Georgia State University
ScholarWorks @ Georgia State University

Computer Science Dissertations

Department of Computer Science

8-13-2019

Game Theory Based Privacy Protection for Context-Aware Services

Yan Huang

GAME THEORY BASED PRIVACY PROTECTION
FOR CONTEXT-AWARE SERVICES

by

YAN HUANG

Under the Direction of Anu G. Bourgeois, Ph.D. and Zhipeng Cai, Ph.D.

ABSTRACT

In the era of context-aware services, users are enjoying remarkable services based on data collected from a multitude of users. To receive services, they are at risk of leaking private information from adversaries possibly eavesdropping on the data and/or the un–trusted service platform selling off its data. Malicious adversaries may use leaked information to violate users' privacy in unpredictable ways. To protect users' privacy, many algorithms are proposed to protect users' sensitive information by adding noise, thus causing context-aware service quality loss. Game theory has been utilized as a powerful tool to balance the tradeoff between privacy protection level and service quality. However, most of the existing schemes fail to depict the mutual relationship between any two parties involved: user, platform, and adversary. There is also an oversight to formulate the interaction occurring between multiple users, as well as the interaction between any two attributes. To solve these issues, this dissertation firstly proposes a three-party game framework to formulate the mutual interaction between three parties and study the optimal privacy protection level for context-aware services, thus optimize the service quality. Next, this dissertation extends the framework to a multi-user scenario and proposes a two-layer three-party game framework. This makes the proposed framework more realistic by further exploring the interaction, not only between different parties, but also between users. Finally, we focus on analyzing the impact of long-term time-serial data and the active actions of the platform and adversary. To achieve this objective, we design a three-party Stackelberg game model to help the user to decide whether to update information and the granularity of updated information.

INDEX WORDS: Privacy Protection, Game Theory, Nash Equilibrium, Three-Party Game, Context-Aware Services

GAME THEORY BASED PRIVACY PROTECTION FOR CONTEXT-AWARE SERVICES

by

YAN HUANG

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2019

GAME THEORY BASED PRIVACY PROTECTION
FOR CONTEXT-AWARE SERVICES

by

YAN HUANG

Committee Chair:	Anu G. Bourgeois
Committee:	Zhipeng Cai Wei Li Ruiyan Luo

Electronic Version Approved:

Office of Graduate Studies
College of Arts and Sciences
Georgia State University
August 2019

DEDICATION

This dissertation is dedicated to my parents Guangyong Huang and Qiulian Yang, my fiancee Nanxi Peng for their endless support and love during my Ph.D. years. I cannot finish my Ph.D. without their love and encouragement.

ACKNOWLEDGMENT

It is a truly life-changing experience for me to pursuing Ph.D. degree in Georgia State University during the past four years. I would never have been able to finish my dissertation without the guidance of my advisors and committee members, help from my group, and support from my family and my friends.

I would like to show my deepest gratitude to my advisor Dr. Anu G. Bourgeois, and Dr. Zhipeng Cai. They provided me with an excellent environment for research, and gave me many opportunities to promote myself. They not only taught and encouraged me in my research but also inspired me to achieve self-actualization.

It is very grateful and a great honor to have Dr. Wei Li, and Dr. Ruiyan Luo in my committee, who gave me great supports for my Ph.D. study and spared time to participate in my defense committee.

Also many thanks go to colleagues in my group and department. Special thanks for my group colleagues Dr. Meng Han, Dr. Yi Liang, Dr. Dongjing Miao, Dr. Xu Zheng, Dr. Zhuojun Duan, Dr. Zaobo He, Dr. Ji Li, who helped me study and live in U.S..

Last but not least, it is a pleasure to thank everybody who made the dissertation possible, as well as express my apologies that I could not mention personally one by one.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

- DVD - Digital Versatile Disc
- LBSs - Lcation-Based Services
- LPPMs - Location-Privacy Preserving Mechanisms
- NE - Nash Equilibrium
- PSO - Particle Swarm Optimization
- PSNE - Pure Strategy Nash Equilibrium
- IR - Individual Rationality
- IC - Incentive Compatibility
- IoT - Internet of Things
- AI - Artificial Intelligence
- DL - Deep Learning

Chapter 1

INTRODUCTION

People are enjoying a huge convenience from context-aware services, such as navigation services provided by Google Map, reviews and recommendations services provided by Yelp, ride-sharing services by Uber and Lyft, online social media services provided by Facebook, etc. However, people are suffering from privacy leakage while enjoying these convenient services. According to the statistics from [1], 55% of iOS applications and 59.7% of Android applications surreptitiously leak user's data. In recent news, we learned of Facebook improperly sharing data that impacted 87 million users [2] and Equifax [3] compromised private information of 143 million users.

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 [4], l-diversity [5], t-closeness [6], and differential privacy [7], protect the data by adding noise. However, the added noise in existing work will decrease the quality of provided services [8]. To maximize the service quality while satisfying the privacy protection demands, researchers made the efforts on two aspects: (i) optimizing the privacy protection algorithm to maintain as much usable information under the privacy protection restriction, and (ii) utilizing Game Theory to find the optimized privacy protection setting for the algorithms by balancing privacy loss and service quality. However, each aspect has its drawbacks.

Current k-anonymity, l-diversity, t-closeness based algorithms suffer from differential attack [9]. In October 2006, Netflix, the worlds largest online Digital Versatile Disc (DVD) rental service, publicly released a data set containing 100 million anonymized movie ratings, created by 500,000 users of Netflix. Narayanan and Shmatikov demonstrated that an adversary could identify the users information with less background knowledge about individual users. They revealed the users sensitive information by using IMDb as the source of the background knowledge [10]. Differential Privacy based algorithms can prevent users from differential attacks [11]. The limitation of Differential Privacy based algorithm is that it cannot be used in the scenario with only one user. To solve this problem, [12] mixed differential privacy and k-anonymity proposed a perturbation method based on local enforcement of differential privacy. However, both k-Anonymity and differential privacy will lead to unavoidable inaccurate service.

To maximize the service quality in the privacy protected context-aware services, many researchers utilize Game Theory to find the best privacy protection level setting by balancing the tradeoff between privacy loss and service quality. Most of the existing game theory based work investigates the interaction between two parties: user/data owner and adversary. In [13–18], games are based on a two-player model, i.e., one-against-one. When there are multiple users trying to maintain a certain privacy preserving level, the user-adversary game can be modeled as an n-player game [19–24], but with the drawback that all users must have the same settings. Another drawback is that the two-party game cannot depict the interactions among three antithetic parties. Recently, three-party game models have been proposed to study complicated privacy issues among user/data owner, service provider/data requester, and adversary. In [25], Li *et. al.* designed a hierarchical game, incorporating a user-service provider game and a user-attacker game, to maximize the service provider's utility while assisting the user in defending the attacker. Adl *et. al.* [26] proposed a three-party sequential game to analyze the interactions among a data provider, a data collector, and a data user (i.e., the adversary), which can guide the data provider and the data collector to find the optimal strategies deciding whether to cooperate with the data user. In [27,28], Wang *et. al.* studied the interactions among a user, an application, and an adversary to answer two questions: whether the user should submit data and whether the application should resell the user's data? To resolve the trade-off between sharing advantages and privacy exposure of cybersecurity information exchange system, Vakilinia *et. al.* [29] designed a three-party game for privacy-preserving cybersecurity information exchange framework consisting of an attacker, an organization, and a cybersecurity information exchange system. However, the three-party games in [25–29] fail to build the mutual interaction between any two of the three parties, and the strategy of each party in [26–28] is coarse-grained, or binary, by indicating "whether to cooperate with opponents or not".

These drawbacks and challenges have not been properly solved. Contrasting from the existing work, we establish our research by proposing a three-party game model to capture the mutual interaction between any two of the three antithetic parties (including user/data owner, service provider/data requester, and adversary) and aim to identify their strategies on *"how to defend (or cooperate with) others"*, which can offer *fine-grained guidance to the three parties*. The main contribution of each part includes:

In the first part, we design a privacy-preserving game to quantify the three parties' concerns and capture interactions between any two of them. We also identify the best strategy for each party at a fine-grained level, i.e. specific settings, not simply binary. Via both theoretical analysis and real-data experiments, the performance of our proposed game model is validated.

In the second part, 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.

The third part analyzes the influence of time-serial data and the possible feedback from a platform with diverse reward for data according to the feedback in three-party game based framework. This work enhances the three-party game based framework by making it more realistic, thus providing more practical application scenario.

The rest of this dissertation proposal is organized as follows. Chapter 2 summarizes the related literature. Chapter 3 studies the problem of "Privacy Protection among Three Antithetic-Party for Context-Aware Services". This work has been submitted to the 39th

IEEE International Conference on Distributed Computing Systems (ICDCS 2019). Chapter 4 “Privacy Protection for Context-Aware Services: A Two-Layer Three-Party Game Model” expands the work in Chapter 3. This work has been accepted by the 14th International conference on algorithms, systems, and applications of wireless networks (WASA 2019). In chapter 5, we introduce game theory based privacy protection for context-aware services with the long-term time series data. Finally, in Chapter 6, we provide future direction and Chapter 7 concludes our work.

Chapter 2

BACKGROUND

Context-aware services has been protected by privacy protection techniques [30–38]. Game Theory is a popular and efficient tool to find out the optimal privacy protection level for the purpose of improve service quality under the protection algorithms. In this section, we survey existing privacy protection level selection based on game theory .

2.1 Two Parties Game model

Two-party game models have two categories: single user centric game model and user group centric game model.

Single user centric game models provide optimal local privacy protection strategy locally by the user. Chorppath and Alpcan [13] propose a game theoretic approach to formulate the interaction between users and companies for mobile commerce. In their game model, users report their locations with granularity to protect privacy. The proposed game is utilized to find the optimal anonymity level for mobile users and the optimal incentive strategy for companies. To balance the service data quality and location privacy preservation in location-based services (LBSs), Shokri [14] et. al. propose a methodology to utilize Stackelberg Bayesian game to enable a designer to find the optimal location-privacy preserving mechanism (LPPM). Given the service quality constraint of the user, the optimal LPPM can provide the best privacy protection. Shokri et. al. improve their work by taking location correlation into account to protect the trajectory privacy of users [16]. The authors use zero-sum Stackelberg Bayesian game to find the optimal LPPM to against adversary subject to a service quality constraint. Sfar et. al. [17] propose a privacy preserving model for retail applications. They utilize a Markov game model to reach a compromise between privacy concessions of users and incentive motivations of data requester. To preserve privacy on mobile phone, Wang and Zhang [39] construct a zero-sum stochastic game to formulate the strategic and dynamic competition between a smart phone user and a malicious adversary. According to the Nash Equilibrium (NE) of the game, the user can find the optimal defense strategy. The users in single user centric game can achieve the privacy protection by themselves. They do not need to consider the affection from other users and do not need to corporate with other users. However, single user centric game models cannot provide theoretical privacy protection guarantees like k-anonymization and differential privacy.

2.2 User Group centric Game Model

A group user based game model needs users in the group working together to maintain a certain anonymization level.

Halkidi and Koutsipoulos [40] propose a game theoretic framework data privacy preservation in recommender systems. In the recommender system, the quality of recommendations depends on the submitted data from all users. Each user prefers to submit less data to preserve data privacy. However, the quality of recommendation will decay if all the users choose to reveal less information to the platform. In [40], the authors propose a game based framework to balance the trade-off between privacy preservation and quality of recommendation. The Nash Equilibrium Point of the proposed game provides the optimal strategy for each user. Wu et. al. [19] utilize game theory to balance the trade-off between privacy and utility for correlated data publication. They find out the payoff of each user is dependent on both its parameters and its neighbors' privacy parameters. Therefore, they build a game model of multiple players to analyze the optimal privacy parameters of data publication of each user. When a user employs k-Anonymity in LBS, it needs to generate fake location records, known as dummy users, to protection its location privacy. However, due to the high cost of dummy user generation, self-interested users do not want generate dummy users but free-ride on other's efforts. Liu et. al [20] propose a distributed approach to guide users to generate dummies according to their privacy demands and utilize Bayesian game model to balance the cost of dummies generation and privacy protection and find the optimal dummy user generation strategy. To protect the pseudonyms used for authentication in mobile networks, users in a mix zone should collectively change their pseudonyms. A Sef-interested user may not cooperate due to the high cost of pseudonym change. Freudiger et. al. [21] define a game theoretical model to help each user in a mix zone to find the optimal time to change their pseudonym by balancing the privacy protection and cost of pseudonym change. Kumari and Chakravarthy [41] propose a Cooperative Game to achieve privacy preserving before data publishing. The proposed game can incentivize users to stay in the coalition to achieve k-anonymity. By using the optimal strategy from Nash Equilibrium, each user can preserve its privacy and also contribute to preserving privacy of other users.

The user group centric game model is utilized to find the optimal strategy of each user or incentivize users to stay in the group. Users in the group can have theoretical privacy protection guarantee. However, they need to consider the interaction between users and need a scheme to ensure all the users in the group are honest.

2.3 Three Parties Game Model

Most existing research only discusses the interaction between two parties: users and adversaries or user and platforms (or service providers). However, privacy issues include three parties: user, platform and adversary.

Adl et. al [26] analyze the trade-off between privacy and utility among three parties: data provider (user), data collector (platform) and data user (adversary). They utilize a sequential game to formulate the action of each party. In their game model, they assume the data user is the leader of the game and can choose the privacy parameter. Data collector and data provider has binary strategy: accept or reject. Obviously, their assumption is not suitable in the real world. Wang et. al. [27] investigate the interaction between three parties: user, application and adversary. They utilize a game model to find out the condition that can make the three parties have good behavior. However, in their game model, the three parties have only binary strategies. In another work of Wang et. al [28], they build a application ecosystem based on quantum game model. The proposed application ecosystem contains three parties: user, application and adversary. However, as the game model in [27], each party only has two states: cooperation state and defection state. The simple strategy space of each party is unpractical in a real scenario.

We build a three-party game model in this dissertation. The strategy of each party is formulated according to their practical actions. Therefore, the proposed model in this paper is more comprehensive and practical than existing works.

Chapter 3

PRIVACY PROTECTION AMONG THREE ANTITHETIC-PARTY FOR CONTEXT-AWARE SERVICES

In this chapter, we study the issues in existing game theory based research on privacy protection for context-aware services and propose a three party game model based framework to solve these issues.

3.1 Motivation

In the past years, privacy-preserving mechanisms have received a lot of attention from researchers. Besides cryptography, game theory has been widely applied as a strategic methodology to search for optimal strategies balancing the trade-off between the benefit of sharing data and cost of privacy disclosure [13–17, 19–29, 42, 43]. Notice that most of the existing research only focuses on the interaction between two opposite parties [13–17, 19–21], i.e., defender-attacker game model. In [25–29], various three-party game models are proposed. But, the game models of [25–29] are not "real" three-party models because they fail to depict the interaction between any two of the three parties, i.e. they considered either data resale by the platform or attacks by the adversary. Additionally, the schemes in [26–28] only provide a binary solution, specifically whether or not the user should submit their data to receive services.

Further exploring the mutual relationships among user, platform, and adversary would be more helpful for the user to defend against both the untrusted platform and the adversary. Moreover, it would be beneficial to produce a more fine-grained solution, so that a user could possibly provide obscured data and still receive adequate service. For this purpose, this chapter aims to *design a three-party game model among the three antithetic parties for users to simultaneously protect their privacy from untrusted service platforms and adversaries.*

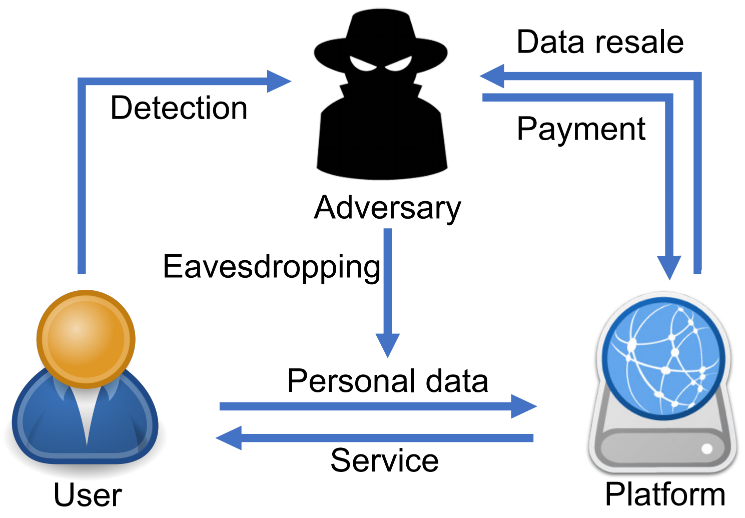

Figure 3.1. Structure of thee-party game.

Such a realistic and complicated game model challenges us in the following aspects: (i) *Complicated game structure.* As shown in Fig. 3.1, the interaction occurs between any two of the three parties, increasing difficulty in addressing the three parties' individual concerns – how does the user assess the potential risk of privacy loss and determine the granularity when submitting personal data; how does the platform determine data resale with consideration of the risk of reputation loss; and how does the adversary make a choice between purchase and eavesdropping? (ii) *Joint threats.* In such a complicated game, the user has to defend the joint threats from both the platform and the adversary, which may be hard to accomplish. (iii) *Multiple data attributes.* For many services, it is common that users need to submit multiple data attributes that could be correlated together. Any obscurity applied to one attribute would need to be correlated accordingly. (iv) *Theoretical analysis & solution.* Designing, analyzing, and solving the proposed three-party game are destined to be difficult due to the complexity of the game structure and correlated data attributes.

Our research endeavor to overcome the above challenges is briefly introduced as follows. Firstly, in our game model, we link the three parties elaborately quantifying their concerns and mutual interactions such that they are inseparable. Secondly, based on our game model, we perform a theoretical analysis to rigorously prove the optimal strategies of the three parties, including the optimal data release granularity for the user, the optimal data resale strategy for the platform, and the optimal probability to purchase data (or launch attack) for the adversary. Finally, we conduct simulations with real datasets under various settings to validate the effectiveness of our proposed game model.

To the best of our knowledge, we are the first to provide a fine-grained analysis on the behaviors and interactions for the user, platform, and adversary with considering resistance to the joint threats. Our major contributions are summarized as below:

- A three-party game is designed to capture the complicated interactions among user, platform, and adversary targeting defending the joint threats from both untrusted platform and adversary.
- An in-depth theoretical analysis is presented to identify the best strategy of each party.
- Comprehensive simulations with real datasets are exploited to evaluate the performance of our game model, regarding optimal strategy, cost, and utility of the three parties.

The rest of this chapter is organized as follows. Our game model is introduced in Section 3.2. The optimal strategy of each party and the performance of our game are analyzed in Section 3.3 and Section 3.4, respectively. Finally, Section 3.5 briefly concludes this chapter and discusses our future work.

3.2 Three-Party Game Model

In this section, the interaction among user, platform, and adversary is modeled as a three-party game, in which their strategies, benefits, and costs are mathematically formulated.

3.2.1 User Model

We consider the following scenario: a user submits personal dataset, denoted by $D = \{d_1, d_2, ..., d_n\}$, to a platform to acquire data-based service, where the dataset could contain one or more attributes and d_i $(1 \leq i \leq n)$ is the data of attribute i. Due to privacy concerns, the user may report data attributes with different data release granularity. Formally, the data release granularity of attribute i is defined as $g_i \in [0,1]$, and the corresponding data granularity set is $G = \{g_1, g_2, \ldots, g_i, \ldots, g_n\}$. Specifically, with a larger g_i, the data of attribute i is less obscured, revealing more personal/sensitive information; for examples, $g_i = 0$ if d_i does not contain any personal data, and $g_i = 1$ if d_i is fully accurate. In this chapter, we use data granularity as a measurement of data quality/obscurity.

As the data release granularity increases, the quality of user's requested service is increased with diminishing marginal benefit [44]. Suppose that the quality of attribute i-based service can achieve a maximum value q_i when $g_i = 1$. Then, the relationship between the quality of attribute i-based service and data release granularity g_i can be formulated to be $2q_i g_i - q_i(g_i)^2$. In addition, any two data attributes may correlate with each other, and such correlation can be exploited to infer more sensitive information [45,46]. Let e_{ij} represent the correlation between attribute i and attribute j. Due to correlations among data attributes, the data of attribute i not only contributes to the quality of attribute i-based service, but also contributes to the quality of attribute j-based service. Thus, given the user's dataset D, data release granularity set G, and data correlation $\{e_{ij}\}$, the overall service quality can be estimated as follows.

$$\sum_{i=1}^{n} \left(1 + \sum_{j=1, j \neq i}^{n} e_{ij} g_j\right) \left(2q_i g_i - q_i(g_i)^2\right). \tag{3.1}$$

While enjoying the service provided by the platform, privacy leakage incurred by data submission brings privacy loss to the user. One possible method for this privacy loss could be due to a malicious attack by an adversary that eavesdrops on the data submitted by the user. In real-world scenarios, the working efficiency of information retrieval is restricted by many factors, such as equipment performance and retrieval technique. The working efficiency of eavesdropping at the adversary side is denoted by $\phi \in [0,1]$, so the granularity of eavesdropped data is ϕg_i. Assume the adversary purchases data from the platform with probability b and the probability of eavesdropping is $1 - b$. Then, the expected cost due to eavesdropping of dataset D is defined as

$$(1-b)\sum_{i=1}^{n} c_i \phi g_i,$$

where c_i is the unit privacy cost when $g_i = 1$.

Another possible method for privacy loss could be that the user's submitted data is resold by the platform to a third-party (e.g., adversary) for more profit. We define the set of platform's resale strategy as $S = \{s_1, s_2, ..., s_n\}$, where $s_i \in [0,1]$ and $s_i g_i$ is the resold data granularity of attribute i. The platform does not resell d_i if $s_i = 0$ but resells all collected d_i if $s_i = 1$. The expected privacy cost due to data resale at the platform side can be computed by

$$b\sum_{i=1}^{n} c_i s_i g_i.$$

By combining the received service quality and the experienced privacy cost, the user's utility can be calculated in Eq. (3.2).

$$U_u = \lambda \sum_{i=1}^{n} \left(1 + \sum_{j=1, j\neq i}^{n} e_{ij} g_j\right) \left(2q_i g_i - q_i (g_i)^2\right) - (1-b)\sum_{i=1}^{n} c_i \phi g_i - b\sum_{i=1}^{n} c_i s_i g_i, \tag{3.2}$$

where λ is the convention rate between service quality and privacy cost, i.e., one unit of privacy cost is equivalent to λ units of service quality loss. Moreover, λ is also used to measure the user's privacy preference; that is, the user would care more about privacy cost than service quality if λ is large, and the service quality outweighs the privacy cost if λ is small.

