Privacy Protection for Context-Aware Services: A Two-Layer Three-Party Game Model

Yan Huang, Zhipeng Cai, and Anu G. Bourgeois

Georgia State University, Atlanta GA 3030, USA
yhuang30@student.gsu.edu, {zcai,abourgeois}@gsu.edu

Abstract. In the era of context-aware services, users are enjoying remarkable services based on data collected from a multitude of users. However, in order to benefit from these services, users are enduring the risk of leaking private information. Game theory is a powerful method that is utilized to balance such tradeoff problems. The drawback is that most schemes consider the tradeoff problem from the aspect of the users, while the platform is the party that dominates the interaction in reality. There is also an oversight to formulate the interaction occurring between multiple users, as well as the mutual influence between any two parties involved, including the user, platform and adversary. In this paper, we propose a platform-centric two-layer three-party game model to protect the users' privacy and provide quality of service. One layer focuses on the interactions among the multiple asymmetric users and the second layer considers the influence between any two of the three parties (user, platform, and adversary). We prove that the Nash Equilibrium exists in the proposed game and find the optimal strategy for the platform to provide quality service, while protecting private data, along with interactions with the adversary. Using real datasets, we present simulations to validate our theoretical analysis.

1 Introduction

Due to the rapid development and popularity of context aware services, people's lives have become more comfortable and convenient than ever before. Applications include health care, smart grid, industrial services [48], social network platforms, and transportation, to name a few [5, 7]. Smart transportation [6,15,32] provides drivers the optimal path based upon current traffic conditions, e-health [46] platforms are able to continuously monitor a patient's health status and facilitate communication with the healthcare specialist, and the smart grid [31] improves power management by monitoring usage patterns and balancing loads. Typically, it is only with the users' information that these context aware applications can provide and maintain any service and the quality of the service is often directly dependent upon the quantity and quality of collected data. As a result, users must consider the cost of leaking private information in order to benefit from such services. There is, of course, always a threat of private information being captured during data transmission. However, in recent years, private data leakage, as well as intentional data sale/reuse is more

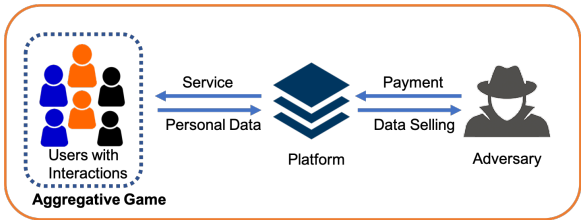

Fig. 1. Two-layer three-party game model

likely from the service provider platform [3,4,16,18,25,26,49]. In recent news we learned of Facebook improperly sharing data that impacted 87 million users [2] and Equifax [1] compromised private information of 143 million users. In spite of this, users still employ these applications, as the services are deemed essential to many people, thus positions the provider platforms in a dominant capacity [14].

This has led to considerable research on techniques to protect a user's private data from being leaked and/or sold. Most of the privacy protection algorithms, e.g. k-anonymity [40, 41], l-diversity [30, 33], t-closeness [22], and differential privacy [36, 50], protect the data by adding noise, but this in turn will decrease the quality of the services provided. Therefore, several game theory based models have been proposed to balance the trade-off between service quality or reward and privacy protection.

Most of the game theory based research has a drawback, in that they only focus on the interaction between two parties, i.e. the user with an untrusted platform, or the user with an adversary (an entity trying to purchase or steal data) [13, 27, 28, 44]. A more realistic model should consider the interactions of the three parties: the user, platform, and adversary. These two-party models ignore or fail to formulate the interaction between each pair of parties (3 such pairings). More recently, diverse three-party game models have been proposed to provide a more realistic interaction analysis [21,23,24,39,42,43]. Yet there is still a shortcoming, as they can only provide binary strategies, meaning the decision for users to submit or not submit their data. Instead, it would be beneficial to have a fine-grained strategy to provide a protection level ranging between 0-1, which is what we propose in this paper.

Another deficiency with the current n-player game models (those with n users) [13, 28, 29, 44, 45, 47] is that they only consider the interaction between the users and other parties (either the platform or adversary). They fail to represent the interaction between asymmetric users, where users have individual privacy protection expectations. To demonstrate the impact, let us consider a transportation application. A user is able to get accurate traffic status without submitting any personal information to the platform, provided other users do submit their information. If multiple users stop submitting their information, the service quality will decrease, and if no users submit their information, minimal service can be provided. Thus multiple users must submit their data to provide enough context to the platform for better quality service.

In this paper, we design a platform centric two-layer three-party game model to provide a balanced fine-grained strategy for the platform, while minimizing users' privacy loss and maximizing quality of service (shown in Fig. 1). To avoid the drawbacks of the existing work, we need to overcome the following challenges: (i) Interaction among asymmetric users. Users of the same service have interactions and each user has a different privacy protection expectation. Thus, the interactions among the users increase the difficulty in addressing the users' strategy selection. (ii) Complicated game structure. Users' strategies are not only influenced by other users, but also by the platform's strategy, as well as the adversary's strategy. We formulate this by using two-layer game model a game model among asymmetric users and a game model among users, platform and adversary. (iii) Theoretical analysis and solution. The complicated game structure and asymmetric users increases the difficulty to perform a theoretical analysis of Nash equilibrium and determining proper strategies for users and platforms.

