A Stochastic Game for Privacy Preserving Context Sensing on Mobile Phone

Wei Wang, Qian Zhang
Department of Computer Science and Engineering,
Hong Kong University of Science and Technology
Email: {gswwang, qianzh}@ust.hk

***Abstract*—The proliferation of sensor-equipped smartphones has enabled an increasing number of context-aware applications that provide personalized services based on users' contexts. However, most of these applications aggressively collect users sensing data without providing clear statements on the usage and disclosure strategies of such sensitive information, which raises severe privacy concerns and leads to some initial investigation on privacy preservation mechanisms design. While most prior studies have assumed static adversary models, we investigate the context dynamics and call attention to the existence of intelligent adversaries. In this paper, we first identify the context privacy problem with consideration of the context dynamics and malicious adversaries with capabilities of adjusting their attacking strategies, and then formulate the interactive competition between users and adversaries as a zero-sum stochastic game. In addition, we propose an efficient minimax learning algorithm to obtain the optimal defense strategy. Our evaluations on real smartphone context traces of 94 users validate the proposed algorithm.**

I. Introduction

The increasing popularity of smartphones equipped with a variety of sensors provides new opportunities for the proliferation of context-aware applications that offer personalized services based on the operating conditions of smartphone users and their surrounding environments. Such applications effectively use sensors such as GPS, accelerometer, proximity sensor and microphone to infer smartphone user's current context including location, mobility mode (e.g., walking or driving), and social activities. Examples of context-aware applications include *GeoNote*[1] that reminds a user of something when he is at a particular location and *AutoSilent*[2] that automatically mutes the phone when the user is in a meeting.

Although context-aware applications improve user experiences on smartphones, severe privacy issues arise with these applications. Nowadays, the growing privacy threats of sharing location-related context information via context-aware applications on smartphones have been concerned by both consumers [1] and governments [2]. Such privacy threats come from the fact that many smartphone applications aggressively collect sensing data without clear statements about how to use the sensing data and whom the sensing data will be shared with. Untrusted applications may sell such personal information to advertisers without user's permission. Enck et al. [3] studied 30 popular Android applications that have access to user's location, camera, microphone data, and found that 15 of them sent users' information to remote advertisement or analytics servers. Moreover, malicious adversaries with criminal intent could hack the applications with such information to pose a threat to individual security and privacy. Being aware of such risks, the smartphone users may not allow the applications to access their sensing data, which, however, disables the functionalities provided by the context-aware applications, and thus, causes inconvenience to the users.

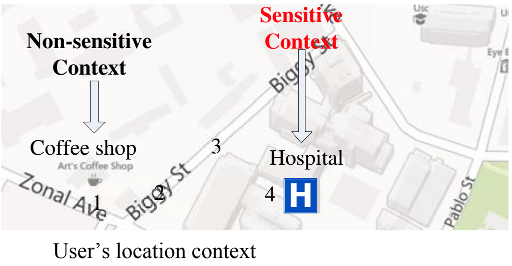


Fig. 1. An illustration on context privacy

To enable smartphone users to enjoy services provided by smartphone applications with privacy protection, many existing privacy preserving approaches have been proposed to explore better tradeoffs between service quality and individual privacy. Most of these approaches focus on location privacy [4]–[7], which, however, fall short when applied to context privacy analysis due to the dynamics of user behaviors and temporal correlations between contexts. Specifically, smartphone users usually transit between different contexts (e.g., a user goes to a particular hospital after eating at a coffee shop), whose sensitivities are different to the users. Moreover, the contexts are usually correlated, which has already been studied for different goals [8]–[10]. Thus, the adversaries can learn the connections between contexts by exploiting the temporal correlations, and then use such correlations to infer user's sensitive contexts based on their observations on non-sensitive contexts. For example, in Figure 1, a context-aware application may learn that a user regularly follows a trajectory $1 \rightarrow 2 \rightarrow 3 \rightarrow 4$. Then, releasing the context information that the user is at the coffee shop at location 1 may reveal that the

[1] GeoNote: http://geonotehelp.blogspot.hk
[2] AutoSilent: http://www.novniv.com

978-1-4799-3360-0/14/$31.00 ©2014 IEEE

user is very likely to go to the hospital, which is sensitive to the user. However, the frameworks on location privacy [4]–[7] do not consider such inference attacks from adversaries knowing temporal correlations, and thus, are not directly applicable to context privacy analysis.