In our proposed three-party game, the user aims to maximize its utility by balancing the trade-off between service quality and privacy cost by strategically setting the granularity set G. Accordingly, the optimization problem at the user side is

$$\max_{G} U_u,$$
$$\text{s.t. } g_i \in [0,1], i \in [1,n].$$

3.2.2 Platform Model

The platform provides users with requested services based on their submitted data. For instance, Google provides navigation service to users based on their input location.

While providing service to the user, the platform has its private valuation, defined to be V_p, for the collected data from the user. With user's data, the platform can obtain profit from data-based production, such as data statistic analysis and new product development. From the viewpoint that data is a type of potential productivity, the value of data can be computed according to the standard form of Cobb-Douglas production function [47] as

$$\theta_p \left(\sum_{i=1}^{n} g_i \right)^{\zeta_p},$$

where θ_p is the total value productivity of the platform, and $\zeta_p \in (0, 1)$ is the platform's value output elasticities of G.

To get extra benefits, the platform may resell the collected data to a third party (i.e., the adversary) [48, 49]. Assume that p_i is the unit data price of attributes i with $g_i = 1$, so the expected payment received from the adversary is

$$b \sum_{i=1}^{n} p_i s_i g_i, \tag{3.3}$$

in which b is the adversary's purchase probability and s_i represents the platform's resale strategy.

However, reselling the user's data may cause the risk of reputation loss at the user side and/or in public. According to the instantaneous risk function [50, 51], we can define the risk of reputation loss due to data resale of attribute i as

$$l_1 s_i g_i + l_2 \left(s_i g_i \right)^2,$$

where l_1 and l_2 are constant parameters of the risk estimation function. Since there may exist a correlation between two data attributes [52], the adversary can infer more personal/sensitive information from one data attribute to another, leading to an increase in the reputation loss at the platform side. Accordingly, the risk of reputation loss can be estimated as

$$\sum_{i=1}^{n} \left(1 + \sum_{j=1, j \neq i}^{n} e_{ij} s_j g_j \right) \left(l_1 s_i g_i + l_2 \left(s_i g_i \right)^2 \right).$$

In addition, there exists a data processing cost c_p at the platform side. Since the data processing cost may be determined by the processing technology, which is out of the scope of this chapter, we assume c_p is a system parameter for simplicity. Therefore, the platform's utility, denoted by U_p, can be defined to be

$$U_p = b\sum_{i=1}^{n} p_i s_i g_i + \theta_p \left(\sum_{i=1}^{n} g_i\right)^{\zeta_p} - c_p - \sum_{i=1}^{n}\left(1 + \sum_{j=1, j\neq i}^{n} e_{ij} s_j g_j\right)\left(l_1 s_i g_i + l_2 \left(s_i g_i\right)^2\right). \tag{3.4}$$

One can see that the platform faces a struggle between benefit and reputation cost from data resale. More specifically, reselling more accurate data can enhance the profit while damaging reputation, but reselling less accurate data can reduce reputation loss while losing attractiveness of data resale. Thus, to improve utility via balancing the trade-off between benefit and cost, the platform needs to choose a proper resale strategy S. Formally, the optimization problem of the platform is formulated as

$$\max_{S} U_p,$$
$$\text{s.t. } s_i \in [0,1], i \in [1,n].$$

3.2.3 Adversary Model

To retrieve the user's private information, the adversary could purchase data from the platform with probability b or eavesdrop on the communication between the user and the platform with probability $1-b$. With respect to each data attribute i, the granularity of purchased data is $s_i g_i$, and that of the eavesdropped data is ϕg_i.

The adversary also has private valuation for the obtained data. With the analysis similar to that in Section 3.2.2, we can utilize Cobb-Douglas production function [47] to compute adversary's private valuation as

$$b\theta_a \sum_{i=1}^{n} \left(s_i g_i\right)^{\zeta_a} + (1-b)\theta_a \sum_{i=1}^{n} \left(\phi g_i\right)^{\zeta_a},$$

where θ_a is the data productivity of the adversary and $\zeta_a \in (0,1)$ is the adversary's value output elasticities of data.

We suppose that the adversary can obtain all the data in D through eavesdropping at a cost (e.g., equipment and time) that can be quantified by a quadratic cost function [53, 54], i.e.,

$$\sigma_1(1-b)^2 + \sigma_2(1-b) + \sigma_3,$$

where $\sigma_1 > 0$, $\sigma_2 \geq 0$, and $\sigma_3 \geq 0$ are constant parameters of the quadratic cost function. Note that when the adversary does not eavesdrop, there still is a cost because it needs to purchase equipment and resources for eavesdropping. If the adversary chooses to purchase data from the platform, the expected payment paid to the platform is formulated in Eq. (3.3).

To sum up, the utility of the adversary, denoted by U_a, can be computed to be

$$U_a = b\theta_a \left(\sum_{i=1}^{n} s_i g_i\right)^{\zeta_a} + (1-b)\theta_a \left(\sum_{i=1}^{n} \phi g_i\right)^{\zeta_a} - b\sum_{i=1}^{n} p_i s_i g_i - \left(\sigma_1(1-b)^2 + \sigma_2(1-b) + \sigma_3\right). \tag{3.5}$$

In the three-party game, the adversary faces the trade-off between data purchase and data eavesdropping, i.e., the probability to purchase/eavesdrop data. Therefore, to improve utility, the adversary has to choose a proper purchase probability b to maximize its utility, which can be formulated as the following optimization problem.

$$\max_{b} U_a,$$
$$\text{s.t. } b \in [0,1].$$

3.3 Nash Equilibrium Analysis

In this section, we conduct in-depth theoretical analysis of the three parties' strategies and the relationships among their strategies.

3.3.1 Nash Equilibrium

In game theory, a Nash equilibrium is a strategy profile E^* with the property that no party can unilaterally do better by choosing an action different from E^*, given that other parties adhere to E^* [53]. Accordingly, the Nash equilibrium of our proposed three-party game can be defined in the following Definition.

Definition 1. *A strategy profile $E^* = (G^*, S^*, b^*)$ is called Nash Equilibrium for the proposed three-party game if the following properties simultaneously hold:*

$$U_u(G^*, S^*, b^*) \geq U_u(G, S^*, b^*);$$

$$U_p(G^*, S^*, b^*) \geq U_p(G^*, S, b^*);$$

$$U_a(G^*, S^*, b^*) \geq U_a(G^*, S^*, b).$$

3.3.2 Strategy Analysis of User

To solve the optimization problem of the user, we analyze the concavity of its utility function. The first-order partial derivative and the second-order partial derivatives of Eq. (3.2) are obtained, respectively.

$$\begin{aligned}\frac{\partial}{\partial g_i} U_u =& -(1-b)c_i\phi - bc_i s_i + \lambda \sum_{j=1, j\neq i}^{n} e_{ij}\left(-q_j(g_j)^2 + 2q_j g_j\right)\\ &+ \lambda\left(1 + \sum_{j=1, j\neq i}^{n} e_{ij} g_j\right)\left(-2q_i g_i + 2q_i\right).\end{aligned}$$

$$\frac{\partial^2}{\partial g_i^2} U_u = -2q_i\lambda\left(1 + \sum_{j=1, j\neq i}^{n} e_{ij} g_j\right).$$

$$\frac{\partial^2}{\partial g_i g_j} U_u = \lambda e_{ij}\left(-2q_j g_j + 2q_j\right) + \lambda e_{ij}\left(-2q_i g_i + 2q_i\right).$$

To find the maximum value, we need to solve the following system of equations.

$$\begin{cases} \frac{\partial}{\partial g_1} U_u = 0; \\ \frac{\partial}{\partial g_2} U_u = 0; \\ \qquad ... \\ \frac{\partial}{\partial g_n} U_u = 0. \end{cases} \tag{3.6}$$

All the solutions of the system of equations are the extreme points of user's utility. To find the global maximum value, we create the corresponding Hessian matrix:

$$H_u = \begin{vmatrix} \frac{\partial^2}{\partial {g_1}^2} U_u & \frac{\partial^2}{\partial g_1 \partial g_2} U_u & \cdots & \frac{\partial^2}{\partial g_1 \partial g_n} U_u \\ \frac{\partial^2}{\partial g_2 \partial g_1} U_u & \frac{\partial^2}{\partial {g_2}^2} U_u & \cdots & \frac{\partial^2}{\partial g_2 \partial g_n} U_u \\ \vdots & \vdots & \ldots & \vdots; \\ \frac{\partial^2}{\partial g_n \partial g_1} U_u & \frac{\partial^2}{\partial g_n \partial g_2} U_u & \cdots & \frac{\partial^2}{\partial {g_n}^2} U_u \end{vmatrix}.$$

The user has a maximum utility only if the matrix is a negative definite matrix. When either of the following two conditions holds, a matrix is negative definite [55]: (1) all its eigenvalues are less than 0; and (2) the even order principal minors are larger than 0 and odd order principal minors are less than 0. In other words, when the Hessian matrix of the user's utility function can meet anyone of the above two conditions, the user's optimal strategy can be found by solving Eq. (3.6).

We take the scenario where $e_{ij} = 0$ for $i, j \in [1, n]$ as an illustrative example. In this scenario, the first-order partial derivative and the second-order partial derivatives of the utility function are as follows.

$$\frac{\partial}{\partial g_i} U_u = \lambda \left(-2q_i g_i + 2q_i \right) - (1-b) c_i \phi - b c_i s_i.$$

$$\frac{\partial^2}{\partial g_i^2} U_u = -2q_i \lambda.$$

$$\frac{\partial^2}{\partial g_i g_j} U_u = 0.$$

Then we derive the corresponding Hessian matrix, i.e.,

$$H_u = \begin{vmatrix} -2q_1\lambda & 0 & ... & 0 \\ 0 & -2q_2\lambda & ... & 0 \\ \vdots & \vdots & ... & \vdots \\ 0 & 0 & ... & -2q_n\lambda \end{vmatrix}.$$

Because the even order principal minors are larger than 0 and the odd order principal minors are less than 0, the matrix H_u is a negative definite matrix. Therefore, the utility function has a maximum value and the maximum points can be calculated by solving Eq. (3.6), i.e.,

$$\begin{cases} g_1 = \frac{-c_1\phi + (\phi - s_1) b c_1 + 2q_1\lambda}{2q_1\lambda}; \\ g_2 = \frac{-c_2\phi + (\phi - s_2) b c_2 + 2q_2\lambda}{2q_2\lambda}; \\ \qquad\qquad ... \\ g_n = \frac{-c_n\phi + (\phi - s_n) b c_n + 2q_n\lambda}{2q_n\lambda}. \end{cases}$$

Since $g_i \in [0,1]$, the best data release granularity for attribute i is $g_i^* = \max\{\min\{g_i, 1\}, 0\}$.

From the results, to preserve data privacy, the user should decrease the data granularity g_i of attribute i if the platform increases data resale strategy s_i.

3.3.3 Strategy Analysis of Platform

We can compute the Hessian matrix to analyze the concavity of the platform's utility function as follows.

$$H_p = \begin{vmatrix} \frac{\partial^2}{\partial {s_1}^2} U_p & \frac{\partial^2}{\partial s_1 \partial s_2} U_p & \cdots & \frac{\partial^2}{\partial s_1 \partial s_n} U_p \\ \frac{\partial^2}{\partial s_2 \partial s_1} U_p & \frac{\partial^2}{\partial {s_2}^2} U_p & \cdots & \frac{\partial^2}{\partial s_2 \partial s_n} U_p \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial^2}{\partial s_n \partial s_1} U_p & \frac{\partial^2}{\partial s_n \partial s_2} U_p & \cdots & \frac{\partial^2}{\partial {s_n}^2} U_p \end{vmatrix}.$$

where

$$\frac{\partial^2}{\partial {s_i}^2} U_p = -2l_2 g_i^2 \left(1 + \sum_{j=1, j\neq i}^{n} e_{ij} s_j g_j\right),$$

and

$$\frac{\partial^2}{\partial s_i \partial g_j} U_p = -e_{ij} g_j \left(l_1 g_i + 2l_2 g_i^2 s_i\right) - e_{ij} g_i \left(l_1 g_j + 2l_2 g_j^2 s_j\right).$$

The platform has a maximum utility only if the Hessian matrix is a negative definite matrix that can satisfy either of the following two conditions [55]:
(1) all eigenvalues of H_p are less than 0; and (2) the even order principal minors of H_p are larger than 0 and odd order principal minors of H_p are less than 0.

If the maximum value exists, we can find the best strategy of the platform by solving the system of equations as shown below.

$$\begin{cases} \frac{\partial}{\partial s_1} U_p = 0; \\ \frac{\partial}{\partial s_2} U_p = 0; \\ \quad \cdots \\ \frac{\partial}{\partial s_n} U_p = 0. \end{cases} \tag{3.7}$$

In Eq. (3.7), we have

$$\frac{\partial}{\partial s_i} U_p = bp_i g_i - \left(1 + \sum_{j=1, j\neq i}^{n} e_{ij} s_j g_j\right) \left(l_1 g_i + 2l_2 g_i^2 s_i\right) - \sum_{j=1, j\neq i}^{n} e_{ij} g_i \left(l_1 s_j g_j + l_2 \left(s_j g_j\right)^2\right).$$

We use the scenario when $e_{ij} = 0$ $(i, j \in [0, 1])$ as an example for demonstration. In this scenario, the first-order partial derivative and the second-order partial derivatives of the utility function are obtained in the following.

$$\frac{\partial}{\partial s_i} U_p = bp_i g_i - l_1 g_i - 2l_2 g_i^2 s_i.$$

$$\frac{\partial^2}{\partial {s_i}^2} U_p = -2l_2 g_i^2.$$

$$\frac{\partial^2}{\partial s_i \partial g_j} U_p = 0.$$

Then we derive the Hessian matrix as:

$$H_u = \begin{vmatrix} -2l_2 g_1^2 & 0 & ... & 0 \\ 0 & -2l_2 g_2^2 & ... & 0 \\ \vdots & \vdots & ... & \vdots \\ 0 & 0 & ... & -2l_2 g_n^2 \end{vmatrix}.$$

Because the even order principal minors are larger than 0 and the odd order principal minors are less than 0, the matrix H_p is a negative definite matrix. Thus, the platform's utility function has a maximum value and the maximum points can be calculated by solving Eq. (3.7), i.e.,

$$\begin{cases} s_1 = \frac{bp_1 - l_1}{2l_2 g_1}; \\ s_2 = \frac{bp_2 - l_1}{2l_2 g_2}; \\ \qquad ... \\ s_n = \frac{bp_n - l_1}{2l_2 g_n}. \end{cases}$$

As $s_i \in [0, 1]$, the best resale strategy for attribute i is $s_i^* = \max\{\min\{s_i, 1\}, 0\}$.

According to the above analysis, to avoid too much reputation loss, the platform should decrease the value of s_i if the user increases g_i. Nevertheless, the granularity of resold data, $s_i g_i$, may be increased, bringing a profit increase to the platform. On the other hand, if the adversary prefers to purchase data rather than eavesdropping (i.e., enhance purchase probability b to a sufficiently large value), the platform can increase the value of s_i to earn more profit, in which the reputation loss maybe compensated by the payment from the adversary.

3.3.4 Strategy Analysis of Adversary

To maximize the utility, the adversary has to find the best strategy b^*. The first-order partial derivative and the second-order partial derivative of U with respect to b are respectively calculated by

$$\frac{dU_a}{db} = \theta_a \left(\sum_{i=1}^{n} s_i g_i \right)^{\zeta_a} - \theta_a \left(\sum_{i=1}^{n} \phi g_i \right)^{\zeta_a} - \sum_{i=1}^{n} p_i s_i g_i - \sigma_1(2b-2) + \sigma_2$$

and

$$\frac{d^2U_a}{db^2} = -2\sigma_1.$$

Since $\frac{d^2U_a}{db^2} = -2\sigma_1 < 0$, the utility function of the adversary is a concave function, which means the maximum value is achievable when $\frac{dU_a}{db} = 0$. Thus, by setting $\frac{dU_a}{db} = 0$, the solution is

$$b = \frac{\theta_a \left(\sum\limits_{i=1}^{n} s_i g_i \right)^{\zeta_a} - \theta_a \left(\sum\limits_{i=1}^{n} \phi g_i \right)^{\zeta_a} - \sum\limits_{i=1}^{n} p_i s_i g_i + 2\sigma_1 + \sigma_2}{2\sigma_1}.$$

Because $b \in [0, 1]$, the best purchase probability is $b^* = \max\{\min\{b, 1\}, 0\}$.

According to the above result, one can find that the purchase probability b is reduced when the working efficiency of eavesdropping ϕ and/or data price p_i increases. This is because the adversary prefers to eavesdrop rather than buying data for cost reduction if the granularity of eavesdropped data is higher than that of the purchased data. However, the relationship between the adversary's strategy and the platform's strategy and the relationship between the adversary's strategy and the user's strategy are not straightforward, because the purchase probability is also affected by the working efficiency ϕ, the price p_i for attribute i, the data productivity of the adversary θ_a, and the data value output elasticities of adversary ζ_a. These complicated relationships will be investigated in our simulations.

3.4 Simulation

In this section, we study the interactions among the user, the platform, and the adversary via intensive simulations. In this chapter, we assume that the user has multiple attributes in its dataset. However, in some cases, the user's dataset has only one attribute. To provide a detailed simulation result, we study the interactions among three parties in two scenarios: i) the user has one attribute in its dataset; ii) the user has more than one attribute in its dataset.

3.4.1 Interactions Among Three Parties with One Attribute

We first discuss the interaction among the three parties when the user's dataset has only one attribute, $D = \{d_1\}$. We select default settings for the parameters as follows:

The granularity of d_1 is $g_1 = 0.6$. The unit privacy cost due to leakage of d_1 is $c_1 = 3$. Based on d_1, the user can get maximum service quality $q_1 = 50$. The convention rate λ of the user is 0.1. The platform resells d_1 by using reselling strategy $s_1 = 0.6$ with the price $p_1 = 20$. The other parameters in the platform's utility function are: $\theta_p = 15, \zeta_p = 0.6, l_1 = 5, l_2 = 10, c_p = 1$. The adversary has a purchase probability $b = 0.6$ and working efficiency $\phi = 0.2$. The other parameters in the adversary's utility function are: $\sigma_1 = 1.5, \sigma_2 = 1, \sigma_3 = 1, \theta_a = 15, \zeta_a = 0.6$. The value of these parameters are chosen to reveal plain interactions among three parties.

The simulations that follow depict different strategies by varying certain parameters from the perspective of each of the three parties.

Simulation Result of User's Utility and Optimal Strategy The utility and optimal response of the user are investigated in this subsection. Fig. 3.2 to Fig. 3.4 reveal the simulation result of the user.

The results of the user's utility are presented in Fig. 3.2 and Fig. 3.3, from which we observe that the utility increases at first and then decreases as the granularity increases. The reason lies in two aspects: (i) when the granularity g_1 increases from 0 to a certain value

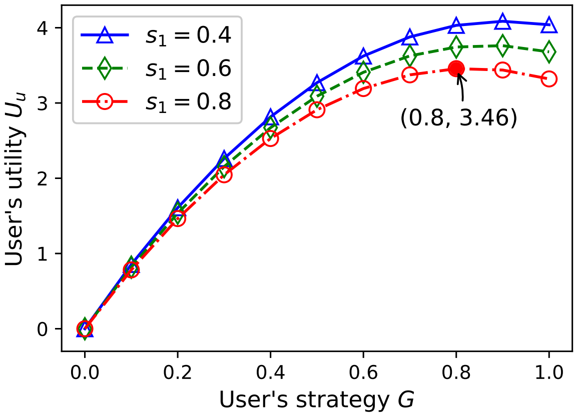

Figure 3.2. Utility of user under various G and S.

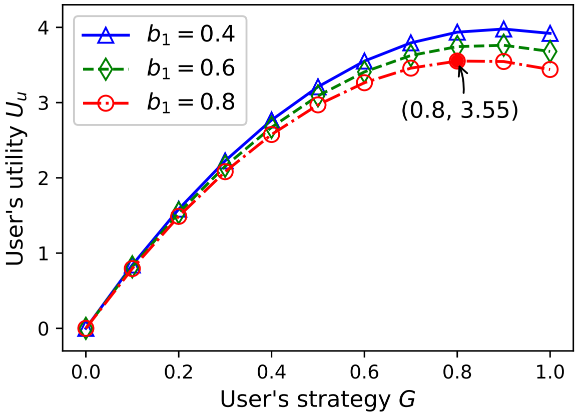

Figure 3.3. Utility of user under various G and b_1.

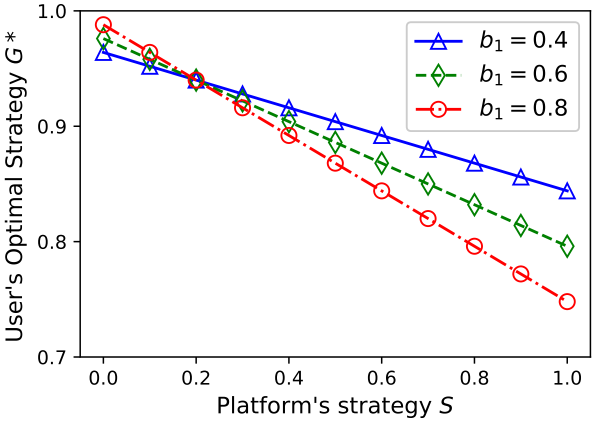

Figure 3.4. Optimal strategy of user under various S and b_1.

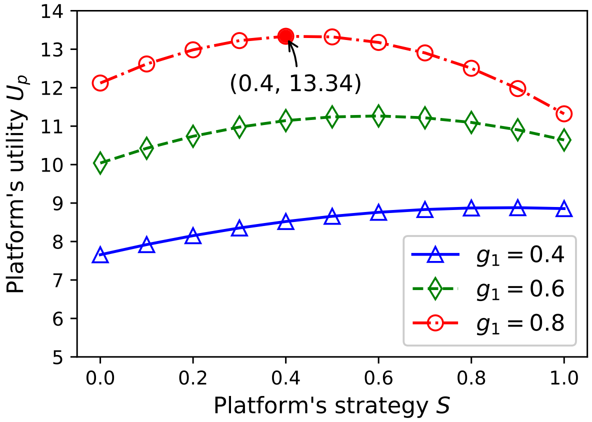

Figure 3.5. Utility of platform under various S and G.