The following methods are implemented to address the above challenges in this paper. Firstly, we utilize a quasi-aggregative game model to formulate the interactions between asymmetric users and utilize a contract model to formulate the interactions between the platform and adversary. Secondly, based on the proposed two-layer three-party game model, we analyze the Nash equilibrium to find the proper fine-grained strategies for all users and the platform. Finally, we perform simulations based on real datasets to validate the theoretical analysis.

To the best of our knowledge, we are the first to provide a privacy protection framework from the perspective of the platform, since the platform is in the dominate position, as described above. The main contributions of this paper are summarized as follows:

- A platform-centric two-layer three-party game model to capture the interactions among asymmetric users, and the interactions between users, the platform, and adversary. This will provide proper guidance for both the users and platforms.
- The theoretical Nash equilibrium analysis to find the proper fine-grained guidance for all the asymmetric users and the platform.
- Simulations with real datasets to validate the theoretical analysis and evaluate the performance of the proposed two-layer three-party game.

The rest of the paper is organized to introduce the system model in Section 3. We analyze the optimal strategies for asymmetric users and platforms in Section 4. Section 5 presents the simulations to validate the theoretical analysis and we conclude the paper and discuss future work in Section 6.

2 Preliminary

In this section, we present previous results that are fundamental to the work proposed in this paper. Let $\Gamma = (\tilde{\pi}_i, S_i)_{i \in \mathscr{I}}$ denotes a non-cooperative, pure strategy game with a finite set of players $\mathscr{I} = \{1, ..., I\}$, and finite dimensional

strategy sets $S_i \subset R^N, s_i \in S_i$. The joint strategy set $S = \prod_{i \in \mathscr{I}} S_i$, is assumed to be a compact metric space, and payoff functions $\tilde{\pi}_i : S \to R, i \in \mathscr{I}$, are assumed to be upper semi-continuous. Then the Quasi-Aggregative Game can be defined as follows.

Definition 1. *(Quasi-Aggregative Game) [17] The game $\Gamma = (\tilde{\pi}_i, S_i)_{i \in \mathscr{I}}$ is a quasi-aggregative game with aggregator $g : S \to \mathbb{R}$, if there exist continuous functions $F_i : \mathbb{R} \times S_i \to \mathbb{R}$ (the shift functions), and $\sigma_i : S_{-i} \to X_{-i} \subset \mathbb{R}, i \in \mathscr{I}$ (the interaction functions) such that each of the payoff functions $i \in \mathscr{I}$ can be written: $\tilde{\pi}_i = \pi_i \left(\sigma_i \left(s_{-i}, s_i\right), s_i\right)$, where $\pi_i : X_{-i} \times S_i \to \mathbb{R}$, and: $g(s) = F_i \left(\sigma_i(s_{-i}), s_i\right)), \forall s \in S, i \in \mathscr{I}$. Agent i's best-replies, depend on $x_{-i} = \sigma_i(s_{-i})$, is given by $R_i(x_{-i}) = arg \max \pi_i(x_{-i}, s_i) : s_i \in S_i$.*

Theorem 1. *The quasi-aggregative game has a pure strategy Nash equilibrium (PSNE) the following two assumptions holds. [17]*

Assumption 1 *Each correspondence $R_i : X_{-i} \to 2^{S_i}$ is strictly decreasing.*

Assumption 2 *The shift-function F_i, $i \in \mathscr{I}$, all exhibit strictly increasing differences in x_{-i} and s_i.*

3 System Model

In this section, we formulate the interactions between asymmetric users, as well as the interactions among the three parties and introduce the proposed game model.

3.1 Users Model

Assume a set of users $N = \{1, 2, ..., n\}$ use a client of a platform to get context-based service. Each user $i \in N$ will submit a dataset $D_i = \{d_{i1}, d_{i2}, ..., d_{im}\}$ with m attributes to the platform. The client has a local privacy protection algorithm installed which satisfies strict privacy protection standards, such as Local Differential Privacy [36]. Thus, the platform can only get anonymized data or noise-added data from users.

Even if the client has a privacy protection algorithm installed, the anonymized data or noise-added data can still leak some information to the platform, the privacy leakage level depends on the privacy protection setting of the client. Without loss of generality, we define the privacy protection level of attribute j as $\delta_j \in [0, 1]$.