To the best of our knowledge, the only existing work on context privacy protection is *MaskIt* [9], which assumes that adversaries take fixed attacking strategies that do not change over time. This assumption is only valid for offline attacks, e.g., analyzing user's personal information and preferences. However, some adversaries launch real-time attacks [7]. For example, a context-aware application may sell user's sensing data to remote advertisement adversaries, who continuously push context-related ads or spam to users based on the user's instant context information. Note that in the real world, context-based ads or spam need to be delivered in real time (e.g., NAVTEQ or AdLocal by Cirius Technologies) as users may lose interest if the ads do not match current context. In such case, it is highly possible that the adversaries will adapt their attacking strategies based on their observations of previous attacking results and context dynamics.

To satisfy the aforementioned requirements, in this paper, we model the strategic and dynamic competition between a smartphone user and a malicious adversary as a zero-sum stochastic game, where the user preserves context-based service quality and context privacy against strategic adversaries. The user's action is to control the released data granularity of each sensor used by context-aware applications in a long-term defense against the adversary, while the adversary's action is to select which sensing data as the source for attacks. The interactive competition between the user and the adversary are considered to last for a number of stages with the contexts dynamics. The previous context and attacking result are observed by the user and the adversary as system state, based on which both players adjust their future strategies. The user's optimal defense strategy is obtained at a Nash Equilibrium (NE) point of this zero-sum game. An efficient minimax learning algorithm with proved convergence is proposed to obtain the NE point. Compared to traditional learning algorithm, the proposed algorithm reduces the computational cost by reducing the dimensions of state values that need to be updated. We give both analytical results and evaluations on real smartphone traces to analyze the factors that affect the user's optimal defense strategy.

The main contributions of this paper are threefold. First, we identify the context privacy problem in context-aware applications. Specifically, we consider the context correlations and powerful adversaries that are capable of adjusting their attacking strategies over time. Second, we analyze the context privacy problem via a stochastic game formulation, and devise an efficient minimax learning algorithm with provable convergence to obtain optimal strategies. We improve the efficiency of the learning algorithm by solving an equivalent problem with reduced dimensions. We also prove that the algorithm converges to an NE point. Finally, we use real smartphone context traces of 94 users to demonstrate the efficacy and efficiency of the proposed algorithm. Promisingly, the results give guidance to the design of context privacy preserving mechanisms.

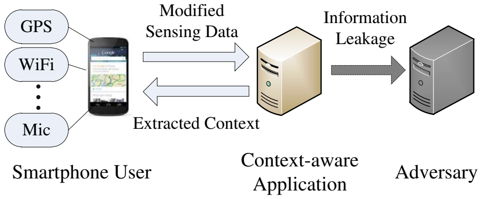


Fig. 2. A mobile phone context sensing system

The rest of the paper is organized as follows. Section II introduces the system model. Section III presents the stochastic game formulation for the context privacy problem. Section IV proposes a minimax learning algorithm to obtain the user's optimal defense strategy. Section V describes the performance evaluations, and Section VI reviews the related works. Finally, Section VII concludes the paper.

II. System Model

In this section, we describe the model of the mobile phone context sensing system and the privacy issue when the context-aware application is untrusted.

Context Sensing Model. Fig. 2 illustrates a mobile phone context sensing system, where a sensor-equipped smartphone runs untrusted context-aware applications. The smartphone senses its environment with multiple sensors and releases the modified sensing data to the context-aware applications periodically for energy-efficiency reasons [8], [9], where a period is referred to as a time slot in the context sensing system. We assume that the smartphone user has installed a privacy-preserving middleware (e.g., *MaskIt* [9]) to control the released data granularity of each sensor. An untrusted application accesses user's data via the privacy-preserving middleware but does not have the permission to access raw sensing data. On the one hand, the untrusted application extracts the user's context using certain context recognition approaches (e.g. [8], [10]). On the other hand, the untrusted application leak the modified sensing data to an adversary.