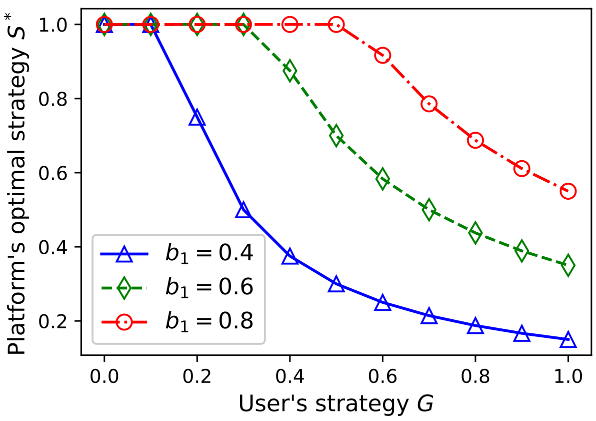

Figure 3.6. Optimal strategy of platform under various G and b_1.

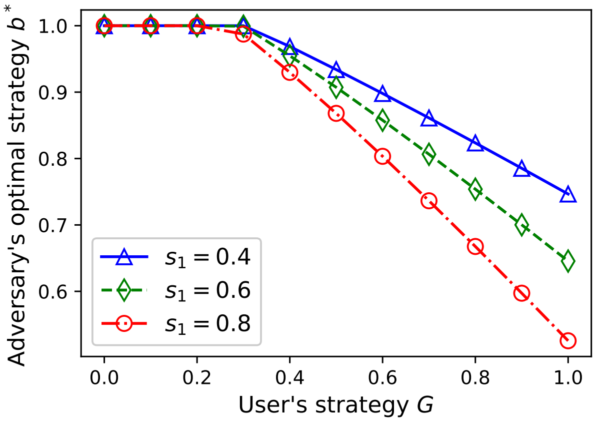

Figure 3.7. Optimal strategy of adversary under various G and b_1.

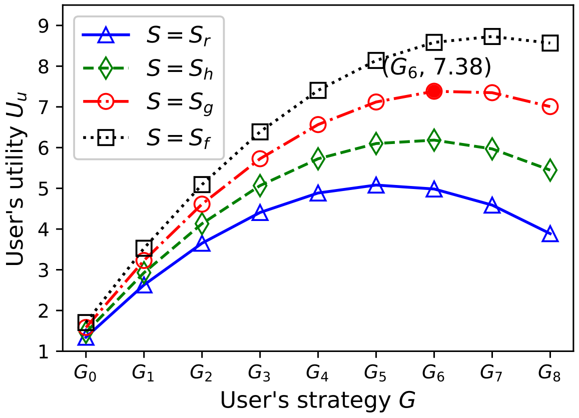

Figure 3.8. Utility of user under various G and S.

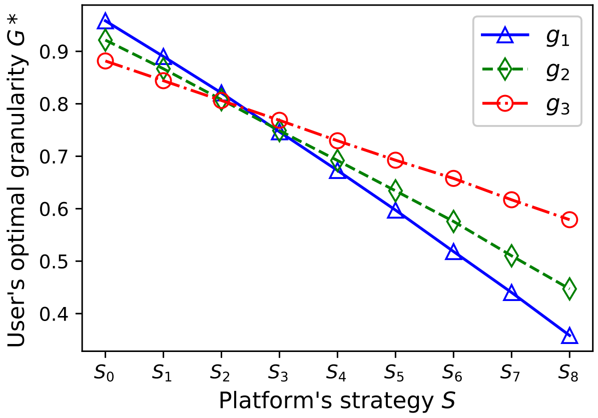

Figure 3.9. Optimal strategy of user under various S.

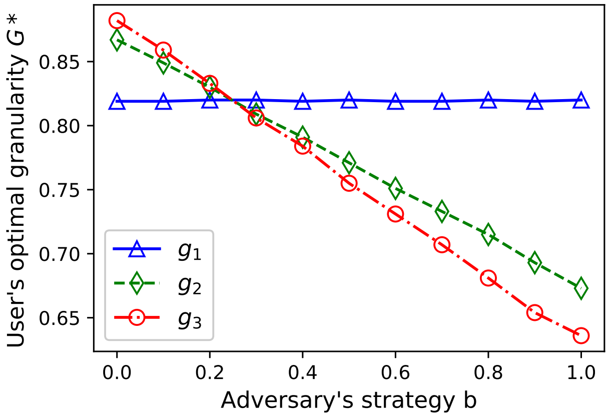

Figure 3.10. Optimal strategy of user under various b.

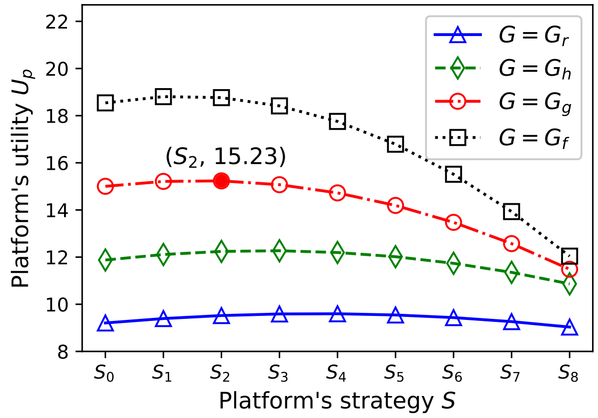

Figure 3.11. Uility of platform under various S and G.

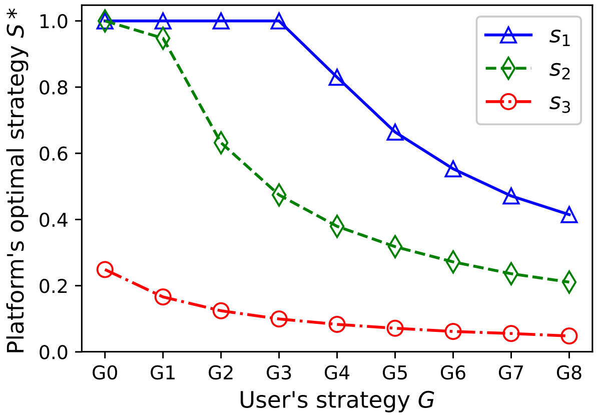

Figure 3.12. Optimal strategy of platform under various G.

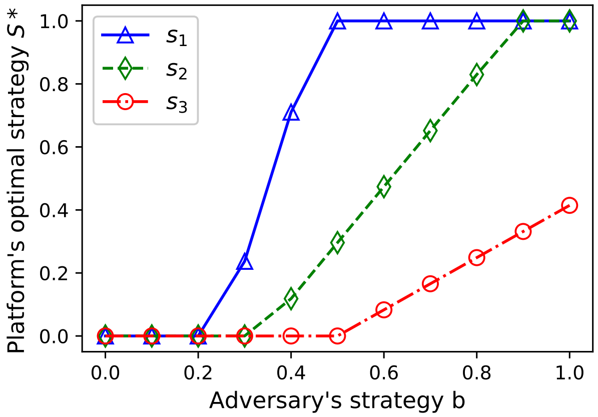

Figure 3.13. Optimal strategy of platform under various b.

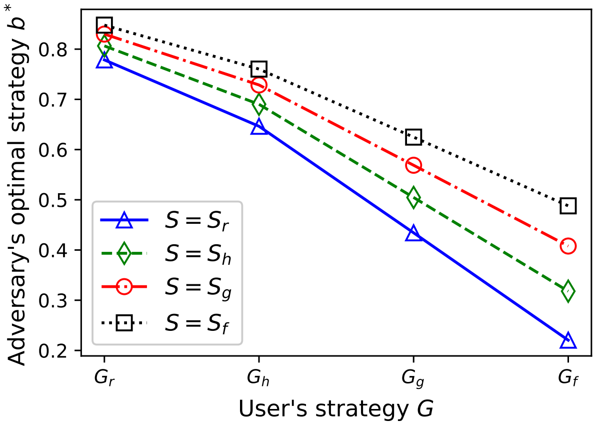

Figure 3.14. Optimal strategy of adversary under various G and b.

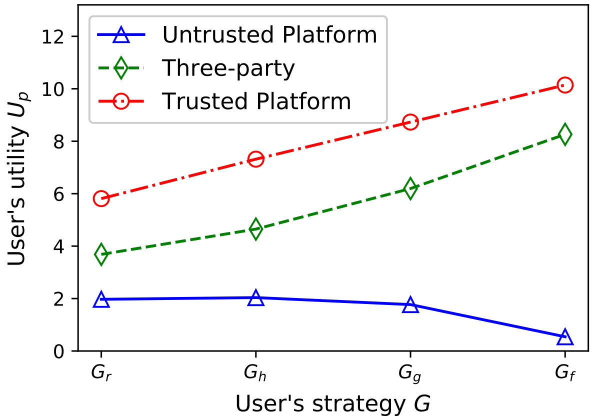

Figure 3.15. Utility comparison of user under various platform.

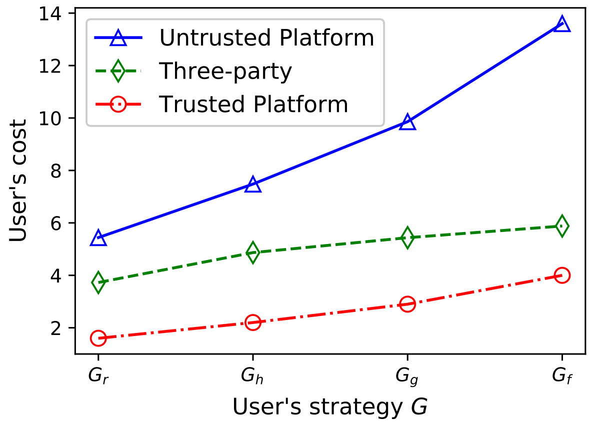

Figure 3.16. Cost comparison of user under various platform.

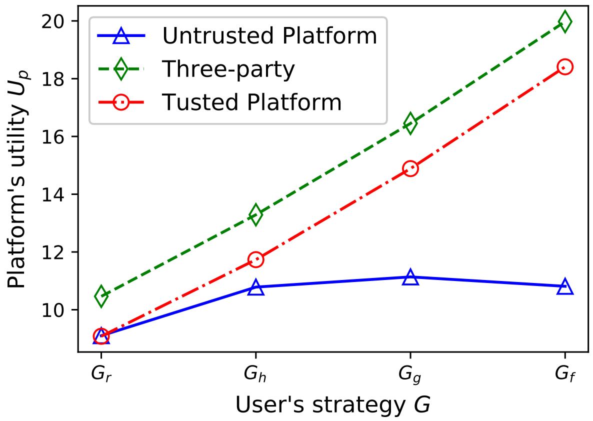

Figure 3.17. Utility comparison of various platforms.

(e.g., 3.46 in line $s_1 = 0.8$ of Fig. 3.2 and 3.55 in line $b_1 = 0.8$ of Fig. 3.3), the increase rate of privacy cost is smaller than that of received service quality, therefore the utility increases; and (ii) when g_1 continues increasing from such a certain value, the increase rate of privacy cost is larger than that of received service quality, leading to a decrease in the utility. In fact, such a certain value corresponds to the optimal granularity.

Besides, as shown in Fig. 3.2, the user's utility U_u decreases as the platform increases the value of its reselling strategy s_1. This is because when the platform increases the value of reselling strategy s_1, the granularity of the reselling data increases. That increases user's privacy cost, resulting in decreasing of user's utility.

Furthermore, the user's utility U_u also decreases as adversary increases its purchase probability as shown in Fig. 3.3. Because the purchased data of the adversary has a higher granularity than eavesdropped data, when the adversary increases the probability of purchase and decreases the probability of eavesdropping, the user has more privacy cost, leading to a decrease in the user's utility.

Fig. 3.4 states the optimal strategy of the user. We can see that the user decreases data granularity if the platform increases the value of reselling strategy s_1. When the platform increases the value of reselling strategy s_1, the granularity of the reselling data increases, thus increasing user's privacy cost. To reduce privacy cost, the user should decrease data granularity as shown in Fig. 3.4.

Fig. 3.4 also reveals how the user adjusts its strategy when the adversary uses different strategies. In Fig. 3.4, we can see that the three lines ($b_1 = 0.4$, $b_1 = 0.6$, and $b_1 = 0.8$) intersect at the point where $s_1 = \phi = 0.2$. When $s_1 = \phi$, the adversary's purchased data has the same data granularity with eavesdropped data, the user has the same privacy cost no matter what the adversary prefers, purchase or eavesdropping. Thus, the user does not need to change its data granularity as the adversary change its strategy when $s_1 = \phi$.

Fig. 3.4 shows that the user adjusts its strategy according to the adversary's strategy as well as the platform's strategy: (i) the user decreases the granularity of the data as the adversary decreases the probability of data purchase and increases the probability of eaves-

dropping when the platform resells data with strategy $s_1 < \phi$. (ii) the user decreases the granularity of the data as the adversary increases the probability of data purchase and decreases the probability of eavesdropping when the platform resells data with strategy $s_1 > \phi$. The reason lies in two aspects: (i) when $s_1 < \phi$, the adversary's eavesdropped data has a higher granularity than purchased data. Thus, the eavesdropping causes more privacy cost than data reselling to the user. To decrease privacy cost, the user should decrease the data granularity if the adversary increases the probability of eavesdropping and decreases the probability of data purchase.
(ii) On the contrary, when $s_1 > \phi$, the adversary's purchased data has a higher granularity than eavesdropped data. Thus, the data reselling causes more privacy cost than eavesdropping to the user. To decrease privacy cost, the user should decrease the data granularity if the adversary decreases the probability of eavesdropping and increases the probability of data purchase.

Simulation Result of Platform's Utility and Optimal Strategy We then study the utility and best response of the platform. The results are shown in Fig. 3.5 and Fig. 3.6.

From Fig. 3.5, we can tell that the platform's utility increases at first and then decreases with the increase of reselling strategy $S = \{s_1\}$. When s_1 increases from 0 to a certain value (e.g., 13.34 in line $g_1 = 0.8$), the increase rate of the cost is smaller than that of the profit, resulting in an improvement of utility; however, when s_1 continues increasing from such a certain value, the increase rate of the cost is larger than that of the profit, further reducing the utility. In other words, there is an optimal value of s_1 for the platform to balance the profit of data resale and cost of reputation loss.

In Fig. 3.5, when the data granularity g_1 increases, the granularity of reselling data increases and brings more profit to the platform, leading to the increase of utility of the platform.

Fig. 3.6 reveals how the platform adjusts its optimal strategy when the user and adversary choose different strategies. When the user's granularity increases, the optimal strategy of the platform decreases. Both the profit and the reputation loss increase if the user increases granularity. However, the higher profit cannot make up the increased reputation loss. Thus, the platform should decrease the value of s_1 to reduce reputation loss.

Moreover, Fig. 3.6 reveals that the optimal strategy of the platform increases if the adversary increases its purchase probability b and decreases its eavesdropping probability $1 - b$. When the adversary increases the purchase probability b, the expected payment to the platform increases. Thus, to earn more profit, the platform increases the value of s_1 as the adversary increases the purchase probability b, as shown in Fig. 3.6.

Simulation Result of Adversary's Optimal Strategy The study of the adversary's optimal strategy is shown in Fig. 3.7. From this figure, we can see that the optimal strategy of the adversary decreases as the user increases the granularity g_1 or the platform increases the value of its reselling strategy s_1. When the user increases the granularity g_1 or the platform increases the value of its reselling strategy s_1, the granularity of reselling data increases correspondingly, resulting in the increasing of data's price. To decrease the payment, the adversary decreases the probability of data purchase (which also increases the probability of eavesdropping) when the user increases the granularity g_1, or the platform increases the value of its reselling strategy s_1 as shown in Fig. 3.7.

3.4.2 Interactions Among Three Parties with Multiple Attributes

According the aforementioned analysis, the theoretical optimal strategies of the user and the platform may not always exist. For computation feasibility, we utilize a parallel machine learning algorithm termed Particle Swarm Optimization (PSO) [56], to find the quazi-optimal strategies for the user and the platform, which is performed in the following manner: (i) we run the simulation 100 times; (ii) in each time, each initial strategy and the update vector in each iteration are randomly generated. and (iii) we use the strategy which has the largest utility as the final output. The results derived from PSO are consistent

Table 3.1. Extracted strategies

Application	Strategy setting of {Income, Age, Race}
Retail	G_r={0.2, 0.3, 0.4}, S_r={0.5, 0.7, 0.8}
Healthcare	G_h={0.3, 0.4, 0.5}, S_h={0.4, 0.6, 0.7}
Government	G_g={0.4, 0.5, 0.7}, S_g={0.3, 0.5, 0.6}
Financial	G_f={0.6, 0.7, 0.8}, S_f={0.2, 0.4, 0.5}

with that in single attribute scenario, thus validating the simulation result. For more details about the implementation of PSO, please refer to [57].

We use real datasets as the inputs of the user and platform. More specifically, based on the privacy survey published by IBM [58] and Data Protection Survey published by SANA [59], we extract four user's strategies and four platform's strategies. As shown in Table 3.1, G_r, G_h, G_g, and G_f, are the user's strategies used for Retail applications, Healthcare applications, Government applications, and Financial applications, respectively. Correspondingly, in Table 3.1, S_r, S_h, S_g, and S_f are the strategies of Retail platforms, Healthcare platforms, Government platforms, and Financial platforms, respectively.

Each extracted strategy contains three data attributes, including income, age, and race. 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. For the three data attributes, the values of maximum service quality are $q_1 = 60$, $q_2 = 50$, and $q_3 = 40$, the unit privacy costs are $c_1 = 15$, $c_2 = 10$, and $c_3 = 5$, and the unit data prices are $p_1 = 20$, $p_2 = 15$, and $p_3 = 10$. The convention rate λ in Eq. (3.2) is 0.1. The other parameters in the platform's utility function are: $\theta_p = 15$, $\zeta_p = 0.6$, $l_1 = 5$, $l_2 = 10$, and $c_p = 1$. The adversary has a purchase probability $b = 0.6$ and working efficiency $\phi = 0.2$. The other parameters in the adversary's utility function are: $\sigma_1 = 1.5$, $\sigma_2 = 1$, $\sigma_3 = 1$, $\theta_a = 15$, and $\zeta_a = 0.6$.

The simulations that follow depict different strategies by varying certain parameters from the perspective of each of the three parties.

Table 3.2. Strategy simulation setting

Notation	Strategy Setting
G_0, S_0	{0.0, 0.1, 0.2}
G_1, S_1	{0.1, 0.2, 0.3}
G_2, S_2	{0.2, 0.3, 0.4}
G_3, S_3	{0.3, 0.4, 0.5}
G_4, S_4	{0.4, 0.5, 0.6}
G_5, S_5	{0.5, 0.6, 0.7}
G_6, S_6	{0.6, 0.7, 0.8}
G_7, S_7	{0.7, 0.8, 0.9}
G_7, S_7	{0.8, 0.9, 1.0}

Results of User's Utility and Optimal Strategy The utility and optimal response of the user are investigated through Fig. 3.8 to Fig. 3.10 in this subsection.

We analyze user's utility by using real platform's strategy S_r, S_h, S_g, and S_f and increasing the granularity of each attribute from G_0 to G_8 as shown in Table 3.2. The results of the user's utility are presented in Fig. 3.8, from which we observe that the utility increases at first and then decreases as the granularity increases. The reason lies in two aspects: (i) when the granularity of each attribute in user's granularity set G increases from G_0 to a certain granularity set (e.g., G_6 in line $S = S_g$), the increase rate of privacy cost is smaller than that of received service quality, therefore the user's utility increases; and (ii) when the granularity of each attribute in user's granularity set G continues increasing from such a certain strategy set, the increase rate of privacy cost is larger than that of received service quality, leading to a decrease in the utility. In fact, such a certain granularity set corresponds to the optimal granularity set among granularity sets from G_0 to G_8.

Also, as shown in Fig. 3.8, the user's utility U_u decreases as the platform increases the value of resale strategy of each attribute from S_r to S_f. This is because when the platform increases the value of resale strategy, the granularity of the resale data increases, enhancing user's privacy cost and reducing user's utility. In particular, the user can gain a larger utility when using Financial Application than other applications because Financial Platform resells user's data with the lowest granularity.

The user's best strategies defending against the platform's different strategies and adversary's different strategies are respectively shown in Fig. 3.9 and Fig. 3.10, where each line indicates the user's optimal data release granularity for one attribute. The results of Fig. 3.9 are obtained when the platform uses the strategies $s_1 = 0.2$, $s_2 = 0.4$, and $s_3 = 0.6$. The results of Fig. 3.10 are obtained when the adversary adopts the data resale strategies $b = 0.2, 0.4, 0.6$.

Fig. 3.9 shows that the user decreases data release granularity to protect data privacy as the platform increases the data resale granularity from S_0 to S_8. On the other hand, when the value of resale strategy is small (e.g., S_0, S_1, S_2), the user's optimal release granularity of data attribute 1 (corresponding to g_1) is larger than those of attributes 2 and 3 (corresponding to g_2 and g_3, respectively). Since data attribute 1 has the largest service quality value q_1, the user can get more profit from submitting data attribute 1, which can compensate the cost of privacy loss. On the contrary, when the value of resale strategy becomes large (e.g., $S_3, S_4, S_5, S_6, S_7, S_8$), releasing attribute 1 causes more privacy cost as data attribute 1 has the largest unit privacy cost c_1. As a result, to reduce privacy cost, the user releases less information regarding data attribute 1, i.e., the optimal release granularity of attribute 1 is less than those of other two attributes.

From Fig. 3.10, we can see that the optimal release granularity g_1 does not change as the adversary changes its strategy, and the optimal release granularities g_2 and g_3 decreases when the adversary increases the data purchase probability and decreases the eavesdropping probability.