When $\delta_j = 1$, the platform cannot retrieve any information about users' attribute j. When $\delta_j = 0$, the platform can retrieve all the information about users' attribute j. To get statistical result from users, the platform has to set the same $\boldsymbol{\delta} = \{\delta_1, \delta_2, ..., \delta_m\}$ for all the users [9,11,20,38]. According to privacy protection laws, such as General Data Protection Regulation within the European Union and the European Economic Area, the platform should use strongest privacy

protection strength in the client by default. Thus, the default setting of privacy protection level vector is $\boldsymbol{\delta} = \{1, 1, ..., 1\}$.

However, by using the strongest privacy protection strength, the platform cannot collect usable information from users, resulting in worst service quality. Thus, to collect information from users, the platform has to offer a $\boldsymbol{\delta}$ with lower privacy protection level.

Users have the right to accept or reject the platform's offer $\boldsymbol{\delta}$. We define user i's strategy for attribute j as $a_{ij} \in [0, 1]$, which defines the probability of user i accept the privacy leakage level δ_j. Therefore, the strategy vector of user i is $\boldsymbol{a}_i = \{a_{i1}, a_{i2}, ..., a_{im}\}$ and the strategy vector of all users is $\boldsymbol{a} = \{\boldsymbol{a_i}, \boldsymbol{a_j}, ..., \boldsymbol{a_n}\}$.

The service quality depends on the users' strategy, and one user's strategy has a marginal impact on service quality. The service quality of user i received from the platform depends not only on its strategy $\boldsymbol{a}_i$, but also on the strategy of other users $\boldsymbol{a}_{-i}$. Formally, for a specific privacy protection level, the expected received service quality of user i is determined by the strategy of user i and other users' strategy, which can be defined as $Q_i(\boldsymbol{a}_{-i}, \boldsymbol{a}_i)$.

Meanwhile, the platform may resell users' data to a adversary resulting in privacy loss to the users. Assume each user has a constant privacy cost estimation vector $\boldsymbol{c_i} = \{c_{i1}, c_{i2}, ..., c_{im}\}$, where c_{ij} defines the privacy cost of attribute j's privacy leakage. We can define the total cost estimation of user i as follows:

$$C_i^u(\boldsymbol{a}_i) = \sum_{j=1}^{m} c_{ij} a_{ij} (s_j + (1 - \delta_j)), \tag{1}$$

where $s_j \leq \delta_j$ is privacy leakage level when the platform resells the users' dataset.

Thus, we can derive the expected utility function of user i as follows.

$$U_i^u(\boldsymbol{a}_i, \boldsymbol{a}_{-i}) = Q_i(\boldsymbol{a}_{-i}, \boldsymbol{a}_i) - C_i^u(\boldsymbol{a}_i). \tag{2}$$

3.2 Platform Model

The quality of service depends upon the number of users that accept the privacy protection level of attributes. For this reason, the platform entices uses to accept the offer with higher privacy leakage level by providing more accurate service quality. We define $\sigma_j(\boldsymbol{a})$ as the expected number of users that accept the information leakage level δ_j for attribute j, and calculate $\sigma_j(\boldsymbol{a})$ as

$$\sigma_j(\boldsymbol{a}) = \sum_{i=1}^{n} a_{ij}. \tag{3}$$

The value of δ_j reveals the privacy leakage of users' attribute j and also reveals the information that can be retrieved by the platform. According to the research of privacy protection algorithms [10,11,37], the service quality based on attribute j can be defined as a logarithmic function of privacy leakage level δ_j, and is affected by the number of users that accept the privacy leakage level δ_j as a law of diminishing marginal utility. Therefore, we can derive that the service

quality depends on a single attribute j as $log((1-\delta_j)+1)\sigma_j^b(\boldsymbol{a})$, where $0 < b < 1$ is the parameter revealing the impact of $\sigma_j^b(\boldsymbol{a})$. The value of b is decided by the local privacy protection algorithm.

Meanwhile, attribute i and attribute j may have a correlation. Thus, the information of attribute i not only contributes to the service which is based on attribute i but also contributes to the service which is based on attribute j, if there is a correlation between attribute i and j. We define the correlation between attribute i and attribute j as e_{ij}. Therefore, the information of attribute i also contributes to the expected service quality which is based on attribute j with the correlation coefficient e_{ij}. Accordingly, we can define the total expected service quality Q as

$$Q\left(\boldsymbol{\delta}, \boldsymbol{a}\right) = \sum_{j=1}^{m} \left(1 + \sum_{k=1, k\neq j}^{m} e_{jk}\right) log((1-\delta_j)+1)\sigma_j^b(\boldsymbol{a}). \tag{4}$$

The collected dataset from users can generate income for the platform. The expected income form data is also affected by the privacy leakage level $\boldsymbol{\delta}$ and the number of users who accept the platform's offer. According to data aggregation research [19] and the standard form of Cobb-Douglas production function [34], the expected data value to the platform can be defined as

$$V^p\left(\boldsymbol{\delta}, \boldsymbol{a}\right) = \alpha \sum_{j=1}^{m} (1-\delta)_j^{\zeta} \sigma_j^b(\boldsymbol{a}), \tag{5}$$

where α is the total value productivity of the platform, and $\zeta \in (0, 1)$ is the platform's value output elasticities of each attribute.