User Model. A smartphone user can encounter a set of contexts $\mathcal{C} = \{c_1, ..., c_n\}$. We adopt the Markov model to capture the transitions between contexts. It has been shown that human behaviors and activities extracted from smartphone sensors can be modeled well with a two-state Markov chain [11], [12]. At time t, the user's context is denoted as $C_t \in \mathcal{C}$, which is generated from a Markov model M. According to the independence property of Markov chains, we have $\Pr[C^t = c_i|C^1, ..., C^{t-1}] = \Pr[C^t = c_i|C^{t-1}]$. A subset of contexts is considered to be private contexts that are sensitive to the user. The user claims a subset of $\mathcal{C}$ to be sensitive via special applications (e.g., Locaccino [13]). The user's context privacy is breached if the adversary successfully infers that the user is in its sensitive context. To protect context privacy, the user can control the released data granularity of each sensor via the privacy-preserving middleware.

Adversary Model. The adversary is able to obtain the released sensing data at the time when the untrusted application accesses the data, and is assumed to know the Markov chain of a user [9]. The sensing data retrieved by the adversary in a time slot is limited due to computational constraints (caused by curse of dimensionality when using private data [5]) or limited bandwidth used for retrieving data. As the contexts and user's released data granularity vary over time, the adversary can adaptively choose different subsets of sensors to maximize its long-term utility. To protect smartphone users against all kinds of adversaries, we make the worst case assumption: the adversary is a *malicious* attacker that aims at minimizing user's utility through a series of strategic attacks [7].

Problem Statement. Our goal is to find the optimal defense strategy for users to preserve privacy against malicious adversary over a serial of correlated contexts. As the user and the malicious adversary have opposite objectives, their dynamic interactions can be modeled as a zero-sum game. Moreover, since the context is considered to keep changing over time and both the user and the adversary make different actions at different times, the zero-sum game is in a stochastic setting, i.e., the context privacy game should be formulated as a stochastic game.

III. Stochastic Game Formulation

A stochastic game is a dynamic game with probabilistic transitions played in a sequence of stages. A two player stochastic game Γ consists of a six-tuple $< \mathcal{S}, \mathcal{A}^1, \mathcal{A}^2, r^1, r^2, P >$. $\mathcal{S}$ is the discrete state space. $\mathcal{A}^k$ is the action space of player k for $k = 1, 2$. $r^k : \mathcal{S} \times \mathcal{A}^1 \times \mathcal{A}^2 \mapsto \mathbb{R}$ is the stage payoff function for player k. $P : \mathcal{S} \times \mathcal{A}^1 \times \mathcal{A}^2 \mapsto \Delta(\mathcal{S})$ is the transition probability map, where $\Delta(\mathcal{S})$ is the set of probability distributions over $\mathcal{S}$. The game Γ is played in a sequence of stages, where each player k receives a stage payoff $r^k(s, a^1, a^2)$ based on players actions $a^k \in \mathcal{A}^k$ and current stage $s \in \mathcal{S}$. Each player k attempts to maximize its expected sum of discounted payoffs.

In this section, we formulate the privacy game in mobile phone context sensing as a two-player stochastic game.

A. States and Actions

1) System States: In each time slot, the smartphone user is in a certain context and releases data of multiple sensors to the context-aware application. The user's context is included in the system state as the user's action depends on its observation of the current context. Note that the current context is only observable to the user, while the adversary can only infer the context based on the modified sensing data and the user's Markov model. Previous attack results should also be included in the system state. As the adversary's strategy is not known by the user, the user can only conjecture the adversary's strategy from previous attack results, which are assumed to be observable to the user. This assumption is reasonable in context-based applications. For instance, if a user receives an advertisement based on its current private context, then the user knows that the adversary successfully inferred this private context; if the user receives an advertisement based on a context that it has never been to, then the user knows that the adversary has failed to infer its true context. Thus, the user should maintain a record of which contexts the adversary has launched attacks on, and which contexts have been successfully attacked. The attack result observed at time t, namely the attack result in the last time slot, is denoted as Ar^t, whose value can be $Ar^t = 1$ meaning the adversary successfully infers the context C_{t-1}, or $Ar^t = 0$ meaning the adversary fails to infer the context C_{t-1}. In summary, the context and attack result are observable to the user and affect the user's decisions. Thus, the state of the privacy stochastic game at time t is defined by $S^t = \{C^t, Ar^t\}$.

