan Jia, Yeqing Ma. (2020). The impact of interactive dependence on privacy protection behavior base d on evolutionary game. *Applied Mathematics and Computation*, 379. DOI: https://doi.org/10.1016/j.amc.2020.125231

Xiaotong Wu, Taotao Wu, Maqbool Khan, Qiang Ni, Wanchun Dou. (2016). Game Theory Based Correlated Privacy Preserving Analysis in Big Data. *IEEE Transactions on Big Data*, 7(4), 643-656.

Lei Xu, Chunxiao Jiang, Jian Wang, Yong Ren, Jian Yuan, Mohsen Guizani. (2015). Game Theoretic Data Privacy Preservation: Equilibrium and Pricing. *Proceedings of The 2015 IEEE International Conference on Communications (ICC)*. DOI: https://doi.org/10.1109/ICC.2015.7249454

Maria Halkidi and Iordanis Koutsopoulos. (2011). A Game Theoretic Framework for Data Privacy Preservation in Recommender Systems. *Proceedings of the 2011 European conference on Machine learning and knowledge discovery in databases*, Volume Part 1, 629–644.

D. G. Arpitha. (2018). A Game Theory Approach to Preserve Privacy in Hospital Management. System *Proceedings of the International Conference on Inventive Research in Computing Applications (ICIRCA 2018)*, 1299-1304. DOI: https://doi.org/10.1109/ICIRCA.2018.8597182

Srinivasa L. Chakravarthya, V. Valli Kumari, Ch. Sarojini. (2012). A Coalitional Game Theoretic Mechanism for Privacy Preserving Publishing Based on k-Anonymity. *Procedia Technology*, 6, 889-896. DOI: https://doi.org/10.1016/j.protcy.2012.10.108

Alfons Schuster and Yoko Yamaguchi. (2010). Application of Game Theory to Neuronal Networks. *Advances in Artificial Intelligence*, 2010, Article ID 521606. DOI: https://doi.org/10.1155/2010/521606

Sergey Zapechnikov. (2020). Privacy-Preserving Machine Learning as a Tool for Secure Personalized Information Services. *Procedia Computer Science*, 169, 393–399. DOI: https://doi.org/10.1016/j.procs.2020.02.235

J. Andrew, J. Karthikeyan, Jeffy Jebastin. (2019). Privacy Preserving Big Data Publication on Cloud Using Mondrian Anonymization Techniques and Deep Neural Networks. *Proceedings of The 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, 722-728. DOI: https://doi.org/10.1109/ICACCS.2019.8728384

Rasim M. Alguliyev, Ramiz M. Aliguliyev, Fargana J. Abdullayeva. (2019). Privacy-Preserving Deep Learning Algorithm for Big Personal Data Analysis. *Journal of Industrial Information Integration*, 15, 1–14. DOI: https://doi.org/10.1016/j.jii.2019.07.002

Julian Stier, Gabriele Gianini, Michael Granitzer, Konstantin Ziegler. (2018). Analyzing Neural Network Topologies: A Game Theoretic Approach. *Procedia Computer Science*, 126, 234–243. DOI: https://doi.org/10.1016/j.procs.2018.07.257