When the platform resells attribute 1 (i.e., the line corresponds to g_1), the platform uses resale strategy $s_1 = 0.2$ that is the same as the working efficiency of eavesdropping ϕ. This means the granularity of purchased data and that of eavesdropped data are equal at the adversary side. Thus, the change of the adversary's strategy will not affect the user's utility. This is the reason why the best release granularity of data attribute 1 does not change when the adversary changes its strategy b under the scenario $s_1 = \phi$ as we discussed for Fig. 3.4.

On the other hand, the adversary can get higher data granularities of attribute 2 and 3 (corresponding to g_2 and g_3, respectively) via purchasing rather than eavesdropping because $s_2 > \phi$ and $s_3 > \phi$. Thus, the data resale from the platform causes more privacy cost than eavesdropping to the user. As a result, the best data release granularities g_2 and g_3 decreases as the adversary increases the data purchase probability and decreases the eavesdropping probability.

The changes of user's optimal release strategies in Fig. 3.9 and Fig. 3.10 confirm that data release strategy is also affected by data diversity (e.g., different data attributes have different privacy costs and resale prices).

Results of Platform's Utility and Optimal Strategy Fig. 3.11 reports the results of the platform's utility, in which the platform's strategies are set to be S_0 to S_8 according to Table 3.2 with user's strategy being G_r, G_h, G_g, and G_f as listed in Table 3.1. From Fig. 3.11, we can tell that the platform's utility increases at first and then decreases over the increase of the value of user's data release granularity, as we discussed for Fig. 3.5. When the data resale granularity of each attribute increases from the value in S_0 to the value in a certain resale strategy set (e.g., S_2 in line $G = G_g$), the increase rate of the cost is smaller than that of the profit, resulting in an improvement of utility to the platform; however, when the data resale granularity of each attribute continues increasing from the value in such a certain strategy set, the increase rate of the cost is larger than that of the profit, further reducing the platform's utility. In other words, there is an optimal resale strategy set for the platform to balance the profit of data resale and cost of reputation loss.

Besides, the platform's utility increases as the user increases the release granularity of each attribute from G_r to G_f. This is because the platform can get more accurate data and more resale profit if the user increases the data release granularity. Particularly, the Financial platform can the highest utility because the user submits data with higher granularity to the Financial platform than other three platforms.

The optimal strategy for reselling each data attribute at the platform side are drawn in Fig. 3.12, from which one can observe that the optimal resale granularity of each attribute decreases when the corresponding data release granularity rises. Notice that as the data release granularity of each attribute increases, the growth rate of reputation cost from data resale becomes larger than the growth rate of the profit from data resale. Therefore, to reduce reputation cost, the platform decreases the resale granularity of each attribute.

Fig. 3.13 presents the changes of platform's strategy when the adversary enhances the purchase probability. We can see that the optimal resale strategy for each attribute increases as the purchase probability is increased (i.e., the eavesdropping probability is reduced). When the adversary increases the probability of purchase (decreases the probability of eavesdropping), the increase rate of data resale profit is larger than the cost of reputation loss. Thus, to get more profit, the platform increases the data resale granularity.

In Fig. 3.12 and Fig. 3.13, the value of optimal resale strategy of attribute 1 is higher than that of other attributes, for which the reasons lie in two aspects: (i) more information regarding data attribute 1 is released from the user (see Fig. 3.10), indicating that less obscured data is available for resale; and (ii) the unit price of attribute 1 is larger than those of other two attributes, implying that more payment can be received by reselling data of attribute 1. These results reflect the fact that a platform may adopt different resale strategies for different data attributes due to data diversity.

Simulation Result of Adversary's Optimal Strategy Fig. 3.14 shows the adversary's optimal strategy when the user and the platform use the strategies from Table 3.1. As shown in Fig. 3.14, the adversary decreases the data purchase probability (i.e., increases the eavesdropping probability) when the user increases the data release granularity of each attribute in the dataset from G_r to G_f, or the platform increases the data resale strategy of each attribute in the dataset from S_f to S_r. The increase of data release/resale granularity will raise the data price, so the adversary need to decreases the data purchase probability to save cost, which is the same as shown in Fig. 3.7.

Comparison with Two-Party Game In this subsection, a comparison between our proposed three-party game and the existing two-party game is performed. According to current research [13–17], there are two types of platforms in two-party game: (i) trusted platform that never resells user's data, and (ii) untrusted platform that resells all collected data.

Fig. 3.15 and Fig. 3.16 show the user's utilities and costs, respectively. On one hand, the user has the highest utility and the lowest cost when the platform is trusted because the potential privacy risk of data resale is ignored and the privacy cost is underestimated. On the other hand, the user has the lowest utility and highest cost when the platform is untrusted because the untrusted platform resells all its collected data and brings more privacy loss to the user. However, the trusted platform is an ideal model, and the untrusted platform, which sells off all data is rare in real life. A more realistic model is a platform that chooses the optimal strategy by balancing the tradeoff between profit and cost. Our three-party game model, captures the actions of this more realistic model.

From the comparison of platform's utility in Fig. 3.17, it can be seen that a platform can get the highest utility by adaptively reselling user's data; specifically, by adaptively reselling user's data, a platform can get more profit than the trusted platform and suffers less reputation cost than the untrusted platform. This is consistent with the fact that a platform usually owns the ability to adjust its strategy for enhancing profit, further confirming the effectiveness of our proposed game model.

3.5 Conclusion

This chapter studies privacy preservation for context-aware services. To provide realistic optimal strategies for both the user and the platform, we propose a *three-party game model* that captures the interactions between any two of the parties: user, platform and adversary. Interactions include privacy leakage and data resale at the platform side, as well as malicious attacks at the adversary side. Our solution determines an *optimal fine-grained strategy* for the user and platform, so that the user can choose an optimal data granularity to balance service quality and privacy leakage and that the platform can choose the optimal reselling strategy to balance profit and reputation loss. Our model also accounts for the correlations between *multiple data attributes* provided by a user.

To find out the best strategy for each party, we conduct a rigorous theoretical analysis. We also perform simulations using real datasets to validate the effectiveness of our proposed game model.

In the next chapter, we extend this work to an m-user scenario, where the interconnected behaviors of the multiple users, the platform, and the adversary become more complex than the proposed model. This extension will also need to involve another layer of interactions between the users themselves, further complicating the model.

Chapter 4

PRIVACY PROTECTION FOR CONTEXT-AWARE SERVICES: A TWO-LAYER THREE-PARTY GAME MODEL

In this chapter, we extend our work to a more realistic model which considers the interaction between multiple users [60–63], thus providing more practical guidance for context-aware service users.

4.1 Motivation

We have studied the three party interaction in the previous chapter. However, in most context-aware services, there are multiple users using a service at the same time. The current n-player game models (those with n users) [19–24] 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 chapter, 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. 4.1). 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

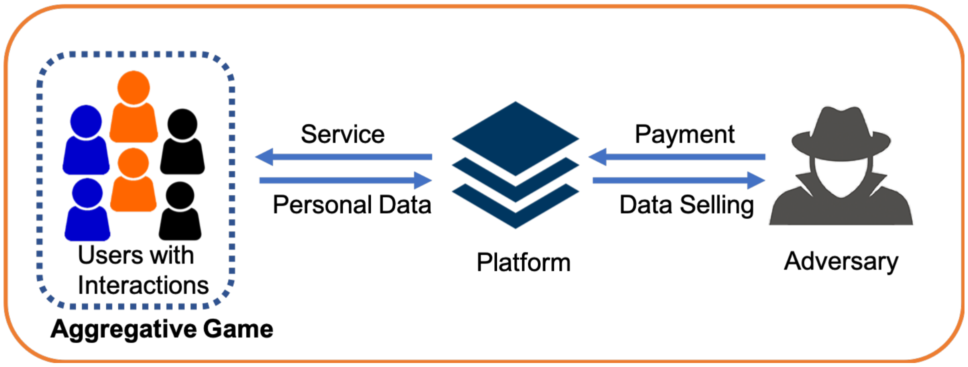

Figure 4.1. Two-layer three-party Game model

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 game layers for the game model a game among asymmetric users and a game 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 chapter. 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 chapter 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 chapter is organized to introduce the system model in Section 5.3. We analyze the optimal strategies for asymmetric users and platforms in Section 4.4. Section 5.5 presents the simulations to validate the theoretical analysis and we conclude the chapter in Section 5.6.

4.2 Preliminary

Let $\Gamma = (\tilde{\pi}_i, S_i)_{i \in \mathscr{I}}$ denote 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 2. *(Quasi-Aggregative Game) [64] 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:*

$$\tilde{\pi}_i = \pi_i \left(\sigma_i \left(s_{-i}, s_i \right), s_i \right), \tag{4.1}$$

where $\pi_i : X_{-i} \times S_i \to \mathbb{R}$, and:

$$g(s) = F_i \left(\sigma_i(s_{-i}), s_i) \right), \forall s \in S, i \in \mathscr{I}. \tag{4.2}$$

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.*

Assumption 1. *Each of the correspondences $R_i : X_{-i} \to 2^{S_i}$ 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.*

4.3 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.

4.3.1 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 [7]. 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 $\vec{\delta} = \{\delta_1, \delta_2, ..., \delta_m\}$ for all the users [65–68]. 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 $\vec{\delta} = \{1, 1, ..., 1\}$.

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 $\vec{\delta}$ with lower privacy protection level.

Users have the right to accept or reject the platform's offer $\vec{\delta}$. 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 $\vec{a_i} = \{a_{i1}, a_{i2}, ..., a_{im}\}$ and the strategy vector of all users is $\vec{a} = \{\vec{a_i}, \vec{a_j}, ..., \vec{a_n}\}$.

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 $\vec{a}_i$, but also on the strategy of other users $\vec{a}_{-i}$. 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 $Q_i\left(\vec{a}_{-i}, \vec{a}_i\right)$.

Meanwhile, the platform may resell users' data to an adversary resulting in privacy loss to the users. Assume each user has a constant privacy cost estimation vector $\vec{c_i} = \{c_{i1}, c_{i2}, ..., c_{im}\}$, 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:

$$C_i^u\left(\vec{a}_i\right) = \sum_{j=1}^{m} c_{ij} a_{ij} \left(s_j + \left(1 - \delta_j\right)\right), \tag{4.3}$$

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.

$$U_i^u\left(\vec{a}_i, \vec{a}_{-i}\right) = Q_i\left(\vec{a}_{-i}, \vec{a}_i\right) - C_i^u\left(\vec{a}_i\right). \tag{4.4}$$

4.3.2 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 users to accept the offer with higher privacy leakage level by providing more accurate service quality. We define $\sigma_j(\vec{a})$ as the expected number of users that accept the information leakage level δ_j for attribute j, and calculate $\sigma_j(\vec{a})$ as

$$\sigma_j(\vec{a}) = \sum_{i=1}^{n} a_{ij}. \tag{4.5}$$

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 [65,69,70], 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

$$log((1-\delta_j)+1)\sigma_j^b(\vec{a}), \tag{4.6}$$

where $0 < b < 1$ is the parameter revealing the impact of $\sigma_j^b(\vec{a})$. 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

$$Q\left(\vec{\delta},\vec{a}\right) = \sum_{j=1}^{m}\left(1+\sum_{k=1,k\neq j}^{m} e_{jk}\right) log((1-\delta_j)+1)\sigma_j^b(\vec{a}). \tag{4.7}$$

The collected dataset from users can generate income for the platform. The expected income form data is also affected by the privacy leakage level $\vec{\delta}$ and the number of users who accept the platform's offer. According to data aggregation research [71] and the standard form of Cobb-Douglas production function [47], the expected data value to the platform can be defined as

$$V^p\left(\vec{\delta},\vec{a}\right) = \alpha\sum_{j=1}^{m}(1-\delta)_j^{\zeta}\sigma_j^b(\vec{a}), \tag{4.8}$$

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 $\vec{s} = \{s_1, s_2, ..., s_m\}$ 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 $\vec{p} = \{p_1, p_2, ...p_m\}$, which is determined in a contract with the adversary. Thus, the total expected price is defined as

$$P\left(\vec{s}, \vec{p}, \vec{a}\right) = \sum_{j=1}^{m} p_j s_j \sigma_j^b(\vec{a}). \tag{4.9}$$

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

$$\sum_{j=1}^{m} r_j s_j \sigma_j^b(\vec{a}). \tag{4.10}$$

Meanwhile, the platform has a constant running cost c_p. Thus, the total expected cost of the platform is

$$C^p\left(\vec{s}, \vec{a}\right) = \sum_{j=1}^{m} r_j s_j \sigma_j^b(\vec{a}) + c_p. \tag{4.11}$$

To sum up, the expected utility of the platform is

$$U^p\left(\vec{\delta}, \vec{s}, \vec{p}, \vec{a}\right) = V^p\left(\vec{\delta}, \vec{a}\right) + P\left(\vec{s}, \vec{p}, \vec{a}\right) - C^p\left(\vec{s}, \vec{a}\right). \tag{4.12}$$

The platform will maximize its utility by achieving a Nash Equilibrium with the users and adversary.

4.3.3 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 [71] and the standard form of Cobb-Douglas production function [47], the expected data value to the adversary can be defined as

$$V_t\left(\vec{s}, \vec{a}\right) = \theta \sum_{j=1}^{m} s_j^{\gamma} \sigma_j^{b}(\vec{a}). \tag{4.13}$$

Thus, the expected utility function of the third party is

$$U^t\left(\left(\vec{p}(\gamma), \vec{s}\right), \vec{a}\right) = V_t\left(\vec{s}, \vec{a}\right) - P\left(\left(\vec{p}(\gamma), \vec{s}\right), \vec{a}\right). \tag{4.14}$$

4.4 Game Model

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

4.4.1 Aggregative Game Model

In this chapter, we assume users do not exchange information with the other users. However, each user's action influences the other users' utility. With a specific privacy leakage level $\vec{\delta}$, 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 [64], 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

$$\tilde{\pi}_{ij} = U_{ij}^{u}(\sigma_{ij}(\vec{a}_{-i}), a_{ij}); \tag{4.15}$$

the aggregator can be defined as

$$g_j(\vec{a}) = F_{ij}(\sigma_{ij}(\vec{a}_{-i}), a_{ij}) = \sigma_{ij}(\vec{a}_{-i}) + a_{ij}; \tag{4.16}$$

the interaction functions vector can be defined as

$$\sigma_{ij}(\vec{a}_{-i}) = \sum_{k \in N, k \neq i} a_{kj}. \tag{4.17}$$

User i in the game Γ_j aims to maximize its utility by properly choosing a strategy vector $\vec{a_i}$ such that

$$\vec{a_i} = arg \max_{a_{ij}} U_i^u(\vec{\sigma}_i(\vec{a}_{-i}), a_{ij}). \tag{4.18}$$

According to the property of quasi-aggregative game theory [64], we can derive the following theorem.

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

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 obvious that the shift function F_i (Eq. 4.16) exhibits strictly increasing differences in x_{-i} and s_i. According to [64], the theorem is proved. □

4.4.2 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 $\{(\vec{p}(\gamma), \vec{s})\}$ to the adversary. According to contract theory [72], 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 $\{(\vec{p}(\gamma), \vec{s})\}$ satisfies the individual rationality constraints if it yields to the adversary a non-negative payoff, i.e., $\forall \gamma \in (0, 1)$,*

$$U^t(\vec{p}(\gamma), \vec{s}) \geq 0, \tag{4.19}$$

where $U^t(\vec{p}(\gamma), \vec{s})$ is the utility of adversary with type γ.

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

$$U^t(\vec{p}(\gamma), \vec{\delta}) \geq U^t(\vec{p}(\hat{\gamma}), \vec{s}). \tag{4.20}$$

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.

$$\begin{gathered} \max_{\{(\vec{p}(\gamma), \vec{s})\}} U^p\left(\vec{\delta}, \vec{s}, \vec{p}(\gamma), \vec{a}\right), \\ \textit{subject to Condition } 1 \textit{ and } 2. \end{gathered} \tag{4.21}$$

According to the aggregative model and contract model, we can see that the platform needs to properly choose the privacy leakage level $\vec{\delta}$ 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:

$$\begin{gathered} \max_{\left(\vec{\delta}, \{(\vec{p}(\gamma), \vec{s})\}\right)} U^p\left(\vec{\delta}, \vec{s}, \vec{p}(\gamma), \vec{a}^*\right), \\ \textit{subject to Condition } 1 \textit{ and } 2. \end{gathered} \tag{4.22}$$

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

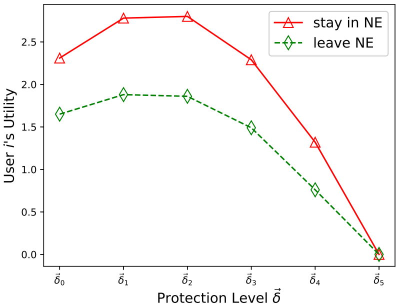

Figure 4.2. User utility vs. protection level.

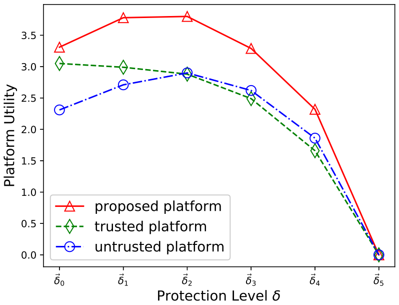

Figure 4.3. Platform utility vs. protection level.

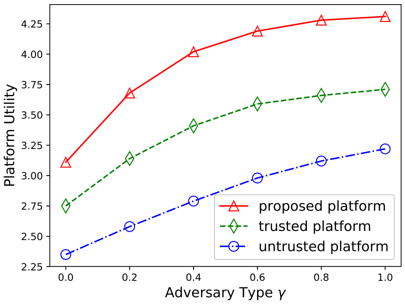

Figure 4.4. Optimal strategy of user under various.

Table 4.1. Extracted strategies

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

4.5 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) [73] to find the optimal strategies for the user and the platform.

4.5.1 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 [59], we extract four protection levels for income, age, and race. As shown in Table 4.1, $\vec{\delta}_1$, $\vec{\delta}_2$, $\vec{\delta}_3$, and $\vec{\delta}_4$, are the protection levels used by Retail platforms, Healthcare platforms, Government platforms, and Financial platforms, respectively. 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 $r_{income} = 3$, $r_{age} = 2$,$r_{race} = 1$. The value of these parameters are chosen to reveal plain interactions among three parties. We choose the best strategy from running the algorithm 100 times, where each run consists of 10,000 iterations.

4.5.2 Users Interaction

Fig. 4.2 shows the utility of user i when it performs different actions under different privacy protection levels. The x axis is the protection level, where $\vec{\delta_0} = \{0, 0, 0\}$ is the lowest protection level and $\vec{\delta_5} = \{1, 1, 1\}$ is the highest protection level. $\vec{\delta_1}$ to $\vec{\delta_4}$ are the increasing protection levels, as in Table 4.1. The solid red line in Fig. 4.2 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 $\vec{\delta_5}$ because the user cannot get any service quality and has no privacy loss. Fig. 4.2 also shows us that the utility of user i when it stays in NE is higher than that when it leaves NE. This demonstrates the existence of NE in the aggregative model and that users cannot get higher utility if they use non-Nash Equilibrium strategies.

4.5.3 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 to the adversary.

As shown in Fig. 4.3, 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 that of reputation loss increasing. This demonstrate the Nash Equilibrium existence of the two-layer three-party game because the platform cannot increase its utility by simply decreasing the privacy protection level. The platform needs to balance the tradeoff between payoff (from data collection and selling data) and reputation loss.

Fig. 4.3 and Fig. 4.4 compare the utility of three types of platforms with different protection levels and different adversary types, respectively. As shown in Fig. 4.3, the trusted platform has higher utility than the untrusted platform with protection level $\vec{\delta}_0$ to $\vec{\delta}_1$ 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 $\vec{\delta}_2$ to $\vec{\delta}_5$ 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. 4.3 and Fig. 4.4, we can see that the proposed platform in this chapter 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.

4.6 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 information 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 chapter 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, we next extend this work 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.

Chapter 5

GAME THEORY BASED PRIVACY PROTECTION FOR CONTEXT-AWARE SERVICES WITH LONG-TERM TIME SERIES DATA

In this chapter, we investigate the influence of long-term time series data and propose a three-party Stackelberg game model with consideration of active actions of platform and adversary.

5.1 Motivation

In context aware services, services providers require accurate personal information from users. However, users' information usually changes as time passes. For example, education information will change after graduation. The income information also changes over time with promotion or employment changes. That is the reason why some applications want users to update their personal information periodically. We can define this kind of data as long-term time-series data. When we consider the long-term time-series data, it is very different from fixed data because the data value is not only influenced by data granularity and data accuracy but also influenced by data freshness. The data value will decay as the time passed which means the fresher data has more value than the older data. There is a paradox on the time series data update frequency. If a user updates their personal data frequently, it will submit its data with high data freshness which means more accurate private information leaked to the platform. On the contrary, if a user refuses upload its information as its status changes, the service provider cannot provide service with good quality. Thus, it is very important to set a proper update frequency for each attribute in addition to the privacy protection level for each update.

Besides, in this kind scenarios, adversaries may not only eavesdrop or purchase users' submitted data, but also actively request users' information by sending surveys, advertisements or giving more reward for specific information. By sending customized mobile advertisements, adversaries can trace a user's location [74]. There is also a possibility to infer users' information from their reaction to different kinds of advertisements [75]. Therefore, to protect users long-term information, we need to consider the active-attack adversaries.

Many researchers propose different game models to protect users' privacy with consideration of time-series data [16, 17, 76–84]. However, all the work only focuses on short-term time series data like location and retail application usage. They investigate the privacy issues from the connection between locations. Different to the short-term time series data, the data freshness is the key privacy issue of long-term time series data. Besides, all the existing work only study the optimal defense strategy in two-party game model without considering the active-attack adversaries.

Therefore, a normal game theory based model cannot be used to formulate the long-term time-series data based context-aware services. To further investigate more realistic privacy protection issues, we extend our work to a model that considers the influence of long-term time-series data and an active-attack adversary.

To design a proper game model to adapt to the scenario with long-term time-series data, we need to consider the influence that does not exist in a scenario without time-series data such as data freshness, active-attack adversary, and decision making for each time phase. (i) *Data freshness.* The user may consider the data freshness for the long-term time-series data. The data freshness determines the privacy leakage as well as the service quality. (ii) *Active-attack adversary.* We should also care about the feedback for previous time phase from platform and adversary. The platform and adversary may have different needs of special data for the future according to the data category they have and the freshness of current data. To make the game model more practical, we need to consider the dynamic and heterogeneous reward for different attributes in the dataset from active-attack adversary. (iii) *Decision making for each time phase.* Because we consider different time phases in the game model, the user needs to make a decision for each time phase about the data submission privacy protection level. That means we need to find the NE strategy for all the parties for each time phase.