2) User's Actions: After observing the state $S^t = \{C^t, Ar^t\}$ at each stage (note that the adversary can only infer C^t based on M), both the user and the adversary decides their actions for the current stage. As discussed in Section II, the user controls the granularity of the released sensing data to protect its context privacy while preserving the quality of context-based services. For simplicity, we use the accuracy of context recognition to measure the granularity of the sensing data, which is assumed to be the weighted summation of the data granularity of each sensor. Formally, the action of the user at time t is defined as $\mathbf{a}_u^t = \{a_{u,1}^t, ..., a_{u,K}^t\}$, with each sensor's data granularity $a_{u,k}^t \in [0, 1], \forall k = 1, ..., K$, where K is the total number of sensors used for recognition. The accuracy of context recognition g $(0 \leq g \leq 1)$ based on $\mathbf{a}_u^t$ is given by $g = \sum_{k=1}^{K} \kappa_k a_{u,k}^t$, where $\{\kappa_k : \forall k\}$ are the weights measuring the sensitivity of the sensor's data granularity to the context recognition accuracy.

3) Adversary's Actions: On the other hand, due to the limited attacking capability, the adversary needs to select a proper subset of sensing data for retrieval. Mathematically, the adversary's actions at time t are defined as $\mathbf{a}_a^t = \{a_{a,1}^t, ..., a_{a,K}^t\}$, where $a_{a,k}^t$ is the probability of retrieving the data of the kth sensor. The power limitation constraints for the adversary's actions are as follows.

$$\begin{aligned} &\sum_k a_{a,i}^t \leq L, \\ &0 \leq a_{a,i}^t \leq 1, \forall k, \end{aligned} \tag{1}$$

where L is the power limitation of the adversary.

4) State Transitions: It can be seen that the state S^t is uncertain (due to the uncertainty of C^t) and depend on the actions of the user and the adversary (Ar^t depends on the player' actions). We assume that user behavior is independent of player's actions. Then, the state transition probability can be computed by

$$\begin{aligned} \Pr[S^{t+1}|S^t, a_u^t, a_a^t] &= \Pr[Ar^{t+1}|Ar^t, a_u^t, a_a^t] \Pr[C^{t+1}|C^t] \\ &= \Pr[Ar^{t+1}|a_u^t, a_a^t] \Pr[C^{t+1}|C^t]. \end{aligned} \tag{2}$$

The second equality holds because Ar^{t+1} is the attack results observed at time $t+1$, which only depends on the actions players made at the last stage.

B. Stage Payoff

After defining the states and actions, we give a concrete expression of stage payoffs. The payoff function of the user is defined to be the quality of the context-based service with weighted penalty on privacy loss, which is written as

$$r_u(S^t, a_u^t, a_a^t) = QoS(a_u^t) - \omega \cdot Pri(S^t), \tag{3}$$

where $QoS(a_u^t)$ is the quality of context-based service the user enjoys, ω the equivalent service quality improvement caused by unit privacy loss, and $Pri(S^t)$ the privacy loss. $QoS(a_u^t)$ is a measure of the user's degree of satisfaction with the context-based service and can be modeled as a sigmoid function of the context recognition accuracy. Sigmoid function has been widely used to approximate the user's satisfaction with respect to service qualities [14]. Concretely, $QoS(a_u^t)$ is measured as

$$QoS(a_u^t) = \frac{1}{1 + e^{-\theta(a_u^t - \eta)}}, \tag{4}$$

where θ decides the steepness of the quality of service satisfactory curve, η the satisfaction threshold below which the user has very limited satisfaction (the function curve is convex) and above which the user's satisfaction rapidly approaches an asymptotic value (the function curve is concave).