To get extra profit, the platform could sell the collected data to an adversary. The platform may choose a different privacy leakage level vector $\boldsymbol{s} = \{s_1, s_2, ..., s_m\}$ for the resale dataset. And for each unit of privacy leakage level of attribute j, the platform asks for a price p_j for each user's data. The price vector of the dataset is defined as $\boldsymbol{p} = \{p_1, p_2, ...p_m\}$, which is determined in a contract with the adversary. Thus, the total expected price is defined as

$$P\left(\boldsymbol{s}, \boldsymbol{p}, \boldsymbol{a}\right) = \sum_{j=1}^{m} p_j s_j \sigma_j^b(\boldsymbol{a}). \tag{6}$$

However, the data resale incurs a cost due to reputation loss to the platform. If we define r_j is the unit cost for reselling one user's attribute j with privacy leakage level s_j, we can derive the expected cost due to reputation loss as

$$\sum_{j=1}^{m} r_j s_j \sigma_j^b(\boldsymbol{a}). \tag{7}$$

Meanwhile, the platform has a constant running cost c_p. Thus, the total expected cost of the platform is $C^p\left(\boldsymbol{s}, \boldsymbol{a}\right) = \sum_{j=1}^{m} r_j s_j \sigma_j^b(\boldsymbol{a}) + c_p$.

To sum up, the expected utility of the platform is $U^p\left(\boldsymbol{\delta}, \boldsymbol{s}, \boldsymbol{p}, \boldsymbol{a}\right) = V^p\left(\boldsymbol{\delta}, \boldsymbol{a}\right) + P\left(\boldsymbol{s}, \boldsymbol{p}, \boldsymbol{a}\right) - C^p\left(\boldsymbol{s}, \boldsymbol{a}\right)$. The platform will maximize its utility by achieving a Nash Equilibrium with the users and adversary.

3.3 Adversary Model

To get users information, the third party can purchase data from the platform. By using purchased data, the adversary can generate value according to its type γ, where θ is its value productivity, and γ is its value output elasticities of each attribute. According to data aggregation research [19] and the standard form of Cobb-Douglas production function [34], the expected data value to the adversary can be defined as

$$V_t\left(\boldsymbol{s}, \boldsymbol{a}\right) = \theta \sum_{j=1}^{m} s_j^{\gamma} \sigma_j^b(\boldsymbol{a}). \tag{8}$$

Thus, the expected utility function of the third party is

$$U^t\left(\left(\boldsymbol{p}(\gamma), \boldsymbol{s}\right), \boldsymbol{a}\right) = V_t\left(\boldsymbol{s}, \boldsymbol{a}\right) - P\left(\left(\boldsymbol{p}(\gamma), \boldsymbol{s}\right), \boldsymbol{a}\right). \tag{9}$$

4 Game Model

In this section, we formulate the problem with a two-layer three-party game and analyze its Nash Equilibrium.

4.1 Aggregative Game Model

In this paper, we assume users do not exchange information with the other users. Each user's action influences the other users' utility. With a specific privacy leakage level $\boldsymbol{\delta}$, we can use quasi-aggregative game model to formulate the interactions among users.

To maximize utility, a user chooses a proper privacy leakage level for each attribute. According to [17], we define the interactions among users as m quasi-aggregative games, e.g., $\Gamma_j = (\tilde{\pi}_{ij}, A_i), \forall j = 1, 2, ...m$, where A_i is user i's strategy space. The payoff function of each player in this game can be defined as $\tilde{\pi}_{ij} = U_{ij}^u(\sigma_{ij}(\boldsymbol{a}_{-i}), a_{ij})$; the aggregator can be defined as $g_j(\boldsymbol{a}) = F_{ij}(\sigma_{ij}(\boldsymbol{a}_{-i}), a_{ij}) = \sigma_{ij}(\boldsymbol{a}_{-i}) + a_{ij}$; the interaction functions vector can be defined as $\sigma_{ij}(\boldsymbol{a}_{-i}) = \sum_{k \in N, k \neq i} a_{kj}$.

User i in the game Γ_j aims to maximize its utility by properly choosing a strategy vector $\boldsymbol{a_i}$ such that $\boldsymbol{a_i} = \underset{a_{ij}}{arg\max}\, U_i^u(\boldsymbol{\sigma}_i(\boldsymbol{a}_{-i}), a_{ij})$.

According to the property of quasi-aggregative game theory [17], we can derive the following theorem.

Theorem 2. *The game $\Gamma_u = (\tilde{\pi}_i, A_i)_{i \in N}$ has a pure strategy Nash equilibrium (PSNE) for any privacy leakage level $\boldsymbol{\delta}$.*

Proof. When the integrated value σ_{-i} increases, user i can get increased payoff. Thus, user i can increase its payoff by decreasing the value of strategy s_i. As a result, the best-reply correspondence of user i is strictly decreasing. It is obviously that the shift function F_i (Eq. 4.1) exhibits strictly increasing differences in x_{-i} and s_i. According to [17], the theorem is proved.