To solve the above challenges for the scenario with consideration of long-term series data and active-attack adversaries, we design a three-party Stackelberg game with the following efforts. (i) We divide time into time slots. In each time slot, to encourage user update fresh data in current time slot. The adversary can actively choose the data purchase price and the platform can actively choose the reward for submitted data. (ii) We use the theta decay formula to take the data freshness into consideration. (iii) The Stackelberg Game model is utilized to formulate the interaction among user, platform and adversary.

The main contributions are concluded as follows:

- We propose a three-party Stackelberg Game model especially for the long-term time series data scenario.
- In the proposed game model, we consider the influence of data freshness and active adversary and platform.
- We utilize machine learning techniques to find the quasi-Nash Equilibrium for the complex game model and provide insight privacy protection guidance for users.

The rest of the chapter is organized as follows. Section 5.2 reviews and compare existing works. We introduce the system model in Section 5.3. The optimal strategies for the user and platform is analyzed in Section 5.4. Section 5.5 presents the simulations to validate the theoretical analysis and we conclude this chapter in Section 5.6.

5.2 Related Work

Many researchers study how time-series data influence the privacy protection strategies in a game model. Cardenas et al. [76] propose a system to detect electricity theft by using data from smart meter which is continuously collect users' electricity usage data. To satisfy users privacy protection demands, they also propose an algorithm based on game theory to decide the proper uploading interval. Rottondi et al. [77] design a Game-theoretic Demand side management system which uses a decentralized approach for collaboratively scheduling the usage of domestic electrical appliances. It requires each user to communicate his/her own energy consumption patterns. By using game theory, the proposed system can decide the proper amount of added noise and the number of users for the purpose of privacy protection. Tefera and Yang [84] propose a framework to preserve location information privacy in location-based service applications by considering both historical access data and current location-based service access scenario, thus determining the probability that the visitors access is honest or not. To protect privacy of mobile users in location-based services, Shokri et al. [16] propose a Game-theoretic framework to alter users' real trace based on Stackelberg Bayesian game. With the purpose of preserving privacy in IoT-based transportation [82], Sfar et al. proposes a algorithm relying on a game theory model between two party (data holder and data requester) to reach a compromise between privacy protection and incentive motivation. Qu et al. [81] design a hybrid privacy-preserving scheme with consideration of both location and identity privacy to against continuous attacks from a dynamic adversary. They establish a game-based Markov decision process model to maximize data utility with high-level privacy protection. Alese et al. [80] propose a game model to protect users location privacy by considering continuous location information from n-users, thus each user can maximize its location privacy at minimum cost by strategically choosing series of actions in the game. In [17], Sfar et al. propose a Markov game privacy preserving model for retail applications. The proposed model considers the privacy leakage issues from diverse components in retail applications and provide detailed states, actions, strategies and transitions for data holder to balance the tradeoff between privacy protection and incentive motivation of data requester. In [79], Xiao et al. design a Stackelberg game model to prevent Mobile crowdsensing server from fake sensing users. The Nash Equilibrium (NE) ensures the smartphone users provide real sensing task with consideration of sensing cost and privacy leakage. The author utilize deep Q-network to find the NE of the proposed Stackelberg game model. Pawlick and Zhu [78] model the conflict between machine learning and data obfuscation with

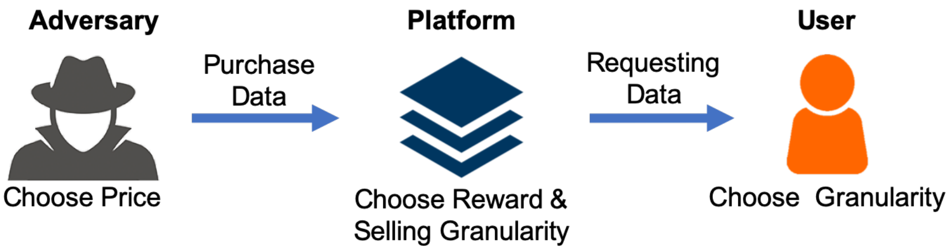

Figure 5.1. Structure of thee-party Stackelberg game model.

a Stackelberg Game. They also analyze the interaction between N+1 users by considering the influence of each user's perturbation. Wang and Zhang [83] study the context privacy problem in the scenario has dynamic context and malicious adversaries with capabilities of attacking adjustment. They use a competitive Markov decision process to model the interactive competition between users and adversaries in a long-term defense. The authors also propose an minimax learning algorithm to calculate the optimal strategies for the users and prove the propose algorithm can converges to the unique Nash equilibrium point quickly.

However, all the existing work only consider the short-term time series data with passive third-party (platform or adversary). In this work, we will focus on the scenario of long-term time-series data with active platform and adversary.

5.3 System Model

In this section, we formulate the interactions between all the three parties with the consideration of data freshness in the scenario of active-attack adversary.

5.3.1 Framework

In this game model, as shown in Fig. 5.1 we assume there are three parties: user, platform, adversary. We consider the adversary can actively change the purchase price of different attributes. That means the adversary can increase the purchase price for needed attributes.

The platform can also actively request information and adjust the reward for different attributes for each time slot. Therefore, the user has to choose a proper data submit/update strategy to maximize its utility with consideration of data freshness. Because user's data d will change as time goes by, the data d may not "fresh" if the user did not update it after it changed. We divide time into time slots. The freshness of attributes is decided by how many time slots has no attributes update, which is defined by Δ.

In this game, we can see that the adversary will choose its strategy first and then the platform. The user will give response to the adversary and platform's strategy. Thus, the adversary is the leader in this game, and the platform is the follower of the adversary and leader of the user. The user is the follower of the platform. We can use Stackelberg Game model to formulate the interaction among user, platform and adversary.

In the following sections, we introduce the model of each party according to the order of play in the Stackelberg Game model.

5.3.2 Adversary Model

As we discussed in former chapters, an adversary can purchase data from platform. According to Cobb-Douglas production function [47], we can derive the expected value of the new purchased data for the adversary:

$$V_a = \gamma_a (sg)^{\alpha_a}. \tag{5.1}$$

where s is the platforms selling strategy decides the selling data granularity sg, γ_a is its value productivity of the adversary, α_a is adversary's output elasticities of purchased data granularity of sg.

Due to time passing, the data my not fresh anymore. If the user does not update in one slot, the adversary can still generate value by using previous collected data as

$$\bar{V_a} = \gamma_a (\bar{sg})^{\alpha_a} (\Delta + 1)^{\beta_a}, \tag{5.2}$$

where $\bar{s}$ denotes the last purchased data granularity and and β_a is adversary's output elasticities of Δ.

As time slot increases, the value derived decreases. Thus, the adversary would like to actively change the purchase price to encourage the platform to request fresh data from users. The price for one unit data granularity is p, thus purchase price is ps, where s is the granularity of purchased data from the platform. Then, we can derive the utility function of the adversary as follows.

$$U_a = V_a - psg. \tag{5.3}$$

The adversary will adjust the price p to maximize its utility. We formalize the optimization problem as follows.

$$\begin{gathered} \max_{p} U_a \\ subject\ to\ U_a > \bar{V_a}. \end{gathered} \tag{5.4}$$

.

5.3.3 Platform Model

After the adversary decides the price for the data, the platform may also change its selling strategy as well as the reward provided to user. The unit granularity reward provided to the user can be defined as r. The total reward provided to the user will be rg, where g is the granularity of user submitted granularity. Meanwhile, the platform will provide service to users according to the submitted data granularity and freshness. The service quality is determined not only by data granularity, but also by data freshness. The data freshness will decay as the time goes by. We can define the service quality lose as option theta decay formulation. Thus, the service quality can be defined as

$$Q = log(g + 1). \tag{5.5}$$

With increasing time slot, the service quality will decay if there is no new data collected. The service quality decay can be formulated by using time decay function as

$$\bar{Q} = log(\bar{g} + 1) - \theta\Delta. \tag{5.6}$$

The function of θ is decided by the type of the data.

The platform can also get value from the data as well. Similarly, according to Cobb-Douglas production function [47], the platform can get value from data as

$$V_p = \gamma_p g^{\alpha_p}, \tag{5.7}$$

where γ_p is its value productivity of the platform, α_p is platform's output elasticities of g.

If there is no new data collected from the user, the generated value by using previous data is

$$\bar{V_p} = \gamma_p \bar{g}^{\alpha_p}(\Delta + 1)^{\beta_p}, \tag{5.8}$$

where β_p is platform's output elasticities of Δ.

Meanwhile, the platform has reputation loss for selling data to adversary. If c is the unit reputation loss for selling data, the total reputation loss for selling data is cs. The platform can also pay unit reward r for submitted data to the user with a total reward rg. Besides, the platform has a reputation loss because selling user's data. We use a exponential function s^{ϵ} to formulate the risk of reputation loss.

Thus, the platform has a utility

$$U_p = V_p - rg + psg - s^{\epsilon} \tag{5.9}$$

As a rational platform, it will choose a proper unit reward r and selling strategy s to maximize its utility as the following optimization problem.

$$\begin{aligned} &\max_{s,r} U_p \\ &subject\ to\ U_p > \bar{V}_p,\ and\ s \leq g. \end{aligned} \tag{5.10}$$

.

5.3.4 User Model

The user can get service with quality Q and reward rg from the platform. However, the user will suffer the joint privacy leakage from both platform and adversary. We define the unit risk cost of privacy leakage is l, the joint privacy leakage is $l(s+g)$.

Thus, the utility of the user is

$$U_u = Q + rg - l(s+g) \tag{5.11}$$

Thus, the user has to choose a proper submission granularity to maximize its utility as

$$\begin{aligned} &\max_{g} U_u \\ &subject\ to\ U_u > \bar{Q} \end{aligned} \tag{5.12}$$

5.4 Nash Equilibrium

To find the best strategy for the user, we need to analyze the interaction among the three parties and find the stable status which is the Nash Equilibrium. Accordingly, the Nash equilibrium can be defined as

Definition 3. *A strategy profile $S^* = (g^*, r^*, s^*, p^*)$ is called Nash Equilibrium if the following properties simultaneously hold:*

$$U_u(g^*, r^*, s^*, p^*) \geq U_u(g, r^*, s^*, p^*), \forall g \in S_u;$$

$$U_p(g^*, r^*, s^*, p^*) \geq U_p(g^*, r, s, p^*), \forall (r, s) \in S_p;$$

$$U_a(g^*, r^*, s^*, p^*) \geq U_a(g^*, r^*, s^*, p), \forall p \in S_a;$$

subject to $U_p^* > \bar{V}_p$, $U_a^* > \bar{V}_a$, *and* $U_u > \bar{Q}$, *where* S_u, S_p, S_a *are the strategy space of the user, platform, and adversary.*

The Stackelberg Game model can be solved by backward induction. However, for most parameters of α_a, α_p, there is no set equation for solving directly, instead we can use machine learning techniques to find the quasi-NE.

5.5 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) [73] to find the optimal strategies for the user and the platform like previous chapters.

5.5.1 Simulation Setting

To get a obvious result, we set the productivity of the adversary $\gamma_a = 4$, the adversary's output elasticities of purchased data granularity $\alpha_a = 0.2$. We set the productivity of the platform $\gamma_p = 6$, the platform's output elasticities of collected data granularity $\alpha_p = 0.3$, and the parameter ϵ for platform's risk function is 2. In the simulation, to analyze the influence of each party's strategy, we set different value of strategies as $p_1 = 6, p_2 = 4, p_3 = 2$ and $r_1 = 3, r_2 = 2, r_3 = 1$. The values for these parameter is choose to show a legible interaction among three parties and the existence of NE. For some other values, the NE may not exist. We choose the best strategy from running the algorithm 100 times, where each run consists of 10,000 iterations.

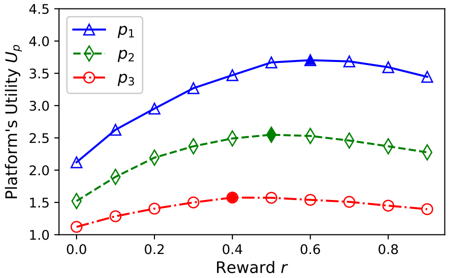

Figure 5.2. Platform Utility vs. Reward.

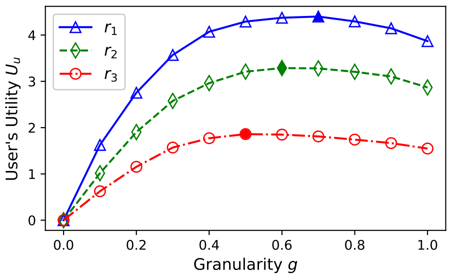

Figure 5.3. User Utility vs. Granularity.

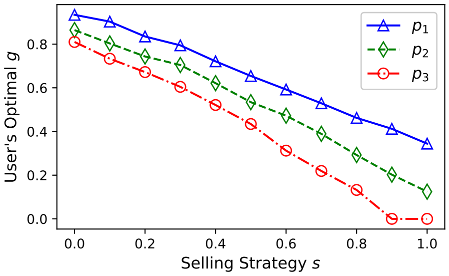

Figure 5.4. User Optimal Strategy.

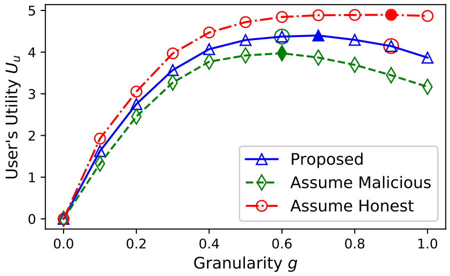

Figure 5.5. User utility Comparison.

5.5.2 Nash Equilibrium

From the simulation result in Fig. 5.2 and Fig. 5.3, we can see that there is a strategy that can lead to maximum utility for both the user and platform. This can demonstrate that the Nash Equilibrium exists in the proposed game model for the chosen parameters. For a specific domain, the user, platform, and the adversary can run simulation under different parameters to analyze the parameter range that can make the existence of NE.

As shown in Fig. 5.2, the platform's utility increases as the reward increases before the maximum point (solid dot) and decreases as the reward increases after the maximum point. The utility increases as the reward increases because the increased reward will encourage the user to increase the granularity of submitted data. But, if the platform gives too much reward to the user, the payoff from the collected data cannot make up the paid reward. That is the reason why the utility decreases as the reward increases after the maximum point. As shown in Fig. 5.3, the user' utility increases as the granularity increases before the maximum point (solid point) and then decreases as the granularity increases. The utility increases as the granularity increases because the increased granularity can bring more service quality and reward. However, a higher granularity can also cause more privacy leakage risk. That is the reason why the utility decreases with the granularity after the maximum point.

From both Fig. 5.2 and Fig. 5.3, we can see that for a given strategy of the other two parties, the platform and the user can find a optimal strategy to get maximum utility. This validates that the NE exist for certain value of parameter.

5.5.3 Interactions Among Three Parties

Fig. 5.2 reveals the interaction between the adversary and the platform. We can states that a higher unit price p will encourage the platform increase the reward given to the user. This is because a higher data granularity will increase the payment from the adversary. Thus, to pursue more payment, the platform would like to increase the reward to get data with high granularity.

Fig. 5.3 and Fig. 5.4 reveals the interaction between the platform and the user. As shown in Fig. 5.3, the higher reward will increase the utility of the user. Thus, the user is willing to increase the data granularity to get higher service quality and higher reward. As shown in Fig. 5.4, the optimal strategy of the user decreases as the platform increases its selling strategy s. This is because the increased selling strategy of the platform will increase the risk of privacy leakage. The user has to decrease the data granularity to avoid to much privacy leakage.

Fig. 5.4 also reveals the indirect interaction between the adversary and the user. We can tell that, a higher purchase unit price p of the adversary leads to a higher data granularity. The reason behind the result is a high unit price encourages the platform increase the reward (as shown in Fig. 5.2), thus encouraging the user to increase the data granularity (as shown in Fig. 5.3).

5.5.4 Significance of Three-Party Game Model

To validate the significance of the proposed three party Stackelberg game model, we calculate the utility of the user in two different scenarios: (i) the user assumes the platform is malicious and selling all the collected data. (ii) the user assume the platform is honest and will not sell any collected data. From Fig. 5.5, we can see that the user will choose a higher granularity (as the solid red dot) for submitted data when it assumes the platform is honest. When the user assumes the platform is malicious, it will choose to decrease the granularity (as the solid green diamond) for the submitted data to avoid risk of privacy leakage. However, in the real world, most of the platforms are the type proposed in the game model. Thus, neither the optimal strategy (left green circle) when assume the platform is malicious nor the strategy (right red circle) when assume the platform is honest can make the user get maximum utility. Therefore, the proposed three-party Stackelberg game model is significant to guide the user to choose the optimal strategy to protect its privacy.

5.6 Conclusion

In this chapter, we consider the long-term time-series data and formulate the influence of time decay of data. In this scenario, we assume both the adversary and platform have active actions to request data. We design a three-party Stackelberg game model to formulate the interactions among user, platform, and adversary. The machine learning based simulation validate the proposed game model and the existence of NE. By using the proposed framework, users can have guidance for whether to submit data for the current time slot and what the proper granularity should be for the submitted data.

Chapter 6

FUTURE WORK

In this dissertation, we propose three-party game models for three different general scenarios. These works provide theoretical guidance for users to protect their privacy while enjoying context-aware services. To apply the derived theoretical result into specific practical scenarios, we need to adjust the general structure and formulations to fit in the target scenario. Furthermore, it is very hard for users to calculate the NE and then choose a proper action based on the theoretical value. Thus, an AI (Artificial Intelligence) based Middleware Powered by Game Theory is necessary for users to automatically apply the theoretical result derived from three-party game model.

Therefore, in the future, we would like to design game model based applications for specific domains such as Smart Healthcare 85–95, Industry 4.0 [96–106], Big Data [107–119], Smart City [120–125], Cloud Computing [126–150], Croudsourcing [151–169], Cognitive Radio Network [170–174], Mobile Network [175–193], Wireless Sensor Network [194–235], Vehicular Network [236–252], Social Network [253–262], 5G Network [263–270], Smart City [271–281], BlockChain [282], etc. We also want to build AI based Middlewares to detect the strategy of platforms (from coupon, survey, advertisement, etc) and adversaries (from advertisement, fraud messages, etc), thus choosing the proper strategy for users automatically based on the result derived from game model.

6.1 Game Theory Based Application for Specific Domain

With the development of CPS (Cyber-Physical System) [283–288] and IoT (Internet of Things) [289, 290], context-aware services has been split to diverse specific scenarios such as Healthcare, Smart City (includes smart grid, water and waste management, traffic and transportation management, Cognitive Radio Network, Mobile Network, Wireless Sensor Network, etc.), Industry 4.0, Retail etc.

As we can see, different scenarios have different type of context and different valuation functions. Thus, to apply the general framework work, we have to go deeper to investigate the risk and valuation functions. Because of the importance of people's health information and the popularity of health related wearable devices, we will investigate the specific needs for healthcare domain in our future work. We will study the common context types collected from users' wearable devices, the services provided from different kind of platform, and possible attacks from adversaries. Besides, we will also investigate the valuation functions for the service quality, risk estimation, and data valuation as well as the proper parameters in these formulations especially for healthcare scenario. Furthermore, according to the interaction in healthcare service, we may modify one of the proposed general three party game model or combine them together. We intend to derive a result which can be easily embedded to the healthcare privacy protection system.

6.2 Privacy Protection Middleware Powered by Game Theory

In this dissertation, we derive the theoretical result of the propose strategy of users. However, the theoretical result like "set granularity to 0.6" is meaningless for the user. Thus, there is a gap between the theoretical result derived from this dissertation and real application. To fill this gap, we would like to build a middleware for context-aware services by utilizing DL techniques [291]. From the coupon, survey and advertisement sent from the platform and adversary, we can use DL based keyword detection and image detection to detect the strategy of the platform and the adversary. Based on the proposed game model, the algorithm can calculate the proper strategy of the user. We can use fuzzy set to map the numerical strategy to proper meaningful strategy. As a privacy protector installed in user's devices, the middleware will not leak precious information to any platform. Therefore, for any type of data, it will submit data with proper granularity to data requester with respect to users' utility.

Chapter 7

CONCLUSION

In this dissertation, we study the problems of privacy protection for context-aware services based on game theory. We build three party game models for three different general scenarios and provide privacy guidance for users accordingly.

Firstly, this dissertation provides realistic optimal strategies for both the user and the platform. We propose a *three-party game model* that captures the interactions between any two of the parties: user, platform and adversary. Our solution determines an *optimal fine-grained strategy* for the user and platform, so that the user can choose an optimal data granularity to balance service quality and privacy leakage and that the platform can choose the optimal reselling strategy to balance profit and reputation loss. Our model also accounts for the correlations between *multiple data attributes* provided by a user.

Secondly, this dissertation 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. 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.

Thirdly, this dissertation consider the long-term time-series data and formulates the influence of time decay of data as well as the active action of the platform and adversary. We design a three-party Stackelberg game model to formulate the interactions among user, platform, and adversary. By using the proposed framework, users can have guidance for whether to submit data for current time and what the proper granularity for data submitting.

All the proposed game models and solutions are thoroughly discussed and validated through simulation. We also discuss some topics for future study in our dissertation. Generally, our dissertation provides a body of solutions for the purpose of privacy protection in diverse context-aware services base on game theory. These solutions could comprehensively provide guidance for user to choose a proper strategy to preserve privacy while enjoy the context-aware services. We believe the work in this dissertation will serve as a core of privacy protection algorithms. Finally, this dissertation will also inspire subsequent research towards the publication of privacy protection in specific applications.

ACKNOWLEDGMENT

This dissertation is partly supported by the National Science Foundation (NSF) under grant NOs. 1252292, 1741277, 1704287, 1829674, and 1912753.

REFERENCES

[1] X. Wang, Y. Yang, C. Tang, Y. Zeng, and J. He, "Droidcontext: Identifying malicious mobile privacy leak using context," in *Proc. IEEE Trustcom/BigDataSE/ISPA*, Aug. 2016.

[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] "The equifax data breach." `https://www.ftc.gov/equifax-data-breach`.