Next, we measure the privacy loss of the user based on the definition of context privacy in [9]. Consider a user over a day with a context space $\mathcal{C}$ and a set of sensitive context $\mathcal{C}_s \subseteq \mathcal{C}$. We say that the released data preserves privacy if the adversary learns little information about the user being in a private state from the released data, meaning that for all sensitive contexts and all times the difference between the posterior and prior beliefs on the user being in a sensitive context at that time is limited. Normally, the adversary values the information of user's recent contexts more highly than the information about user's contexts in the faraway future. Based on the above intuition, we define *context sensitivity* as follows.

Definition 1 (Context Sensitivity) *The sensitivity of a context c is defined to be the sum of the discounted differences between the prior belief and the posterior belief after observing current context on the user being in each sensitive context in the future, that is,*

$$Sens(c) = \sum_{t=0}^{\infty} \sum_{c^s \in \mathcal{C}_s} \gamma^t \left| \Pr[C^t = c_s | C^0 = c] - \Pr[C^t = c_s] \right|, \tag{5}$$

where $0 < \gamma < 1$ is the discount factor of the context privacy.

The sensitivity of a context c measures the maximum information that the adversary can learn about the user's sensitive contexts in the future by observing the user being in c.

Based on the context sensitivity, we define user's privacy loss. If an adversary successfully infers a user's current context, the user's privacy loss is the sensitivity of the current context. Otherwise, the privacy loss is zero, as the user's true context is still unknown to the adversary. Thus, the privacy loss is expressed as

$$Pri(S^t, \mathbf{a}_u^t, \mathbf{a}_a^t) = Sens(C^t) Ar^{t+1}, \tag{6}$$

where Ar^{t+1} is the attack result known at time $t+1$, i.e., the attack result for context C^t. The probability of a successful attack at time t is $\Pr[Ar^{t+1}] = \sum_i \kappa_i a_{u,i}^t a_{a,i}^t$.

Then, we decide ω, i.e., the equivalent service quality improvement caused by unit privacy loss. For each context, we measure service quality improvement and privacy loss when the adversary can access all user's raw sensing data, compared with the case that the adversary knows nothing. We assume that the adversary has prior belief of a user's context based on its background knowledge (e.g., the adversary knows the user's behavior pattern or the Markov chain of the user's contexts). Therefore, we express ω as (7), where $\Pr[C^t = c]$ is the adversary's prior belief on user's context. Substituting (4) (5) (6) (7) back into (3), we can obtain the stage payoff for the user, while the stage payoff for the adversary is the negative of (3).

Generally, context applications run continuously on a smartphone all day long [8], [9]. Thus, we assume that there is an infinite number of time slots, i.e., the context privacy stochastic game is played for an infinite number of stages. Normally, the smartphone users care more about the current context or near future contexts than the faraway future contexts. For example, a user's current context is more private since the adversaries can cause immediate damage to the user. Therefore, the user's utility is to the expected sum of discounted stage payoffs, where the delayed payoffs value less to the user

$$U_u = \mathbb{E}[\sum_{t=0}^{\infty} \gamma^t r_u(S^t, \mathbf{a}_u^t, \mathbf{a}_a^t)], \tag{8}$$

where γ is the discount factor of the context privacy. Then, the user's objective is to derive an *optimal defense strategy* that maximizes U_u, which is discussed in the following section.

IV. Learning the Optimal Defense Strategy

Based on the context privacy stochastic game formulation in Section III, we will discuss the algorithm to derive the NE of the stochastic game, so as to obtain the optimal defense strategy of the user.