4.2 Contract Model

The platform makes a contract with the adversary. Assume the adversary announces its type is γ, $\gamma \in (0,1)$. The platform provides a menu of contracts $\{(\boldsymbol{p}(\gamma), \boldsymbol{s})\}$to the adversary. According to contract theory [35], to incentivize the adversary to accept the contract designated for him rather than choosing other contracts or refusing any contract, the menu of contracts should satisfy both the individual rationality condition and the incentive compatibility condition defined below.

Condition 1 *(Individual Rationality (IR)) A menu of contracts* $\{(\boldsymbol{p}(\gamma), \boldsymbol{s})\}$ *satisfies the individual rationality constraints if it yields to the adversary a non-negative payoff, i.e.,* $\forall \gamma \in (0,1)$, $U^t(\boldsymbol{p}(\gamma), \boldsymbol{s}) \geq 0$, *where* $U^t(\boldsymbol{p}(\gamma), \boldsymbol{s})$ *is the utility of adversary with type* γ.

Condition 2 *(Incentive Compatibility (IC)) A menu of contracts* $\{(\boldsymbol{p}(\gamma), \boldsymbol{\delta})\}$ *satisfies the individual compatibility constraints if the best response for the adversary with type* γ *is to choose the contract* $(\boldsymbol{p}(\gamma), \boldsymbol{s})$ *rather than other contracts, i.e.,* $\forall \gamma, \hat{\gamma} \in (0,1)$, $U^t(\boldsymbol{p}(\gamma), \boldsymbol{\delta}) \geq U^t(\boldsymbol{p}(\hat{\gamma}), \boldsymbol{s})$.

Therefore, the objective of the platform is to maximize its utility by properly creating a menu of contracts. We formalize the optimization problem of the platform as follows.

$$\begin{aligned} &\max_{\{(\boldsymbol{p}(\gamma), \boldsymbol{s})\}} U^p\left(\boldsymbol{\delta}, \boldsymbol{s}, \boldsymbol{p}(\gamma), \boldsymbol{a}\right), \\ &\textit{subject to Condition } 1 \textit{ and } 2. \end{aligned} \tag{10}$$

According to the aggregative model and contract model, we can see that the platform needs to properly choose the privacy leakage level $\boldsymbol{\delta}$ for all users and create the contract menu for the adversary to maximize its utility. Therefore, the Nash Equilibrium can be derived by solving the combined optimization problem:

$$\begin{aligned} &\max_{(\boldsymbol{\delta}, \{(\boldsymbol{p}(\gamma), \boldsymbol{s})\})} U^p\left(\boldsymbol{\delta}, \boldsymbol{s}, \boldsymbol{p}(\gamma), \boldsymbol{a}^*\right), \\ &\textit{subject to Condition } 1 \textit{ and } 2. \end{aligned} \tag{11}$$

where $\boldsymbol{a}^*$ is the PSNE of the aggregative game.

5 Simulation

In this section, we study the interactions in the proposed two-layer three-party game. In the simulation, we utilize a parallel machining learning algorithm termed Particle Swarm Optimization (PSO) [8] to find the optimal strategies for the user and the platform.

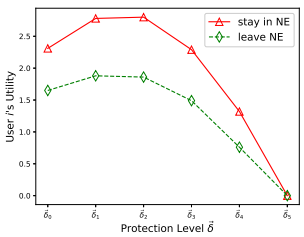

Fig. 2. User utility vs. protection level.

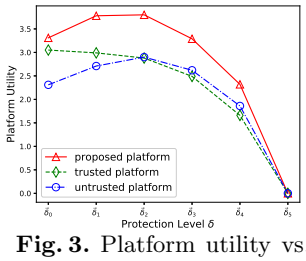

Fig. 3. Platform utility vs. protection level.

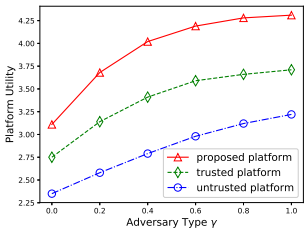

Fig. 4. Optimal strategy of user under various.

5.1 Simulation Setting

We use real datasets as the inputs of the user and platform. More specifically, based on the Data Protection Survey published by SANA [12], we extract four protection levels for income, age, and race. As shown in Table 1, $\boldsymbol{\delta}_1$, $\boldsymbol{\delta}_2$, $\boldsymbol{\delta}_3$, and $\boldsymbol{\delta}_4$, are the protection levels used by Retail platforms, Healthcare platforms, Government platforms, and Financial platforms, respectively.