[4] J. Wang, Z. Cai, Y. Li, D. Yang, J. Li, and H. Gao, "Protecting query privacy with differentially private k-anonymity in location-based services," *Personal and Ubiquitous Computing*, pp. 1–17, 2018.

[5] A. Machanavajjhala, M. Venkitasubramaniam, D. Kifer, and J. Gehrke, "l-diversity: Privacy beyond k-anonymity," in *ICDE 2016*, vol. 00, p. 24, 04 2006.

[6] N. Li, T. Li, and S. Venkatasubramanian, "t-closeness: Privacy beyond k-anonymity and l-diversity," in *ICDE 2007*, pp. 106–115, April 2007.

[7] A. Pastore and M. Gastpar, "Locally differentially-private distribution estimation," in *IEEE ISIT 2016*, pp. 2694–2698, July 2016.

[8] N. Li, W. Qardaji, and D. Su, "On sampling, anonymization, and differential privacy or, k-anonymization meets differential privacy," in *Proceedings of the 7th ACM Symposium on Information, Computer and Communications Security*, ASIACCS '12, (New York, NY, USA), pp. 32–33, ACM, 2012.

[9] A. Narayanan and V. Shmatikov, "Robust de-anonymization of large sparse datasets," in *2008 IEEE Symposium on Security and Privacy (sp 2008)*, pp. 111–125, May 2008.

[10] A. Narayanan and V. Shmatikov, “How to break anonymity of the netflix prize dataset,” *CoRR*, vol. abs/cs/0610105, 2006.

[11] A. McGregor, I. Mironov, T. Pitassi, O. Reingold, K. Talwar, and S. Vadhan, “The limits of two-party differential privacy,” in *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, pp. 81–90, Oct 2010.

[12] R. Dewri, “Local differential perturbations: Location privacy under approximate knowledge attackers,” *IEEE Transactions on Mobile Computing*, vol. 12, pp. 2360–2372, Dec 2013.

[13] A. K. Chorppath and T. Alpcan, “Trading privacy with incentives in mobile commerce: A game theoretic approach,” *Pervasive and Mobile Computing*, vol. 9, pp. 598–612, Aug. 2013.

[14] R. Shokri, G. Theodorakopoulos, C. Troncoso, J.-P. Hubaux, and J.-Y. Le Boudec, “Protecting location privacy: Optimal strategy against localization attacks,” in *CCS 2012*, Oct. 2012.

[15] C. Rottondi, A. Barbato, L. Chen, and G. Verticale, “Enabling privacy in a distributed game-theoretical scheduling system for domestic appliances,” *IEEE TSG*, vol. 8, pp. 1220–1230, May 2017.

[16] R. Shokri, G. Theodorakopoulos, and C. Troncoso, “Privacy games along location traces: A game-theoretic framework for optimizing location privacy,” *ACM TPS*, vol. 19, pp. 11–31, Feb. 2017.

[17] A. Sfar, E. Natalizio, Y. Challal, and Z. Chtourou, “A game-theoretic approach: A markov game privacy preserving model in retail applications,” in *IEEE MowNet 2017*, May 2017.

[18] K. Li, L. Tian, W. Li, G. Luo, and Z. Cai, "Incorporating social interaction into three-party game towards privacy protection in iot," *Computer Networks*, vol. 150, pp. 90–101, 2019.

[19] X. Wu, W. Dou, and Q. Ni, "Game theory based privacy preserving analysis in correlated data publication," in *Proc. ACM ACSW*, Feb. 2017.

[20] X. Liu, K. Liu, L. Guo, X. Li, and Y. Fang, "A game-theoretic approach for achieving k-anonymity in location based services," in *IEEE INFOCOM 2013*, Apr. 2013.

[21] J. Freudiger, M. H. Manshaei, J.-P. Hubaux, and D. C. Parkes, "Non-cooperative location privacy," *TDSC*, vol. 10, pp. 84–98, Mar. 2013.

[22] R. Ma, J. Xiong, M. Lin, Z. Yao, H. Lin, and A. Ye, "Privacy protection-oriented mobile crowdsensing analysis based on game theory," in *2017 IEEE Trustcom/BigDataSE/ICESS*, pp. 990–995, Aug 2017.

[23] B. Ying and A. Nayak, "Location privacy-protection based on p-destination in mobile social networks: A game theory analysis," in *2017 IEEE Conference on Dependable and Secure Computing*, pp. 243–250, Aug 2017.

[24] L. Xu, C. Jiang, Y. Qian, J. Li, Y. Zhao, and Y. Ren, "Privacy-accuracy trade-off in differentially-private distributed classification: A game theoretical approach," *IEEE Transactions on Big Data*, pp. 1–1, 2017.

[25] W. Li, C. Hu, T. Song, J. Yu, X. Xing, and Z. Cai, "Preserving data privacy in context-aware applications through hierarchical game," in *accepted by IEEE Symposium on Privacy-Aware Computing*, (Washington DC, USA), Sep. 2018.

[26] R. Karimi Adl, M. Askari, K. Barker, and R. Safavi-Naini, *Privacy consensus in anonymization systems via game theory*, pp. 74–89. Springer Berlin Heidelberg, 2012.

[27] S. Wang, L. Li, W. Sun, J. Guo, R. Bie, and K. Lin, "Context sensing system analysis for privacy preservation based on game theory," *Sensors*, vol. 17, p. 339, Feb. 2017.

[28] S. Wang, J. Huang, L. Li, L. Ma, and X. Cheng, "Quantum game analysis of privacy-leakage for application ecosystems," in *Proc. ACM MobiHoc*, Jul. 2017.

[29] I. Vakilinia, D. K. Tosh, and S. Sengupta, "3-way game model for privacy-preserving cybersecurity information exchange framework," in *MILCOM 2017*, pp. 829–834, Oct 2017.

[30] C. Stergiou, K. E. Psannis, B.-G. Kim, and B. Gupta, "Secure integration of iot and cloud computing," *Future Generation Computer Systems*, vol. 78, pp. 964 – 975, 2018.

[31] B. Farahani, F. Firouzi, V. Chang, M. Badaroglu, N. Constant, and K. Mankodiya, "Towards fog-driven iot ehealth: Promises and challenges of iot in medicine and healthcare," *Future Generation Computer Systems*, vol. 78, pp. 659 – 676, 2018.

[32] H. Li, K. Ota, and M. Dong, "Learning iot in edge: Deep learning for the internet of things with edge computing," *IEEE Network*, vol. 32, pp. 96–101, Jan 2018.

[33] M. Mohammadi, A. Al-Fuqaha, S. Sorour, and M. Guizani, "Deep learning for iot big data and streaming analytics: A survey," *IEEE Communications Surveys Tutorials*, vol. 20, pp. 2923–2960, Fourthquarter 2018.

[34] M. Elhoseny, G. Ramrez-Gonzlez, O. M. Abu-Elnasr, S. A. Shawkat, A. N, and A. Farouk, "Secure medical data transmission model for iot-based healthcare systems," *IEEE Access*, vol. 6, pp. 20596–20608, 2018.

[35] O. Novo, "Blockchain meets iot: An architecture for scalable access management in iot," *IEEE Internet of Things Journal*, vol. 5, pp. 1184–1195, April 2018.

[36] A. P. Plageras, K. E. Psannis, C. Stergiou, H. Wang, and B. Gupta, "Efficient iot-based sensor big data collectionprocessing and analysis in smart buildings," *Future Generation Computer Systems*, vol. 82, pp. 349 – 357, 2018.

[37] M. Ammar, G. Russello, and B. Crispo, “Internet of things: A survey on the security of iot frameworks,” *Journal of Information Security and Applications*, vol. 38, pp. 8 – 27, 2018.

[38] K. Muhammad, R. Hamza, J. Ahmad, J. Lloret, H. Wang, and S. W. Baik, “Secure surveillance framework for iot systems using probabilistic image encryption,” *IEEE Transactions on Industrial Informatics*, vol. 14, pp. 3679–3689, Aug 2018.

[39] W. Wang and Q. Zhang, “A stochastic game for privacy preserving context sensing on mobile phone,” in *IEEE INFOCOM 2014 - IEEE Conference on Computer Communications*, pp. 2328–2336, April 2014.

[40] M. Halkidi and I. Koutsopoulos, *A game theoretic framework for data privacy preservation in recommender systems*, pp. 629–644. Springer Berlin Heidelberg, 2011.

[41] V. Kumari and S. Chakravarthy, “Cooperative privacy game: a novel strategy for preserving privacy in data publishing,” *Human-centric Computing and Information Sciences*, vol. 6, p. 12, Jul. 2016.

[42] Z. Wang, S. C. S. Cheung, and Y. Luo, “Information-theoretic secure multi-party computation with collusion deterrence,” *IEEE Transactions on Information Forensics and Security*, vol. 12, pp. 980–995, April 2017.

[43] Y. Wang, Z. Cai, G. Yin, Y. Gao, X. Tong, and Q. Han, “A game theory-based trust measurement model for social networks,” *Computational Social Networks*, vol. 3, no. 1, p. 2, 2016.

[44] L. Xu, C. Jiang, Y. Chen, Y. Ren, and K. J. R. Liu, “Privacy or utility in data collection? a contract theoretic approach,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, pp. 1256–1269, Oct 2015.

[45] T. Zhu, P. Xiong, G. Li, and W. Zhou, "Correlated differential privacy: Hiding information in non-iid data set," *IEEE Transactions on Information Forensics and Security*, vol. 10, pp. 229–242, Feb 2015.

[46] L. Zhang, Z. Cai, and X. Wang, "Fakemask: A novel privacy preserving approach for smartphones," *IEEE Transactions on Network and Service Management*, vol. 13, pp. 335–348, June 2016.

[47] W. Meeusen and J. van Den Broeck, "Efficiency estimation from cobb-douglas production functions with composed error," *International Economic Review*, vol. 18, pp. 435–444, Jun. 1977.

[48] L. Tian, J. Li, W. Li, B. Ramesh, and Z. Cai, "Optimal contract-based mechanisms for online data trading markets," *IEEE Internet of Things Journal*, 2019.

[49] Z. Cai and Z. He, "Trading private range counting over big iot data," in *The 39th IEEE International Conference on Distributed Computing Systems (ICDCS 2019)*, IEEE, 2019.

[50] C. Fershtman and M. I. Kamien, "Dynamic duopolistic competition with sticky prices," *Econometrica*, vol. 55, pp. 1151–1164, Sep. 1987.

[51] P. Hu, H. Li, H. Fu, D. Cansever, and P. Mohapatra, "Dynamic defense strategy against advanced persistent threat with insiders," in *Proc. IEEE INFOCOM*, Apr. 2015.

[52] X. Zheng, Z. Cai, and Y. Li, "Data linkage in smart internet of things systems: A consideration from a privacy perspective," *IEEE Communications Magazine*, vol. 56, no. 9, pp. 55–61, 2018.

[53] M. Osborne, *Introduction to Game Theory: International Edition.* Oxford University Press, 2009.

[54] A. H. Mohsenian-Rad, V. W. S. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia, “Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid,” *IEEE Transactions on Smart Grid*, vol. 1, pp. 320–331, Dec. 2010.

[55] R. A. Horn and C. R. Johnson, *Matrix Analysis.* New York, NY, USA: Cambridge University Press, 2nd ed., 2012.

[56] R. Eberhart and J. Kennedy, “A new optimizer using particle swarm theory,” in *Micro Machine and Human Science, 1995. MHS ’95., Proceedings of the Sixth International Symposium on*, pp. 39–43, Oct 1995.

[57] `https://github.com/chnhuangyan/ThreePartyGameSimulationPSO`.

[58] IBM, “Ibm multi-national consumer privacy survey,” Otc. 1999.

[59] B. Filkins, “Sensitive data at risk: The sans 2017 data protection survey,” Sep. 2017.

[60] X. Zheng, J. Li, H. Gao, and Z. Cai, “Capacity of wireless networks with multiple types of multicast sessions,” in *The 15th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MOBIHOC 2014)*, pp. 135–144, ACM, 2014.

[61] Z. He, Z. Cai, S. Cheng, and X. Wang, “Approximate aggregation for tracking quantiles in wireless sensor networks,” in *The 8th Annual International Conference on Combinatorial Optimization and Applications (COCOA2014)*, pp. 161–172, Springer, 2014.

[62] Z. Cai, R. Goebel, and G. Lin, “Size-constrained tree partitioning: approximating the multicast k-tree routing problem,” *Theoretical Computer Science*, vol. 412, no. 3, pp. 240–245, 2011.

[63] Z. Cai, Z.-Z. Chen, G. Lin, and L. Wang, “An improved approximation algorithm for the capacitated multicast tree routing problem,” *The 2nd Annual International Conference on Combinatorial Optimization and Applications (COCOA2008)*, vol. 5165, pp. 286–295, 2008.

[64] M. K. Jensen, “Aggregative games and best-reply potentials,” *Economic Theory*, vol. 43, pp. 45–66, Apr 2010.

[65] U. Erlingsson, V. Pihur, and A. Korolova, “Rappor: Randomized aggregatable privacy-preserving ordinal response,” in *CCS 2014*, pp. 1054–1067, ACM, 2014.

[66] P. Kairouz, S. Oh, and P. Viswanath, “Extremal mechanisms for local differential privacy,” in *Advances in Neural Information Processing Systems 27*, pp. 2879–2887, Curran Associates, Inc., 2014.

[67] G. Cormode, S. Jha, T. Kulkarni, N. Li, D. Srivastava, and T. Wang, “Privacy at scale: Local differential privacy in practice,” in *SIGMOD 2018*, SIGMOD ’18, pp. 1655–1658, ACM, 2018.

[68] A. G. Thakurta, A. H. Vyrros, U. S. Vaishampayan, G. Kapoor, J. Freudinger, V. V. Prakash, A. Legendre, and S. Duplinsky, “Emoji frequency detection and deep link frequency.”

[69] Z. Qin, Y. Yang, T. Yu, I. Khalil, X. Xiao, and K. Ren, “Heavy hitter estimation over set-valued data with local differential privacy,” in *CCS 2016*, CCS ’16, pp. 192–203, ACM, 2016.

[70] R. Dewri, “Local differential perturbations: Location privacy under approximate knowledge attackers,” *IEEE TMC*, vol. 12, pp. 2360–2372, Dec 2013.

[71] U. Jugel, Z. Jerzak, G. Hackenbroich, and V. Markl, “M4: A visualization-oriented time series data aggregation,” *Proc. VLDB Endow.*, vol. 7, pp. 797–808, June 2014.

[72] M. Miltiadis, “The theory of incentives: The principalagent model.,” *The Economic Journal*, vol. 113, no. 488, pp. F394–F395, 2001.

[73] M. Clerc and J. Kennedy, “The particle swarm - explosion, stability, and convergence in a multidimensional complex space,” *IEEE TEC*, vol. 6, pp. 58–73, Feb 2002.

[74] "It takes just $1,000 to track someone's location with mobile ads." `https://www.wired.com/story/track-location-with-mobile-ads-1000-dollars-study/`.

[75] M. A. Bashir, S. Arshad, W. Robertson, and C. Wilson, "Tracing information flows between ad exchanges using retargeted ads," in *25th USENIX Security Symposium (USENIX Security 16)*, (Austin, TX), pp. 481–496, 2016.

[76] A. A. Crdenas, S. Amin, G. Schwartz, R. Dong, and S. Sastry, "A game theory model for electricity theft detection and privacy-aware control in ami systems," in *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 1830–1837, Oct 2012.

[77] C. Rottondi, A. Barbato, L. Chen, and G. Verticale, "Enabling privacy in a distributed game-theoretical scheduling system for domestic appliances," *IEEE Transactions on Smart Grid*, vol. 8, pp. 1220–1230, May 2017.

[78] J. Pawlick and Q. Zhu, "A stackelberg game perspective on the conflict between machine learning and data obfuscation," in *2016 IEEE International Workshop on Information Forensics and Security (WIFS)*, pp. 1–6, Dec 2016.

[79] L. Xiao, Y. Li, G. Han, H. Dai, and H. V. Poor, "A secure mobile crowdsensing game with deep reinforcement learning," *IEEE Transactions on Information Forensics and Security*, vol. 13, pp. 35–47, Jan 2018.

[80] B. K. Alese, A. F. Thompson, and P. Y. Oni, "A location privacy system in mobile network using game theory," in *2017 International Conference On Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*, pp. 1–5, June 2017.

[81] Y. Qu, S. Yu, L. Gao, W. Zhou, and S. Peng, "A hybrid privacy protection scheme in cyber-physical social networks," *IEEE Transactions on Computational Social Systems*, vol. 5, pp. 773–784, Sep. 2018.

[82] A. Riahi Sfar, Y. Challal, P. Moyal, and E. Natalizio, “A game theoretic approach for privacy preserving model in iot-based transportation,” *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–10, 2019.

[83] W. Wang and Q. Zhang, “Privacy preservation for context sensing on smartphone,” *IEEE/ACM Trans. Netw.*, vol. 24, pp. 3235–3247, Dec. 2016.

[84] M. K. Tefera and X. Yang, “A game-theoretic framework to preserve location information privacy in location-based service applications,” *Sensors*, vol. 19, no. 7, 2019.

[85] G. Muhammad, S. M. M. Rahman, A. Alelaiwi, and A. Alamri, “Smart health solution integrating iot and cloud: A case study of voice pathology monitoring,” *IEEE Communications Magazine*, vol. 55, pp. 69–73, January 2017.

[86] Y. Zhang, D. Zheng, and R. H. Deng, “Security and privacy in smart health: Efficient policy-hiding attribute-based access control,” *IEEE Internet of Things Journal*, vol. 5, pp. 2130–2145, June 2018.

[87] M. I. Pramanik, R. Y. Lau, H. Demirkan, and M. A. K. Azad, “Smart health: Big data enabled health paradigm within smart cities,” *Expert Systems with Applications*, vol. 87, pp. 370 – 383, 2017.

[88] Q. Jiang, M. K. Khan, X. Lu, J. Ma, and D. He, “A privacy preserving three-factor authentication protocol for e-health clouds,” *The Journal of Supercomputing*, vol. 72, pp. 3826–3849, Oct 2016.

[89] D. Kotz, C. A. Gunter, S. Kumar, and J. P. Weiner, “Privacy and security in mobile health: A research agenda,” *Computer*, vol. 49, pp. 22–30, June 2016.

[90] C. Lin, Z. Song, H. Song, Y. Zhou, Y. Wang, and G. Wu, “Differential privacy preserving in big data analytics for connected health,” *Journal of Medical Systems*, vol. 40, p. 97, Feb 2016.

[91] P. Sundaravadivel, E. Kougianos, S. P. Mohanty, and M. K. Ganapathiraju, “Everything you wanted to know about smart health care: Evaluating the different technologies and components of the internet of things for better health,” *IEEE Consumer Electronics Magazine*, vol. 7, pp. 18–28, Jan 2018.

[92] E. C. Nelson, T. Verhagen, and M. L. Noordzij, “Health empowerment through activity trackers: An empirical smart wristband study,” *Computers in Human Behavior*, vol. 62, pp. 364 – 374, 2016.

[93] D. He, R. Ye, S. Chan, M. Guizani, and Y. Xu, “Privacy in the internet of things for smart healthcare,” *IEEE Communications Magazine*, vol. 56, pp. 38–44, April 2018.

[94] L. Liu, E. Stroulia, I. Nikolaidis, A. Miguel-Cruz, and A. R. Rincon, “Smart homes and home health monitoring technologies for older adults: A systematic review,” *International Journal of Medical Informatics*, vol. 91, pp. 44 – 59, 2016.

[95] V. Sucasas, G. Mantas, A. Radwan, and J. Rodriguez, “An oauth2-based protocol with strong user privacy preservation for smart city mobile e-health apps,” in *2016 IEEE International Conference on Communications (ICC)*, pp. 1–6, May 2016.

[96] M. Ghobakhloo, “The future of manufacturing industry: a strategic roadmap toward industry 4.0,” *Journal of Manufacturing Technology Management*, vol. 29, no. 6, pp. 910–936, 2018.

[97] Z. Cai, X. Zheng, and J. Yu, “A differential-private framework for urban traffic flows estimation via taxi companies,” *IEEE Transactions on Industrial Informatics*, pp. 1–1, 2019.

[98] H. Ahuett-Garza and T. Kurfess, “A brief discussion on the trends of habilitating technologies for industry 4.0 and smart manufacturing,” *Manufacturing Letters*, vol. 15, pp. 60 – 63, 2018. Industry 4.0 and Smart Manufacturing.

[99] H. Ahuett-Garza and T. Kurfess, “A brief discussion on the trends of habilitating technologies for industry 4.0 and smart manufacturing,” *Manufacturing Letters*, vol. 15, pp. 60 – 63, 2018. Industry 4.0 and Smart Manufacturing.

[100] S. S. Kamble, A. Gunasekaran, and S. A. Gawankar, “Sustainable industry 4.0 framework: A systematic literature review identifying the current trends and future perspectives,” *Process Safety and Environmental Protection*, vol. 117, pp. 408 – 425, 2018.

[101] M.-L. Tseng, R. R. Tan, A. S. Chiu, C.-F. Chien, and T. C. Kuo, “Circular economy meets industry 4.0: Can big data drive industrial symbiosis?,” *Resources, Conservation and Recycling*, vol. 131, pp. 146 – 147, 2018.

[102] M. Aazam, S. Zeadally, and K. A. Harras, “Deploying fog computing in industrial internet of things and industry 4.0,” *IEEE Transactions on Industrial Informatics*, vol. 14, pp. 4674–4682, Oct 2018.

[103] C. Lin, D. He, X. Huang, K.-K. R. Choo, and A. V. Vasilakos, “Bsein: A blockchain-based secure mutual authentication with fine-grained access control system for industry 4.0,” *Journal of Network and Computer Applications*, vol. 116, pp. 42 – 52, 2018.

[104] P. Pace, G. Aloi, R. Gravina, G. Caliciuri, G. Fortino, and A. Liotta, “An edge-based architecture to support efficient applications for healthcare industry 4.0,” *IEEE Transactions on Industrial Informatics*, vol. 15, pp. 481–489, Jan 2019.

[105] P. O’Donovan, C. Gallagher, K. Bruton, and D. T. O’Sullivan, “A fog computing industrial cyber-physical system for embedded low-latency machine learning industry 4.0 applications,” *Manufacturing Letters*, vol. 15, pp. 139 – 142, 2018. Industry 4.0 and Smart Manufacturing.