A. Minimax Equilibrium in the Context Privacy Game

Formally, a stratefy in a stochastic game is defined to be a probability distribution over the action set at any state. A strategy π is said to be *stationary* if $\pi^t = \pi$ for all t, that is, the strategy is fixed over time. In this paper, we are interested in stationary policies. In the context privacy stochastic game, the user's strategy is denoted by $\pi_u : \mathcal{S} \mapsto \Delta(\mathcal{A}_u)$ and the adversary's strategy is denoted by $\pi_a : \mathcal{S} \mapsto \Delta(\mathcal{A}_a)$, where $\mathcal{S}$ is the state space for S^t, $\Delta(\mathcal{A}_u)$ and $\Delta(\mathcal{A}_a)$ the probability distributions over the user's action space $\mathcal{A}_u$ and the adversary's action space $\mathcal{A}_a$, respectively.

In stochastic games, utilities are expressed in the form of *state value*. Here, the initial state is defined to be the state at

$$\omega = \frac{\sum_{c:c\in\mathcal{C}} \Pr[C^t = c](QoS(a_u = \mathbf{1}) - QoS(\mathbf{a}_u^t = \mathbf{0}))}{\sum_{c:c\in\mathcal{C}} \Pr[C^t = c](Pri(\mathbf{a}_u^t = \mathbf{1}, \mathbf{a}_a^t = \mathbf{1}, C^t = c) - Pri(\mathbf{a}_u^t = \mathbf{0}, \mathbf{a}_a^t = \mathbf{1}, C^t = c))}, \tag{7}$$

time $t = 0$, denoted by S^0. Given policies π_u, π_a and a state $s \in \mathcal{S}$, the user's utility can be written as

$$V^\pi(s) = \sum_{t=0}^{\infty} \gamma^t \mathbb{E}[r_u(S^t, \mathbf{a}_u^t, \mathbf{a}_a^t) | \pi_u, \pi_a, S^0 = s]. \tag{9}$$

Denote the actions $\mathbf{a}_u^t, \mathbf{a}_a^t$ determined by policies π_u, π_a to be $\mathbf{a}_u^\pi, \mathbf{a}_a^\pi$, respectively. Then, we can rewrite (9) as

$$V^\pi(s) = r_u(s, \mathbf{a}_u^\pi, \mathbf{a}_a^\pi) + \gamma \sum_{s' \neq s} \Pr[s' | s, \mathbf{a}_u^\pi, \mathbf{a}_a^\pi] V^\pi(s'). \tag{10}$$

Both user and adversary follow their optimal policies $\{\pi_u^*, \pi_a^*\}$ that maximize their own utilities, where the optimal policies are called an *optimal strategy pair* $\pi^* = \{\pi_u^*, \pi_a^*\}$. An optimal strategy pair in a stochastic game are the policies at a Nash equilibrium point, which is defined as follows.

Definition 2 (NE in Stochastic Game) *In a zero-sum stochastic game* Γ*, a Nash Equilibrium (NE) point is an optimal strategy pair* $\pi^* = \{\pi_u^*, \pi_a^*\}$*, such that for all state* $s \in \mathcal{S}$

$$V^{\pi^*}(s) \geq V^{\pi^a}(s), \tag{11}$$

and

$$V^{\pi^*}(s) \leq V^{\pi^u}(s), \tag{12}$$

where $\pi^a = \{\pi_u, \pi_a^*\}, \forall \pi_u$*, and* $\pi^u = \{\pi_u^*, \pi_a\}, \forall \pi_a$*.*

In the context privacy stochastic game, the user aims to find the minimax equilibria, where the user tries to determine an optimal strategy π_u^* that maximizes $\{V^\pi(s) : \forall s\}$, while the adversary tries to find an optimal strategy π_u^* that minimizes $\{V^\pi(s) : \forall s\}$. Thus, based on (10), we have

$$\begin{aligned} V^{\pi^*}(s) = \max_{\pi_u} \min_{\pi_a} \Big\{ & r_u(s, \mathbf{a}_u^\pi, \mathbf{a}_a^\pi) \\ & + \gamma \sum_{s' \neq s} \Pr[s' | s, \mathbf{a}_u^\pi, \mathbf{a}_a^\pi] V^{\pi^*}(s') \Big\}, \end{aligned} \tag{13}$$

where $V^{\pi^*}(s)$ is referred to as the value of state s.