Table 1. Extracted strategies

Application	{Income, Age, Race}
Retail	$\boldsymbol{\delta_1}$={0.2, 0.3, 0.4}
Healthcare	$\boldsymbol{\delta_2}$={0.3, 0.4, 0.5}
Government	$\boldsymbol{\delta_3}$={0.4, 0.5, 0.7}
Financial	$\boldsymbol{\delta_4}$={0.6, 0.7, 0.8}

We set the correlation coefficient between income and age as 0.1, the correlation coefficient between income and race as 0.01, and the correlation coefficient between age and race as 0. We also tried the other correlation coefficient values and find out that the correlation coefficient is not a key factor. The privacy costs of users have normal distribution with parameters: $\mu_{income} = 10, \mu_{age} = 6, \mu_{race} = 2$ and $\sigma^2 = 1$. The total value productivity of the platform is $\alpha = 6$ and the output elasticity is $\zeta = 0.6$. The total value productivity of the adversary is $\theta = 8$ and the output elasticity is $b = 0.6$. The reputation cost for the attributes are $r_{income} = 3$, $r_{age} = 2$,$r_{race} = 1$. We choose the best strategy from running the algorithm 100 times, where each run consists of 10,000 iterations.

5.2 Users Interaction

Fig. 2 shows the utility of user i when it performs different actions under different privacy protection levels. The x axis is the protection level, where $\boldsymbol{\delta}_0 = \{0, 0, 0\}$ is the lowest protection level and $\boldsymbol{\delta}_5 = \{1, 1, 1\}$ is the highest protection level. $\boldsymbol{\delta}_1$ to $\boldsymbol{\delta}_4$ are the increasing protection levels, as in Table 1. The solid red line in Fig. 2

shows the utility of user i when it stays in the Nash Equilibrium, and the dashed green line is when it leaves the Nash Equilibrium. As we can see, the user's utility increases at first and then decreases as the protection level increases. The reason for utility increasing, is that the rate of the user's privacy loss decreasing is larger than that of service quality decreasing. However, the user's utility decreases after the maximum point, because the rate of service quality decreasing is larger than that of privacy loss decreasing. User i has utility 0 with the strongest protection level $\boldsymbol{\delta}_5$ because the user cannot get any service quality and has no privacy loss. Fig. 2 also shows us that the utility of user i when it stays in NE is higher than that when it leaves NE. This proves the existence of NE in the aggregative model and that users cannot get higher utility if they use non-NE strategies.

5.3 Platform Comparison

We compare the proposed platform with a trusted platform and an untrusted platform. We assume the trusted platform keeps users' data safe and will not trade the data, while the untrusted platform sells all its collected data.

As shown in Fig. 3, the utility of the proposed platform (solid red line) increases at first and then decreases as the protection level increases. The utility increases because the rate of payoff increasing is larger than that of reputation loss increasing and the utility decreases because the rate of payoff increasing is less than that of reputation loss increasing. This proves the NE existence of the two-layer three-party game because the platform cannot increase its utility by simply decreasing the privacy protection level.

Fig. 3 and Fig. 4 compare the utility of three types of platforms with different protection levels and different adversary types, respectively. As shown in Fig. 3, the trusted platform has higher utility than the untrusted platform with protection level $\boldsymbol{\delta}_0$ to $\boldsymbol{\delta}_1$ because the trusted platform has no reputation loss and the selling profit of untrusted platform cannot make up its reputation loss. The untrusted platform has higher utility than the trusted platform with protection level $\boldsymbol{\delta}_2$ to $\boldsymbol{\delta}_5$ because the payment from selling data can dominate the reputation loss, thus has more profit than the trusted platform. This explains why the platforms usually sell users data in real life.

However, the platform does not need to sell all the users' data to maximize its utility. From Fig. 3 and Fig. 4, we can see that the proposed platform in this paper has the highest utility because it balances the tradeoff between payoff (from data collection and selling data) and reputation loss. It will choose a proper protection level and selling strategy to maximize its utility. Therefore, we can conclude that the proposed framework can provide balanced strategies for the platform. By using the proposed model, the platform will properly choose the data selling strategy, thus decreasing users' privacy loss.

6 Conclusion

The use of context-aware services are integrated into the majority of people's daily lives. By utilizing these services, one must provide certain private infor-

mation in order to receive better outcomes. Users risk leaking private data, as service platforms are sometimes willing to sell this information to a third party, or adversary to gain more profit, thus resulting in conflicting goals.

This paper studies the interactions among the three parties by proposing a platform-centric two-layer three party game. In the proposed game model, we theoretically formulate the behaviors of each party and the interactions among the three parties by using an aggregate game model and contract model. We run simulations with real datasets to validate the effectiveness of the proposed game model. We show that the proposed model can provide the proper strategy for the platform to balance the payoff and reputation loss, thus increasing privacy protection of the users. This work will enable platforms, such as Facebook, to provide quality service and protection to its users, but also provide a means to profit from a balanced strategy. To further investigate more realistic privacy protection issues, this work will be extended to a model that considers the influence of temporal data. Therefore, the users and platform need to consider the privacy protection for not only the current status, but also previous and future conditions.