[106] L. D. Xu, E. L. Xu, and L. Li, “Industry 4.0: state of the art and future trends,” *International Journal of Production Research*, vol. 56, no. 8, pp. 2941–2962, 2018.

[107] Y. Wang, L. Kung, and T. A. Byrd, "Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations," *Technological Forecasting and Social Change*, vol. 126, pp. 3 – 13, 2018.

[108] K. Zhang, H. Gao, X. Han, Z. Cai, and J. Li, "Modeling and computing probabilistic skyline on incomplete data," *IEEE Transactions on Knowledge and Data Engineering*, pp. 1–1, 2019.

[109] A. L. Beam and I. S. Kohane, "Big Data and Machine Learning in Health CareBig Data and Machine Learning in Health CareBig Data and Machine Learning in Health Care," *JAMA*, vol. 319, pp. 1317–1318, 04 2018.

[110] Z. Cai, Y. Shi, M. Song, R. Goebel, and G. Lin, "Smoothing blemished gene expression microarray data via missing value imputation," in *The 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 5688–5691, IEEE, 2008.

[111] V. Grover, R. H. Chiang, T.-P. Liang, and D. Zhang, "Creating strategic business value from big data analytics: A research framework," *Journal of Management Information Systems*, vol. 35, no. 2, pp. 388–423, 2018.

[112] Y. Yang, X. Zheng, W. Guo, X. Liu, and V. Chang, "Privacy-preserving smart iot-based healthcare big data storage and self-adaptive access control system," *Information Sciences*, vol. 479, pp. 567 – 592, 2019.

[113] A. Oussous, F.-Z. Benjelloun, A. A. Lahcen, and S. Belfkih, "Big data technologies: A survey," *Journal of King Saud University - Computer and Information Sciences*, vol. 30, no. 4, pp. 431 – 448, 2018.

[114] X. Wang, Y. Zhang, V. C. M. Leung, N. Guizani, and T. Jiang, "D2d big data: Content deliveries over wireless device-to-device sharing in large-scale mobile networks," *IEEE Wireless Communications*, vol. 25, pp. 32–38, February 2018.

[115] R. Jiang, R. Lu, and K.-K. R. Choo, “Achieving high performance and privacy-preserving query over encrypted multidimensional big metering data,” *Future Generation Computer Systems*, vol. 78, pp. 392 – 401, 2018.

[116] Z. Cai, Z. He, X. Guan, and Y. Li, “Collective data-sanitization for preventing sensitive information inference attacks in social networks,” *IEEE Transactions on Dependable and Secure Computing*, vol. 15, no. 4, pp. 577–590, 2018.

[117] S. J. Mooney and V. Pejaver, “Big data in public health: Terminology, machine learning, and privacy,” *Annual Review of Public Health*, vol. 39, no. 1, pp. 95–112, 2018.

[118] D. Zhang, “Big data security and privacy protection,” in *8th International Conference on Management and Computer Science (ICMCS 2018)*, Atlantis Press, 2018/10.

[119] L. Yu, L. Chen, Z. Cai, H. Shen, Y. Liang, and Y. Pan, “Stochastic load balancing for virtual resource management in data centers,” *IEEE Transactions on Cloud Computing*, vol. PP, no. 99, pp. 1–1, 2017.

[120] R. Khatoun and S. Zeadally, “Cybersecurity and privacy solutions in smart cities,” *IEEE Communications Magazine*, vol. 55, pp. 51–59, March 2017.

[121] C. Patsakis, P. Laird, M. Clear, M. Bouroche, and A. Solanas, “Interoperable privacy-aware e-participation within smart cities,” *Computer*, vol. 48, pp. 52–58, Jan 2015.

[122] K. Zhang, J. Ni, K. Yang, X. Liang, J. Ren, and X. S. Shen, “Security and privacy in smart city applications: Challenges and solutions,” *IEEE Communications Magazine*, vol. 55, pp. 122–129, January 2017.

[123] A. W. Burange and H. D. Misalkar, “Review of internet of things in development of smart cities with data management privacy,” in *2015 International Conference on Advances in Computer Engineering and Applications*, pp. 189–195, March 2015.

[124] D. Eckhoff and I. Wagner, “Privacy in the smart cityapplications, technologies, challenges, and solutions,” *IEEE Communications Surveys Tutorials*, vol. 20, pp. 489–516, Firstquarter 2018.

[125] L. van Zoonen, “Privacy concerns in smart cities,” *Government Information Quarterly*, vol. 33, no. 3, pp. 472 – 480, 2016.

[126] M. Mukherjee, R. Matam, L. Shu, L. Maglaras, M. A. Ferrag, N. Choudhury, and V. Kumar, “Security and privacy in fog computing: Challenges,” *IEEE Access*, vol. 5, pp. 19293–19304, 2017.

[127] X. Zheng, A. Chen, G. Luo, L. Tian, and Z. Cai, “Privacy-preserved distinct content collection in human-assisted ubiquitous computing systems,” *Information Sciences*, vol. 493, pp. 91–104, 2019.

[128] X. Liang, S. Shetty, D. Tosh, C. Kamhoua, K. Kwiat, and L. Njilla, “Provchain: A blockchain-based data provenance architecture in cloud environment with enhanced privacy and availability,” in *Proceedings of the 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, CCGrid ’17, (Piscataway, NJ, USA), pp. 468–477, IEEE Press, 2017.

[129] T. Zhu, T. Shi, J. Li, Z. Cai, and X. Zhou, “Task scheduling in deadline-aware mobile edge computing systems,” *IEEE Internet of Things Journal*, 2018.

[130] J. Xiong, F. Li, J. Ma, X. Liu, Z. Yao, and P. S. Chen, “A full lifecycle privacy protection scheme for sensitive data in cloud computing,” *Peer-to-Peer Networking and Applications*, vol. 8, pp. 1025–1037, Nov 2015.

[131] A. Alrawais, A. Alhothaily, C. Hu, and X. Cheng, “Fog computing for the internet of things: Security and privacy issues,” *IEEE Internet Computing*, vol. 21, pp. 34–42, Mar 2017.

[132] Z. Tari, X. Yi, U. S. Premarathne, P. Bertok, and I. Khalil, “Security and privacy in cloud computing: Vision, trends, and challenges,” *IEEE Cloud Computing*, vol. 2, pp. 30–38, Mar 2015.

[133] K. Gai, M. Qiu, and H. Zhao, “Privacy-preserving data encryption strategy for big data in mobile cloud computing,” *IEEE Transactions on Big Data*, pp. 1–1, 2018.

[134] P. Manuel, “A trust model of cloud computing based on quality of service,” *Annals of Operations Research*, vol. 233, pp. 281–292, Oct 2015.

[135] Q. Zhang, L. T. Yang, Z. Chen, and P. Li, “Pphopcm: Privacy-preserving high-order possibilistic c-means algorithm for big data clustering with cloud computing,” *IEEE Transactions on Big Data*, pp. 1–1, 2018.

[136] K. Gai, M. Qiu, H. Zhao, and J. Xiong, “Privacy-aware adaptive data encryption strategy of big data in cloud computing,” in *2016 IEEE 3rd International Conference on Cyber Security and Cloud Computing (CSCloud)*, pp. 273–278, June 2016.

[137] Q. Jiang, J. Ma, and F. Wei, “On the security of a privacy-aware authentication scheme for distributed mobile cloud computing services,” *IEEE Systems Journal*, vol. 12, pp. 2039–2042, June 2018.

[138] J. Tsai and N. Lo, “A privacy-aware authentication scheme for distributed mobile cloud computing services,” *IEEE Systems Journal*, vol. 9, pp. 805–815, Sep. 2015.

[139] M. Satyanarayanan, “The emergence of edge computing,” *Computer*, vol. 50, pp. 30–39, Jan 2017.

[140] S. K. Pasupuleti, S. Ramalingam, and R. Buyya, “An efficient and secure privacy-preserving approach for outsourced data of resource constrained mobile devices in cloud computing,” *Journal of Network and Computer Applications*, vol. 64, pp. 12 – 22, 2016.

[141] M. B. Mollah, M. A. K. Azad, and A. Vasilakos, "Security and privacy challenges in mobile cloud computing: Survey and way ahead," *Journal of Network and Computer Applications*, vol. 84, pp. 38 – 54, 2017.

[142] Z. Xia, Y. Zhu, X. Sun, Z. Qin, and K. Ren, "Towards privacy-preserving content-based image retrieval in cloud computing," *IEEE Transactions on Cloud Computing*, vol. 6, pp. 276–286, Jan 2018.

[143] J. Li, Z. Liu, X. Chen, F. Xhafa, X. Tan, and D. S. Wong, "L-encdb: A lightweight framework for privacy-preserving data queries in cloud computing," *Knowledge-Based Systems*, vol. 79, pp. 18 – 26, 2015.

[144] W. Shi and S. Dustdar, "The promise of edge computing," *Computer*, vol. 49, pp. 78–81, May 2016.

[145] Z. Xia, N. N. Xiong, A. V. Vasilakos, and X. Sun, "Epcbir: An efficient and privacy-preserving content-based image retrieval scheme in cloud computing," *Information Sciences*, vol. 387, pp. 195 – 204, 2017.

[146] W. Zhang, Y. Lin, S. Xiao, J. Wu, and S. Zhou, "Privacy preserving ranked multi-keyword search for multiple data owners in cloud computing," *IEEE Transactions on Computers*, vol. 65, pp. 1566–1577, May 2016.

[147] Y. Zhang, X. Chen, J. Li, D. S. Wong, H. Li, and I. You, "Ensuring attribute privacy protection and fast decryption for outsourced data security in mobile cloud computing," *Information Sciences*, vol. 379, pp. 42 – 61, 2017.

[148] P. Li, J. Li, Z. Huang, T. Li, C.-Z. Gao, S.-M. Yiu, and K. Chen, "Multi-key privacy-preserving deep learning in cloud computing," *Future Generation Computer Systems*, vol. 74, pp. 76 – 85, 2017.

[149] P. Li, J. Li, Z. Huang, C.-Z. Gao, W.-B. Chen, and K. Chen, "Privacy-preserving outsourced classification in cloud computing," *Cluster Computing*, vol. 21, pp. 277–286, Mar 2018.

[150] Z. Xia, X. Wang, L. Zhang, Z. Qin, X. Sun, and K. Ren, "A privacy-preserving and copy-deterrence content-based image retrieval scheme in cloud computing," *IEEE Transactions on Information Forensics and Security*, vol. 11, pp. 2594–2608, Nov 2016.

[151] J. Sun, R. Zhang, X. Jin, and Y. Zhang, "Securefind: Secure and privacy-preserving object finding via mobile crowdsourcing," *IEEE Transactions on Wireless Communications*, vol. 15, pp. 1716–1728, March 2016.

[152] Y. Wang, Z. Cai, Z.-H. Zhan, Y.-J. Gong, and X. Tong, "An optimization and auction-based incentive mechanism to maximize social welfare for mobile crowdsourcing," *IEEE Transactions on Computational Social Systems*, 2019.

[153] G. Zhuo, Q. Jia, L. Guo, M. Li, and P. Li, "Privacy-preserving verifiable data aggregation and analysis for cloud-assisted mobile crowdsourcing," in *IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications*, pp. 1–9, April 2016.

[154] Z. Duan, M. Yan, Z. Cai, X. Wang, M. Han, and Y. Li, "Truthful incentive mechanisms for social cost minimization in mobile crowdsourcing systems," *Sensors*, vol. 16, no. 4, p. 481, 2016.

[155] Y. Gong, L. Wei, Y. Guo, C. Zhang, and Y. Fang, "Optimal task recommendation for mobile crowdsourcing with privacy control," *IEEE Internet of Things Journal*, vol. 3, pp. 745–756, Oct 2016.

[156] J. Li, Z. Cai, J. Wang, M. Han, and Y. Li, "Truthful incentive mechanisms for geographical position conflicting mobile crowdsensing systems," *IEEE Transactions on Computational Social Systems*, vol. 5, no. 2, pp. 324–334, 2018.

[157] B. Zhang, C. H. Liu, J. Lu, Z. Song, Z. Ren, J. Ma, and W. Wang, "Privacy-preserving qoi-aware participant coordination for mobile crowdsourcing," *Computer Networks*, vol. 101, pp. 29 – 41, 2016. Industrial Technologies and Applications for the Internet of Things.

[158] Q. Zhang, L. T. Yang, Z. Chen, P. Li, and M. J. Deen, "Privacy-preserving double-projection deep computation model with crowdsourcing on cloud for big data feature learning," *IEEE Internet of Things Journal*, vol. 5, pp. 2896–2903, Aug 2018.

[159] Y. Wang, Z. Cai, X. Tong, Y. Gao, and G. Yin, "Truthful incentive mechanism with location privacy-preserving for mobile crowdsourcing systems," *Computer Networks*, vol. 135, pp. 32–43, 2018.

[160] S. Wilson, F. Schaub, R. Ramanath, N. Sadeh, F. Liu, N. A. Smith, and F. Liu, "Crowdsourcing annotations for websites' privacy policies: Can it really work?," in *Proceedings of the 25th International Conference on World Wide Web*, WWW '16, (Republic and Canton of Geneva, Switzerland), pp. 133–143, International World Wide Web Conferences Steering Committee, 2016.

[161] K. Yang, K. Zhang, J. Ren, and X. Shen, "Security and privacy in mobile crowdsourcing networks: challenges and opportunities," *IEEE Communications Magazine*, vol. 53, pp. 75–81, August 2015.

[162] Y. Wang, Z. Cai, G. Yin, Y. Gao, X. Tong, and G. Wu, "An incentive mechanism with privacy protection in mobile crowdsourcing systems," *Computer Networks*, vol. 102, pp. 157 – 171, 2016.

[163] S. Egelman, R. Kannavara, and R. Chow, "Is this thing on?: Crowdsourcing privacy indicators for ubiquitous sensing platforms," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, (New York, NY, USA), pp. 1669–1678, ACM, 2015.

[164] Z. Duan, W. Li, and Z. Cai, "Distributed auctions for task assignment and scheduling in mobile crowdsensing systems," in *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, pp. 635–644, IEEE, 2017.

[165] M. Li, J. Weng, A. Yang, W. Lu, Y. Zhang, L. Hou, J. Liu, Y. Xiang, and R. H. Deng, "Crowdbc: A blockchain-based decentralized framework for crowdsourcing," *IEEE Transactions on Parallel and Distributed Systems*, vol. 30, pp. 1251–1266, June 2019.

[166] A. Liu, W. Wang, S. Shang, Q. Li, and X. Zhang, "Efficient task assignment in spatial crowdsourcing with worker and task privacy protection," *GeoInformatica*, vol. 22, pp. 335–362, Apr 2018.

[167] J. Shu, X. Jia, K. YANG, and H. Wang, "Privacy-preserving task recommendation services for crowdsourcing," *IEEE Transactions on Services Computing*, pp. 1–1, 2018.

[168] J. Shu, X. Liu, X. Jia, K. Yang, and R. H. Deng, "Anonymous privacy-preserving task matching in crowdsourcing," *IEEE Internet of Things Journal*, vol. 5, pp. 3068–3078, Aug 2018.

[169] Z. Duan, W. Li, X. Zheng, and Z. Cai, "Mutual-preference driven truthful auction mechanism in mobile crowdsensing," in *The 39th IEEE International Conference on Distributed Computing Systems (ICDCS 2019)*, IEEE, 2019.

[170] L. Zhang, Z. Cai, P. Li, L. Wang, and X. Wang, "Spectrum-availability based routing for cognitive sensor networks," *IEEE Access*, vol. 5, pp. 4448–4457, 2017.

[171] S. Ji, M. Yan, R. Beyah, and Z. Cai, "Semi-structure routing and analytical frameworks for cognitive radio networks," *IEEE Transactions on Mobile Computing*, vol. 15, no. 4, pp. 996–1008, 2016.

[172] L. Yu and Z. Cai, “Dynamic scaling of virtual clusters with bandwidth guarantee in cloud datacenters,” in *The 35th Annual IEEE International Conference on Computer Communications (INFOCOM 2016)*, pp. 1–9, April 2016.

[173] Z. Cai, S. Ji, J. He, L. Wei, and A. G. Bourgeois, “Distributed and asynchronous data collection in cognitive radio networks with fairness consideration,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 8, pp. 2020–2029, 2014.

[174] Z. Cai, S. Ji, J. He, and A. G. Bourgeois, “Optimal distributed data collection for asynchronous cognitive radio networks,” in *The 32nd Proceeding. of International Conference on Distributed Computing Systems 2012 (ICDCS 2012)*, pp. 245–254, IEEE, 2012.

[175] L. Yang, Z. Han, Z. Huang, and J. Ma, “A remotely keyed file encryption scheme under mobile cloud computing,” *Journal of Network and Computer Applications*, vol. 106, pp. 90 – 99, 2018.

[176] G. Luo, K. Yan, X. Zheng, L. Tian, and Z. Cai, “Preserving adjustable path privacy for task acquisition in mobile crowdsensing systems,” *Information Sciences*, 2018.

[177] R. Roman, J. Lopez, and M. Mambo, “Mobile edge computing, fog et al.: A survey and analysis of security threats and challenges,” *Future Generation Computer Systems*, vol. 78, pp. 680 – 698, 2018.

[178] L. Zhang, Z. Cai, J. Lu, and X. Wang, “Mobility-aware routing in delay tolerant networks,” *Personal and Ubiquitous Computing*, vol. 19, no. 7, pp. 1111–1123, 2015.

[179] X. Jin and Y. Zhang, “Privacy-preserving crowdsourced spectrum sensing,” *IEEE/ACM Trans. Netw.*, vol. 26, pp. 1236–1249, June 2018.

[180] F. H. McKay, C. Cheng, A. Wright, J. Shill, H. Stephens, and M. Uccellini, “Evaluating mobile phone applications for health behaviour change: A systematic review,” *Journal of Telemedicine and Telecare*, vol. 24, no. 1, pp. 22–30, 2018.

[181] K. Gai, K. R. Choo, M. Qiu, and L. Zhu, "Privacy-preserving content-oriented wireless communication in internet-of-things," *IEEE Internet of Things Journal*, vol. 5, pp. 3059–3067, Aug 2018.

[182] A. Papageorgiou, M. Strigkos, E. Politou, E. Alepis, A. Solanas, and C. Patsakis, "Security and privacy analysis of mobile health applications: The alarming state of practice," *IEEE Access*, vol. 6, pp. 9390–9403, 2018.

[183] X. Wang, Y. Lin, Y. Zhao, L. Zhang, J. Liang, and Z. Cai, "A novel approach for inhibiting misinformation propagation in human mobile opportunistic networks," *Peer-to-Peer Networking and Applications*, vol. 10, no. 2, pp. 377–394, 2017.

[184] Y. Wang, Z. Cai, X. Tong, Y. Gao, and G. Yin, "Truthful incentive mechanism with location privacy-preserving for mobile crowdsourcing systems," *Computer Networks*, vol. 135, pp. 32 – 43, 2018.

[185] Z. He, Z. Cai, Q. Han, W. Tong, L. Sun, and Y. Li, "An energy efficient privacy-preserving content sharing scheme in mobile social networks," *Personal and Ubiquitous Computing*, vol. 20, no. 5, pp. 833–846, 2016.

[186] Z. Wang, J. Hu, R. Lv, J. Wei, Q. Wang, D. Yang, and H. Qi, "Personalized privacy-preserving task allocation for mobile crowdsensing," *IEEE Transactions on Mobile Computing*, vol. 18, pp. 1330–1341, June 2019.

[187] Y. Wang, Z. Cai, G. Yin, Y. Gao, X. Tong, and G. Wu, "An incentive mechanism with privacy protection in mobile crowdsourcing systems," *Computer Networks*, vol. 102, pp. 157–171, 2016.

[188] C.-z. Gao, Q. Cheng, X. Li, and S.-b. Xia, "Cloud-assisted privacy-preserving profile-matching scheme under multiple keys in mobile social network," *Cluster Computing*, Feb 2018.

[189] J. Wang, J. Zhang, W. Bao, X. Zhu, B. Cao, and P. S. Yu, "Not just privacy: Improving performance of private deep learning in mobile cloud," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery &#38; Data Mining*, KDD '18, pp. 2407–2416, 2018.

[190] X. Xiao, C. Chen, A. K. Sangaiah, G. Hu, R. Ye, and Y. Jiang, "Cenlocshare: A centralized privacy-preserving location-sharing system for mobile online social networks," *Future Generation Computer Systems*, vol. 86, pp. 863 – 872, 2018.

[191] Y. Huang, Z. Cai, and A. G. Bourgeois, "Search locations safely and accurately: A location privacy protection algorithm with accurate service," *Journal of Network and Computer Applications*, vol. 103, pp. 146–156, 2018.

[192] L. Zhang, X. Wang, J. Lu, P. Li, and Z. Cai, "An efficient privacy preserving data aggregation approach for mobile sensing," *Security and Communication Networks*, vol. 9, no. 16, pp. 3844–3853, 2016.

[193] X. Wang, Y. Lin, S. Zhang, and Z. Cai, "A social activity and physical contact-based routing algorithm in mobile opportunistic networks for emergency response to sudden disasters," *Enterprise Information Systems*, vol. 11, no. 5, pp. 597–626, 2017.

[194] K. Yoshigoe, W. Dai, M. Abramson, and A. Jacobs, "Overcoming invasion of privacy in smart home environment with synthetic packet injection," in *2015 TRON Symposium (TRONSHOW)*, pp. 1–7, Dec 2014.

[195] Q. Chen, H. Gao, Z. Cai, L. Cheng, and J. Li, "Distributed low-latency data aggregation for duty-cycle wireless sensor networks," *IEEE/ACM Transactions on Networking (TON)*, vol. 26, no. 5, pp. 2347–2360, 2018.

[196] S. B. Othman, A. A. Bahattab, A. Trad, and H. Youssef, "Confidentiality and integrity for data aggregation in wsn using homomorphic encryption," *Wireless Personal Communications*, vol. 80, pp. 867–889, Jan 2015.

[197] Q. Chen, Z. Cai, L. Cheng, H. Gao, and J. Li, “Low-latency concurrent broadcast scheduling in duty-cycled multihop wireless networks,” in *The 39th IEEE International Conference on Distributed Computing Systems (ICDCS 2019)*, IEEE, 2019.