It has been shown [15] that the equilibrium in a zero-sum stochastic game is the unique minimax equilibrium, and thus the optimal strategy pair in the context privacy game is unique.

B. Equivalent State Value

Based on (13), the optimal strategy pair can be derived via existing *reinforcement learning* algorithms, e.g. minimax-Q learning [16]. However, since cardinality of $\mathcal{S}$ could be very large, the complexity of deriving π^* according to (13) would be very high. For example, the minimax-Q learning needs to solve $|\mathcal{S}|$ bimatrix games, where $|\mathcal{S}|$ is the cardinality of $\mathcal{S}$. In order to reduce the computational complexity, we solve an equivalent problem instead.

The equivalent state value $\tilde{V}_u^{\pi^*}(Ar)$ is defined to be the expected state value over the context variable, i.e., $\tilde{V}_u^{\pi^*}(Ar) = \mathbb{E}_c[V_u^{\pi^*}(s)]$ where $s = \{Ar, c\}$. Then, we have the following observation.

Lemma 1 *The equivalent state value* $\tilde{V}_u^{\pi^*}(Ar)$ *can be derived from* (13) *and enjoys an expression where context* c *is eliminated, i.e.,*

$$\tilde{V}^{\pi^*}(Ar) = \mathbb{E}_c \left[r_u(s, \mathbf{a}^{\pi^*}) + \gamma \sum_{Ar'} \left(\Pr[Ar' | \mathbf{a}^{\pi^*}] \tilde{V}^{\pi^*}(Ar') \right) \right], \tag{14}$$

where $\mathbf{a}^{\pi^*} = \{\mathbf{a}_u^{\pi^*}, \mathbf{a}_a^{\pi^*}\}$ *is the action pair following the optimal strategy pair* π^**.*

Proof: By taking expectation over c on both sides of (13), we have

$$\begin{aligned} & \tilde{V}^{\pi^*}(Ar) \\ & = \mathbb{E}_c \left[r_u(s, \mathbf{a}^{\pi^*}) + \gamma \sum_{s' \neq s} \Pr[s' | s, \mathbf{a}^{\pi^*}] V^{\pi^*}(s') \right] \\ & = \gamma \sum_{Ar', c'} \mathbb{E}_c \left[r_u(s, \mathbf{a}^{\pi^*}) + \Pr[Ar' | \mathbf{a}^{\pi^*}] \Pr[c' | c] V_u^{\pi^*}(Ar', c') \right] \\ & = \mathbb{E}_c \left[r_u(s, \mathbf{a}^{\pi^*}) \right] \\ & \qquad + \gamma \sum_{Ar', c', c} \left(\Pr[Ar' | \mathbf{a}^{\pi^*}] \Pr[c' | c] \Pr[c] V_u^{\pi^*}(Ar', c') \right) \\ & = \mathbb{E}_c \left[r_u(s, \mathbf{a}^{\pi^*}) \right] + \gamma \sum_{Ar', c'} \left(\Pr[Ar' | \mathbf{a}^{\pi^*}] \Pr[c'] V_u^{\pi^*}(Ar', c') \right) \\ & = \mathbb{E}_c \left[r_u(s, \mathbf{a}^{\pi^*}) + \gamma \sum_{Ar'} \left(\Pr[Ar' | \mathbf{a}^{\pi^*}] \tilde{V}^{\pi^*}(Ar') \right) \right]. \end{aligned} \tag{15}$$

This completes the proof. ∎

We can see that $\tilde{V}_u^{\pi^*}(Ar)$ largely reduces the number of state values from $|\mathcal{S}|$ or $2|\mathcal{C}|$ to 2 (since Ar is a binary variable). The following theorem proves that we can derive the optimal action pair from $\tilde{V}_u^{\pi^*}(Ar)$.