Acknowledgments

This work is partly supported by the National Science Foundation (NSF) under grant NOs. 1252292, 1741277, 1704287, and 1829674.

References

1. The equifax data breach. https://www.ftc.gov/equifax-data-breach
2. Facebook security breach exposes accounts of 50 million users. https://www.nytimes.com/2018/09/28/technology/facebook-hack-data-breach.html, sept. 28, 2018
3. Cai, Z., He, Z.: Trading private range counting over big iot data. In: The 39th IEEE International Conference on Distributed Computing Systems (July 2019)
4. Cai, Z., He, Z., Guan, X., Li, Y.: Collective data-sanitization for preventing sensitive information inference attacks in social networks. IEEE Transactions on Dependable and Secure Computing **15**(4), 577–590 (July 2018)
5. Cai, Z., Zheng, X.: A private and efficient mechanism for data uploading in smart cyber-physical systems. IEEE TNSE pp. 1–1 (2018)
6. Cai, Z., Zheng, X., Yu, J.: A differential-private framework for urban traffic flows estimation via taxi companies. IEEE Transactions on Industrial Informatics (2019)
7. Capurso, N., Mei, B., Song, T., Cheng, X., Yu, J.: A survey on key fields of context awareness for mobile devices. JNCA **118**, 44 – 60 (2018)
8. Clerc, M., Kennedy, J.: The particle swarm - explosion, stability, and convergence in a multidimensional complex space. IEEE TEC **6**(1), 58–73 (Feb 2002)
9. Cormode, G., Jha, S., Kulkarni, T., Li, N., Srivastava, D., Wang, T.: Privacy at scale: Local differential privacy in practice. In: SIGMOD. pp. 1655–1658 (2018)
10. Dewri, R.: Local differential perturbations: Location privacy under approximate knowledge attackers. IEEE TMC **12**(12), 2360–2372 (Dec 2013)

11. Erlingsson, U., Pihur, V., Korolova, A.: Rappor: Randomized aggregatable privacy-preserving ordinal response. In: CCS. ACM (2014)
12. Filkins, B.: Sensitive data at risk: The sans 2017 data protection survey (Sep 2017)
13. Freudiger, J., Manshaei, M.H., Hubaux, J.P., Parkes, D.C.: Non-cooperative location privacy. TDSC **10**(2), 84–98 (Mar 2013)
14. He, Z., Cai, Z., Yu, J.: Latent-data privacy preserving with customized data utility for social network data. IEEE Transactions on Vehicular Technology **67**(1), 665–673 (Jan 2018)
15. Hu, Q., Wang, S., Hu, C., Huang, J., Li, W., Cheng, X.: Messages in a concealed bottle: Achieving query content privacy with accurate location-based services. IEEE Transactions on Vehicular Technology **67**(8), 7698–7711 (Aug 2018)
16. Huang, Y., Cai, Z., Bourgeois, A.G.: Search locations safely and accurately: A location privacy protection algorithm with accurate service. Journal of Network and Computer Applications **103**, 146 – 156 (2018)
17. Jensen, M.K.: Aggregative games and best-reply potentials. Economic Theory **43**(1), 45–66 (Apr 2010)
18. Jia, Y., Chen, Y., Dong, X., Saxena, P., Mao, J., Liang, Z.: Man-in-the-browser-cache: Persisting https attacks via browser cache poisoning. Computers & Security **55**, 62 – 80 (2015)
19. Jugel, U., Jerzak, Z., Hackenbroich, G., Markl, V.: M4: A visualization-oriented time series data aggregation. VLDB **7**(10), 797–808 (Jun 2014)
20. Kairouz, P., Oh, S., Viswanath, P.: Extremal mechanisms for local differential privacy. In: Advances in Neural Information Processing Systems 27, pp. 2879–2887. Curran Associates, Inc. (2014)
21. Karimi Adl, R., Askari, M., Barker, K., Safavi-Naini, R.: Privacy consensus in anonymization systems via game theory, pp. 74–89. Springer (2012)
22. Li, N., Li, T., Venkatasubramanian, S.: t-closeness: Privacy beyond k-anonymity and l-diversity. In: ICDE. pp. 106–115 (April 2007)
23. Li, W., Song, T., Li, Y., Ma, L., Yu, J., Cheng, X.: A hierarchical game framework for data privacy preservation in context-aware iot applications. In: 2017 IEEE Symposium on Privacy-Aware Computing (PAC). pp. 176–177 (Aug 2017)
24. Li, W., Hu, C., Song, T., Yu, J., Xing, X., Cai, Z.: Preserving data privacy in context-aware applications through hierarchical game. In: SPAC. Washington DC, USA (Sep 2018)
25. Liang, Y., Cai, Z., Han, Q., Li, Y.: Location privacy leakage through sensory data. Security and Communication Networks (2017)
26. Liang, Y., Cai, Z., Yu, J., Han, Q., Li, Y.: Deep learning based inference of private information using embedded sensors in smart devices. IEEE Network **32**(4), 8–14 (July 2018)
27. Liu, C., Wang, S., Ma, L., Cheng, X., Bie, R., Yu, J.: Mechanism design games for thwarting malicious behavior in crowdsourcing applications. In: IEEE INFOCOM 2017 - IEEE Conference on Computer Communications. pp. 1–9 (May 2017)
28. Liu, X., Liu, K., Guo, L., Li, X., Fang, Y.: A game-theoretic approach for achieving k-anonymity in location based services. In: IEEE INFOCOM (Apr 2013)
29. Ma, R., Xiong, J., Lin, M., Yao, Z., Lin, H., Ye, A.: Privacy protection-oriented mobile crowdsensing analysis based on game theory. In: IEEE TBDI. pp. 990–995 (Aug 2017)
30. Machanavajjhala, A., Venkitasubramaniam, M., Kifer, D., Gehrke, J.: l-diversity: Privacy beyond k-anonymity. In: ICDE. vol. 00, p. 24 (04 2006)