[198] F. Wu, L. Xu, S. Kumari, and X. Li, “A privacy-preserving and provable user authentication scheme for wireless sensor networks based on internet of things security,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 8, pp. 101–116, Feb 2017.

[199] T. Shi, Z. Cai, J. Li, and H. Gao, “The energy-data dual coverage in battery-free sensor networks,” in *The 39th IEEE International Conference on Distributed Computing Systems (ICDCS 2019)*, IEEE, 2019.

[200] R. Rios, J. Cuellar, and J. Lopez, “Probabilistic receiver-location privacy protection in wireless sensor networks,” *Information Sciences*, vol. 321, pp. 205 – 223, 2015. Security and privacy information technologies and applications for wireless pervasive computing environments.

[201] K. Chen, H. Gao, Z. Cai, Q. Chen, and J. Li, “Distributed energy-adaptive aggregation scheduling with coverage guarantee for battery-free wireless sensor networks,” in *The 38th Annual IEEE International Conference on Computer Communications (INFOCOM 2019)*, IEEE, 2019.

[202] M. Bradbury, M. Leeke, and A. Jhumka, “A dynamic fake source algorithm for source location privacy in wireless sensor networks,” in *2015 IEEE Trustcom/BigDataSE/ISPA*, vol. 1, pp. 531–538, Aug 2015.

[203] T. Shi, J. Li, H. Gao, and Z. Cai, “Coverage in battery-free wireless sensor networks,” in *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*, pp. 108–116, IEEE, 2018.

[204] W. Feng, Z. Yan, H. Zhang, K. Zeng, Y. Xiao, and Y. T. Hou, “A survey on security, privacy, and trust in mobile crowdsourcing,” *IEEE Internet of Things Journal*, vol. 5, pp. 2971–2992, Aug 2018.

[205] X. Zheng, Z. Cai, J. Li, and H. Gao, “A study on application-aware scheduling in wireless networks,” *IEEE Transactions on Mobile Computing*, vol. 16, no. 7, pp. 1787–1801, 2016.

[206] M. Elhoseny, H. Elminir, A. Riad, and X. Yuan, “A secure data routing schema for wsn using elliptic curve cryptography and homomorphic encryption,” *Journal of King Saud University - Computer and Information Sciences*, vol. 28, no. 3, pp. 262 – 275, 2016.

[207] K. Zhang, Q. Han, Z. Cai, and G. Yin, “Rippas: a ring-based privacy-preserving aggregation scheme in wireless sensor networks,” *Sensors*, vol. 17, no. 2, p. 300, 2017.

[208] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao, “A survey on internet of things: Architecture, enabling technologies, security and privacy, and applications,” *IEEE Internet of Things Journal*, vol. 4, pp. 1125–1142, Oct 2017.

[209] Q. Chen, H. Gao, Z. Cai, L. Cheng, and J. Li, “Energy-collision aware data aggregation scheduling for energy harvesting sensor networks,” in *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*, pp. 117–125, IEEE, 2018.

[210] E. Al Alkeem, C. Y. Yeun, and M. J. Zemerly, “Security and privacy framework for ubiquitous healthcare iot devices,” in *2015 10th International Conference for Internet Technology and Secured Transactions (ICITST)*, pp. 70–75, Dec 2015.

[211] T. Shi, S. Cheng, J. Li, and Z. Cai, “Constructing connected dominating sets in battery-free networks,” in *IEEE INFOCOM 2017-IEEE Conference on Computer Communications*, pp. 1–9, IEEE, 2017.

[212] A. Liu, X. Liu, Z. Tang, L. T. Yang, and Z. Shao, “Preserving smart sink-location privacy with delay guaranteed routing scheme for wsns,” *ACM Trans. Embed. Comput. Syst.*, vol. 16, pp. 68:1–68:25, May 2017.

[213] Z. Cai, D. Miao, and Y. Li, “Deletion propagation for multiple select-join queries: Approximations and complexity,” in *The 35th IEEE International Conference on Data Engineering (ICDE 2019)*, IEEE, 2019.

[214] F. Dalipi and S. Y. Yayilgan, “Security and privacy considerations for iot application on smart grids: Survey and research challenges,” in *2016 IEEE 4th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW)*, pp. 63–68, Aug 2016.

[215] J. Gao, J. Li, Z. Cai, and H. Gao, “Composite event coverage in wireless sensor networks with heterogeneous sensors,” in *The 34th Annual IEEE International Conference on Computer Communications (INFOCOM 2015)*, pp. 217–225, April 2015.

[216] H. Huang, T. Gong, P. Chen, R. Malekian, and T. Chen, “Secure two-party distance computation protocol based on privacy homomorphism and scalar product in wireless sensor networks,” *Tsinghua Science and Technology*, vol. 21, pp. 385–396, Aug 2016.

[217] K. Zhang, Q. Han, Z. Cai, G. Yin, and J. Lin, “Doami: A distributed on-line algorithm to minimize interference for routing in wireless sensor networks,” *Theoretical Computer Science*, 2016.

[218] Y. Yang, L. Wu, G. Yin, L. Li, and H. Zhao, “A survey on security and privacy issues in internet-of-things,” *IEEE Internet of Things Journal*, vol. 4, pp. 1250–1258, Oct 2017.

[219] M. Ren, J. Li, L. Guo, and Z. Cai, “Data collection with probabilistic guarantees in opportunistic wireless networks,” *International Journal of Sensor Networks*, vol. 24, no. 2, pp. 125–137, 2017.

[220] J. Lopez, R. Rios, F. Bao, and G. Wang, "Evolving privacy: From sensors to the internet of things," *Future Generation Computer Systems*, vol. 75, pp. 46 – 57, 2017.

[221] S. Cheng, Z. Cai, J. Li, and H. Gao, "Extracting kernel dataset from big sensory data in wireless sensor networks," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 4, pp. 813–827, 2016.

[222] A. Alaiad and L. Zhou, "Patients' adoption of wsn-based smart home healthcare systems: An integrated model of facilitators and barriers," *IEEE Transactions on Professional Communication*, vol. 60, pp. 4–23, March 2017.

[223] T. Zhu, S. Cheng, Z. Cai, and J. Li, "Critical data points retrieving method for big sensory data in wireless sensor networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2016, no. 1, p. 18, 2016.

[224] X. Li, J. Niu, M. Z. A. Bhuiyan, F. Wu, M. Karuppiah, and S. Kumari, "A robust ecc-based provable secure authentication protocol with privacy preserving for industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 14, pp. 3599–3609, Aug 2018.

[225] S. Sicari, A. Rizzardi, L. Grieco, and A. Coen-Porisini, "Security, privacy and trust in internet of things: The road ahead," *Computer Networks*, vol. 76, pp. 146 – 164, 2015.

[226] U. Senthil kumaran and P. Ilango, "Secure authentication and integrity techniques for randomized secured routing in wsn," *Wireless Networks*, vol. 21, pp. 443–451, Feb 2015.

[227] A. Abuzneid, T. Sobh, and M. Faezipour, "An enhanced communication protocol for anonymity and location privacy in wsn," in *2015 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, pp. 91–96, March 2015.

[228] Z. Cai, S. Ji, and J. Li, "Data caching-based query processing in multi-sink wireless sensor networks," *Int. J. Sen. Netw.*, vol. 11, pp. 109–125, Mar. 2012.

[229] H. Chen and W. Lou, "On protecting end-to-end location privacy against local eavesdropper in wireless sensor networks," *Pervasive and Mobile Computing*, vol. 16, pp. 36 – 50, 2015.

[230] S. Cheng, Z. Cai, J. Li, and X. Fang, "Drawing dominant dataset from big sensory data in wireless sensor networks," in *The 34th Annual IEEE International Conference on Computer Communications (INFOCOM 2015)*, pp. 531–539, April 2015.

[231] T. Gong, H. Huang, P. Li, K. Zhang, and H. Jiang, "A medical healthcare system for privacy protection based on iot," in *2015 Seventh International Symposium on Parallel Architectures, Algorithms and Programming (PAAP)*, pp. 217–222, Dec 2015.

[232] Q. Chen, H. Gao, S. Cheng, J. Li, and Z. Cai, "Distributed non-structure based data aggregation for duty-cycle wireless sensor networks," in *IEEE INFOCOM 2017-IEEE Conference on Computer Communications*, pp. 1–9, IEEE, 2017.

[233] A.-S. Abuzneid, T. Sobh, M. Faezipour, A. Mahmood, and J. James, "Fortified anonymous communication protocol for location privacy in wsn: A modular approach," *Sensors*, vol. 15, no. 3, pp. 5820–5864, 2015.

[234] D. Miao, Z. Cai, and J. Li, "On the complexity of bounded view propagation for conjunctive queries," *IEEE Transactions on Knowledge and Data Engineering*, vol. 30, no. 1, pp. 115–127, 2017.

[235] P. Kumar, J. P. Singh, P. Vishnoi, and M. P. Singh, "Source location privacy using multiple-phantom nodes in wsn," in *TENCON 2015 - 2015 IEEE Region 10 Conference*, pp. 1–6, Nov 2015.

[236] H. Li, R. Lu, J. Misic, and M. Mahmoud, "Security and privacy of connected vehicular cloud computing," *IEEE Network*, vol. 32, pp. 4–6, May 2018.

[237] X. Wang, Z. Ning, M. Zhou, X. Hu, L. Wang, Y. Zhang, F. R. Yu, and B. Hu, "Privacy-preserving content dissemination for vehicular social networks: Challenges

and solutions," *IEEE Communications Surveys Tutorials*, vol. 21, pp. 1314–1345, Secondquarter 2019.

[238] Y. Huang, X. Guan, Z. Cai, and T. Ohtsuki, "Multicast capacity analysis for social-proximity urban bus-assisted vanets," in *Communications (ICC), 2013 IEEE International Conference on*, pp. 6138–6142, IEEE, 2013.

[239] L. Li, J. Liu, L. Cheng, S. Qiu, W. Wang, X. Zhang, and Z. Zhang, "Creditcoin: A privacy-preserving blockchain-based incentive announcement network for communications of smart vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, pp. 2204–2220, July 2018.

[240] Z. Yang, K. Yang, L. Lei, K. Zheng, and V. C. M. Leung, "Blockchain-based decentralized trust management in vehicular networks," *IEEE Internet of Things Journal*, vol. 6, pp. 1495–1505, April 2019.

[241] J. Kang, R. Yu, X. Huang, M. Wu, S. Maharjan, S. Xie, and Y. Zhang, "Blockchain for secure and efficient data sharing in vehicular edge computing and networks," *IEEE Internet of Things Journal*, vol. 6, pp. 4660–4670, June 2019.

[242] J. Ni, K. Zhang, Y. Yu, X. Lin, and X. Shen, "Privacy-preserving smart parking navigation supporting efficient driving guidance retrieval," *IEEE Transactions on Vehicular Technology*, vol. 67, pp. 6504–6517, July 2018.

[243] X. Guan, Y. Huang, Z. Cai, and T. Ohtsuki, "Intersection-based forwarding protocol for vehicular ad hoc networks," *Telecommunication Systems*, vol. 62, no. 1, pp. 67–76, 2016.

[244] K. Wang, H. Yin, W. Quan, and G. Min, "Enabling collaborative edge computing for software defined vehicular networks," *IEEE Network*, vol. 32, pp. 112–117, Sep. 2018.

[245] X. Zheng, Z. Cai, J. Li, and H. Gao, "An application-aware scheduling policy for real-time traffic," in *The 35th IEEE International Conference on Distributed Computing Systems (ICDCS 2015)*, pp. 421–430, June 2015.

[246] Z. Lu, G. Qu, and Z. Liu, "A survey on recent advances in vehicular network security, trust, and privacy," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, pp. 760–776, Feb 2019.

[247] Y. Huang, M. Chen, Z. Cai, X. Guan, T. Ohtsuki, and Y. Zhang, "Graph theory based capacity analysis for vehicular ad hoc networks," in *2015 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–5, Dec 2015.

[248] P. Vijayakumar, V. Chang, L. J. Deborah, B. Balusamy, and P. Shynu, "Computationally efficient privacy preserving anonymous mutual and batch authentication schemes for vehicular ad hoc networks," *Future Generation Computer Systems*, vol. 78, pp. 943 – 955, 2018.

[249] P. Asuquo, H. Cruickshank, J. Morley, C. P. A. Ogah, A. Lei, W. Hathal, S. Bao, and Z. Sun, "Security and privacy in location-based services for vehicular and mobile communications: An overview, challenges, and countermeasures," *IEEE Internet of Things Journal*, vol. 5, pp. 4778–4802, Dec 2018.

[250] S. Cheng, Z. Cai, and J. Li, "Curve query processing in wireless sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 11, pp. 5198–5209, 2015.

[251] L. Zhang, X. Men, K. R. Choo, Y. Zhang, and F. Dai, "Privacy-preserving cloud establishment and data dissemination scheme for vehicular cloud," *IEEE Transactions on Dependable and Secure Computing*, pp. 1–1, 2018.

[252] X. Zheng, Z. Cai, J. Li, and H. Gao, "Location-privacy-aware review publication mechanism for local business service systems," in *IEEE INFOCOM 2017-IEEE Conference on Computer Communications*, pp. 1–9, IEEE, 2017.

[253] X. Zheng, Z. Cai, G. Luo, L. Tian, and X. Bai, "Privacy-preserved community discovery in online social networks," *Future Generation Computer Systems*, vol. 93, pp. 1002–1009, 2019.

[254] X. Zheng, G. Luo, and Z. Cai, "A fair mechanism for private data publication in online social networks," *IEEE Transactions on Network Science and Engineering*, 2018.

[255] T. Shi, S. Cheng, Z. Cai, Y. Li, and J. Li, "Retrieving the maximal time-bounded positive influence set from social networks," *Personal and Ubiquitous Computing*, vol. 20, no. 5, pp. 717–730, 2016.

[256] M. Siddula, Z. Cai, and D. Miao, "Privacy preserving online social networks using enhanced equicardinal clustering," in *2018 IEEE 37th International Performance Computing and Communications Conference (IPCCC)*, pp. 1–8, IEEE, 2018.

[257] G. Li, Z. Cai, G. Yin, Z. He, and M. Siddula, "Differentially private recommendation system based on community detection in social network applications," *Security and Communication Networks*, vol. 2018, 2018.

[258] J. Lu, Z. Cai, X. Wang, L. Zhang, P. Li, and Z. He, "User social activity-based routing for cognitive radio networks," *Personal and Ubiquitous Computing*, vol. 22, no. 3, pp. 471–487, 2018.

[259] M. Han, M. Yan, Z. Cai, Y. Li, X. Cai, and J. Yu, "Influence maximization by probing partial communities in dynamic online social networks," *Transactions on Emerging Telecommunications Technologies*, vol. 28, no. 4, 2017.

[260] J. Li, Z. Cai, M. Yan, and Y. Li, "Using crowdsourced data in location-based social networks to explore influence maximization," in *The 35th Annual IEEE International Conference on Computer Communications (INFOCOM 2016)*, pp. 1–9, April 2016.

[261] Z. He, Z. Cai, and X. Wang, "Modeling propagation dynamics and developing optimized countermeasures for rumor spreading in online social networks," in *The 35th*

IEEE International Conference on Distributed Computing Systems (ICDCS 2015), pp. 205–214, June 2015.

[262] Y. Wang, G. Yin, Z. Cai, Y. Dong, and H. Dong, “A trust-based probabilistic recommendation model for social networks,” *Journal of Network and Computer Applications*, vol. 55, pp. 59–67, 2015.

[263] S. Li, L. D. Xu, and S. Zhao, “5g internet of things: A survey,” *Journal of Industrial Information Integration*, vol. 10, pp. 1 – 9, 2018.

[264] Y. Zhang, J. Li, D. Zheng, P. Li, and Y. Tian, “Privacy-preserving communication and power injection over vehicle networks and 5g smart grid slice,” *Journal of Network and Computer Applications*, vol. 122, pp. 50 – 60, 2018.

[265] Y. Zhong, X. Ge, H. H. Yang, T. Han, and Q. Li, “Traffic matching in 5g ultra-dense networks,” *IEEE Communications Magazine*, vol. 56, pp. 100–105, August 2018.

[266] I. Ahmad, T. Kumar, M. Liyanage, J. Okwuibe, M. Ylianttila, and A. Gurtov, “Overview of 5g security challenges and solutions,” *IEEE Communications Standards Magazine*, vol. 2, pp. 36–43, MARCH 2018.

[267] D. Basin, J. Dreier, L. Hirschi, S. Radomirovic, R. Sasse, and V. Stettler, “A formal analysis of 5g authentication,” in *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, CCS ’18, (New York, NY, USA), pp. 1383–1396, ACM, 2018.

[268] M. Matinmikko, M. Latva-aho, P. Ahokangas, and V. Seppnen, “On regulations for 5g: Micro licensing for locally operated networks,” *Telecommunications Policy*, vol. 42, no. 8, pp. 622 – 635, 2018.

[269] J. Cheng, W. Chen, F. Tao, and C.-L. Lin, “Industrial iot in 5g environment towards smart manufacturing,” *Journal of Industrial Information Integration*, vol. 10, pp. 10 – 19, 2018.

[270] V. Sharma, I. You, F.-Y. Leu, and M. Atiquzzaman, “Secure and efficient protocol for fast handover in 5g mobile xhaul networks,” *Journal of Network and Computer Applications*, vol. 102, pp. 38 – 57, 2018.

[271] S. Kakran and S. Chanana, “Smart operations of smart grids integrated with distributed generation: A review,” *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 524 – 535, 2018.

[272] K. K. Zame, C. A. Brehm, A. T. Nitica, C. L. Richard, and G. D. S. III, “Smart grid and energy storage: Policy recommendations,” *Renewable and Sustainable Energy Reviews*, vol. 82, pp. 1646 – 1654, 2018.

[273] R. Zafar, A. Mahmood, S. Razzaq, W. Ali, U. Naeem, and K. Shehzad, “Prosumer based energy management and sharing in smart grid,” *Renewable and Sustainable Energy Reviews*, vol. 82, pp. 1675 – 1684, 2018.

[274] K. Mahmood, S. A. Chaudhry, H. Naqvi, S. Kumari, X. Li, and A. K. Sangaiah, “An elliptic curve cryptography based lightweight authentication scheme for smart grid communication,” *Future Generation Computer Systems*, vol. 81, pp. 557 – 565, 2018.

[275] F. Knirsch, A. Unterweger, and D. Engel, “Privacy-preserving blockchain-based electric vehicle charging with dynamic tariff decisions,” *Computer Science - Research and Development*, vol. 33, pp. 71–79, Feb 2018.

[276] X. Zheng, Z. Cai, and Y. Li, “Data linkage in smart internet of things systems: A consideration from a privacy perspective,” *IEEE Communications Magazine*, vol. 56, pp. 55–61, Sep. 2018.

[277] S. Tonyali, K. Akkaya, N. Saputro, A. S. Uluagac, and M. Nojoumian, “Privacy-preserving protocols for secure and reliable data aggregation in iot-enabled smart metering systems,” *Future Generation Computer Systems*, vol. 78, pp. 547 – 557, 2018.

[278] Z. Guan, G. Si, X. Zhang, L. Wu, N. Guizani, X. Du, and Y. Ma, "Privacy-preserving and efficient aggregation based on blockchain for power grid communications in smart communities," *IEEE Communications Magazine*, vol. 56, pp. 82–88, July 2018.

[279] L. Lyu, K. Nandakumar, B. Rubinstein, J. Jin, J. Bedo, and M. Palaniswami, "Ppfa: Privacy preserving fog-enabled aggregation in smart grid," *IEEE Transactions on Industrial Informatics*, vol. 14, pp. 3733–3744, Aug 2018.

[280] Y. Liu, W. Guo, C. Fan, L. Chang, and C. Cheng, "A practical privacy-preserving data aggregation (3pda) scheme for smart grid," *IEEE Transactions on Industrial Informatics*, vol. 15, pp. 1767–1774, March 2019.

[281] T. Shi, S. Cheng, Z. Cai, and J. Li, "Adaptive connected dominating set discovering algorithm in energy-harvest sensor networks," in *The 35th Annual IEEE International Conference on Computer Communications (INFOCOM 2016)*, pp. 1–9, April 2016.

[282] S. Zhu, W. Li, H. Li, L. Tian, G. Luo, and Z. Cai, "Coin hopping attack in blockchain-based iot," *IEEE Internet of Things Journal*, 2018.

[283] Z. Cai and X. Zheng, "A private and efficient mechanism for data uploading in smart cyber-physical systems," *IEEE Transactions on Network Science and Engineering*, pp. 1–1, 2018.

[284] Z. He, Z. Cai, Y. Sun, Y. Li, and X. Cheng, "Customized privacy preserving for inherent data and latent data," *Personal and Ubiquitous Computing*, vol. 21, no. 1, pp. 43–54, 2017.

[285] M. Han, M. Yan, Z. Cai, and Y. Li, "An exploration of broader influence maximization in timeliness networks with opportunistic selection," *Journal of Network and Computer Applications*, vol. 63, pp. 39–49, 2016.

[286] J. Li, S. Cheng, H. Gao, and Z. Cai, “Approximate physical world reconstruction algorithms in sensor networks,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 12, pp. 3099–3110, 2014.

[287] S. Cheng, J. Li, and Z. Cai, “O (ε)-approximation to physical world by sensor networks,” in *The 32rd Annual IEEE International Conference on Computer Communications (IEEE INFOCOM 2013)*, pp. 3084–3092, IEEE, 2013.

[288] Z. Cai, C. Wang, and A. Bourgeois, “Preface: Special issue on computing and combinatorics conference and wireless algorithms, systems, and applications conference,” 2016.

[289] X. Fang, J. Luo, G. Luo, W. Wu, Z. Cai, and Y. Pan, “Big data transmission in industrial iot systems with small capacitor supplying energy,” *IEEE Transactions on Industrial Informatics*, vol. 15, pp. 2360–2371, April 2019.

[290] X. Zheng, Z. Cai, J. Li, and H. Gao, “Scheduling flows with multiple service frequency constraints,” *IEEE Internet of Things Journal*, vol. 4, no. 2, pp. 496–504, 2017.

[291] Y. Liang, Z. Cai, J. Yu, Q. Han, and Y. Li, “Deep learning based inference of private information using embedded sensors in smart devices,” *IEEE Network*, vol. 32, no. 4, pp. 8–14, 2018.