31. Maharjan, S., Zhu, Q., Zhang, Y., Gjessing, S., Basar, T.: Dependable demand response management in the smart grid: A stackelberg game approach. IEEE TSG **4**(1), 120–132 (March 2013)
32. Mahrsi, M.K.E., Cme, E., Oukhellou, L., Verleysen, M.: Clustering smart card data for urban mobility analysis. IEEE TITSystems **18**(3), 712–728 (March 2017)
33. Mao, J., Tian, W., Jiang, J., He, Z., Zhou, Z., Liu, J.: Understanding structure-based social network de-anonymization techniques via empirical analysis. EURASIP JWCN **2018**(1) (Dec 2018)
34. Meeusen, W., van Den Broeck, J.: Efficiency estimation from cobb-douglas production functions with composed error. International Economic Review **18**(2), 435–444 (Jun 1977)
35. Miltiadis, M.: The theory of incentives: The principalagent model. The Economic Journal **113**(488), F394–F395 (2001)
36. Pastore, A., Gastpar, M.: Locally differentially-private distribution estimation. In: IEEE ISIT. pp. 2694–2698 (July 2016)
37. Qin, Z., Yang, Y., Yu, T., Khalil, I., Xiao, X., Ren, K.: Heavy hitter estimation over set-valued data with local differential privacy. In: CCS. ACM (2016)
38. Thakurta, A.G., Vyrros, A.H., Vaishampayan, U.S., Kapoor, G., Freudinger, J., Prakash, V.V., Legendre, A., Duplinsky, S.: Emoji frequency detection and deep link frequency
39. Vakilinia, I., Tosh, D.K., Sengupta, S.: 3-way game model for privacy-preserving cybersecurity information exchange framework. In: MILCOM (Oct 2017)
40. Wang, J., Cai, Z., Li, Y., Yang, D., Li, J., Gao, H.: Protecting query privacy with differentially private k-anonymity in location-based services. Personal and Ubiquitous Computing pp. 1–17 (2018)
41. Wang, S., Hu, Q., Sun, Y., Huang, J.: Privacy preservation in location-based services. IEEE Communications Magazine **56**(3), 134–140 (March 2018)
42. Wang, S., Huang, J., Li, L., Ma, L., Cheng, X.: Quantum game analysis of privacy-leakage for application ecosystems. In: MobiHoc (Jul 2017)
43. Wang, S., Li, L., Sun, W., Guo, J., Bie, R., Lin, K.: Context sensing system analysis for privacy preservation based on game theory. Sensors **17**(2), 339 (Feb 2017)
44. Wu, X., Dou, W., Ni, Q.: Game theory based privacy preserving analysis in correlated data publication. In: ACSW (Feb 2017)
45. Xu, L., Jiang, C., Qian, Y., Li, J., Zhao, Y., Ren, Y.: Privacy-accuracy trade-off in differentially-private distributed classification: A game theoretical approach. IEEE TBD pp. 1–1 (2017)
46. Yi, C., Cai, J.: A priority-aware truthful mechanism for supporting multi-class delay-sensitive medical packet transmissions in e-health networks. IEEE TMC **16**(9), 2422–2435 (Sept 2017)
47. Ying, B., Nayak, A.: Location privacy-protection based on p-destination in mobile social networks: A game theory analysis. In: IEEE CDSC. pp. 243–250 (Aug 2017)
48. Zheng, X., Cai, Z., Li, J., Gao, H.: Location-privacy-aware review publication mechanism for local business service systems. In: IEEE INFOCOM. pp. 1–9 (May 2017)
49. Zheng, X., Cai, Z., Li, Y.: Data linkage in smart internet of things systems: A consideration from a privacy perspective. IEEE Communications Magazine **56**(9), 55–61 (Sep 2018)
50. Zheng, X., Cai, Z., Yu, J., Wang, C., Li, Y.: Follow but no track: Privacy preserved profile publishing in cyber-physical social systems. IEEE Internet of Things Journal **4**(6), 1868–1878 (Dec 2017)