Theorem 1 *The optimal strategy pair* $\pi^* = \{\pi_u^*, \pi_a^*\}$ *for the context privacy stochastic game* (13) *can be obtained by solving the following equivalent problem*

$$\begin{aligned} \pi^* = \arg \max_{\pi_u} \min_{\pi_a} \Big\{ & r_u(s, \mathbf{a}^{\pi^*}) \\ & + \gamma \sum_{Ar'} \left(\Pr[Ar' | \mathbf{a}^{\pi^*}] \tilde{V}^{\pi^*}(Ar') \right) \Big\}, \end{aligned} \tag{16}$$

Proof: By standard Markov decision process (MDP) techniques [17], [18], the problem (13) can be expressed as an equivalent MDP $\min_{\pi_a} \max_{\pi_u} \mathbb{E}_c[V_u^\pi(s)]$ with the state

Algorithm 1 Minimax Learning Algorithm

```
Input: The context privacy stochastic game $\Gamma$
Output: $\pi^*$
    // 1. initialization
    $t \leftarrow 0, Ar^t = 0$;
    $\tilde{V}^t(Ar = 0) \leftarrow 1, \tilde{V}^t(Ar = 1) \leftarrow 1$;
    Initialize strategy pair $\pi^t$: two uniform distributions where $a^t_{u,i} = \frac{1}{K}, a^t_{a,i} = \frac{L}{K}, \forall i$;
    // 2. iteration
    repeat
        Select an action pair $\{\mathbf{a}^t_u, \mathbf{a}^t_a\}$ based on $\pi^t$;
        Update $Ar^{t+1}$ after both players take their actions $\{\mathbf{a}^t_u, \mathbf{a}^t_a\}$;
        Update equivalent state value $\tilde{V}^{t+1}(Ar)$ according to (18);
        Update optimal strategy $\pi^{t+1}$ according to (16) with updated state values;
        $t \leftarrow t + 1$;
    until Converge
```

space $\mathcal{S}$, the action space $\{\{\mathbf{a}_u\}, \{\mathbf{a}_a\}\}$, the transition kernel $\Pr[Ar'|\mathbf{a}^{\pi^*}] = \mathbb{E}_c\left[\Pr[s'|s]\right]$, and the stage payoff function $\mathbb{E}_c\left[r_u(s, \mathbf{a}^{\pi^*})\right]$. It is known that the optimal strategy pair π^* can be obtained by solving

$$\begin{aligned}\min_{\pi_a}\max_{\pi_u}\mathbb{E}_c[V^\pi(s)] = \mathbb{E}_c\Bigg[\min_{\pi_a}\max_{\pi_u}\Big\{r_u(s, \mathbf{a}^{\pi^*})\\ +\gamma\sum_{Ar'}\left(\Pr[Ar'|\mathbf{a}^{\pi^*}]\tilde{V}^{\pi^*}(Ar')\right)\Big\}\Bigg],\end{aligned} \tag{17}$$

which completes the proof. ■

C. A Minimax Learning Algorithm

According to Theorem 1, we can derive π^* by learning $\tilde{V}^{\pi^*}_u(Ar)$, which can be obtained by the following update rule, which is modified from Q-learning [19].

$$\begin{aligned}\tilde{V}^{t+1}(Ar) =&(1 - \alpha^{t+1})\tilde{V}^t(Ar)\\ &+ \alpha^{t+1}\mathbb{E}_c\left[r_u(s, \mathbf{a}^t_u, \mathbf{a}^t_a) + \gamma\tilde{V}^t(Ar')\right],\end{aligned} \tag{18}$$

where $\alpha^t \in [0, 1)$ is the learning rate, which needs to decay over time in order for the learning algorithm to converge. In this paper, we set $\alpha^t = \frac{1}{t}$. $\tilde{V}^{t+1}_u(Ar)$ is used as an approximate of $\tilde{V}^{\pi^*}_u(Ar)$ and iteratively updates according to (18) until converges.

Then, the learning algorithm for equivalent state value $\tilde{V}^{\pi^*}_u(Ar)$ is described in Algorithm 1. First we initialize equivalent state values to be 1, and the strategy of each player to be uniform distribution. Then, we iteratively update equivalent state values and strategy pair according to (18) and (16), respectively, until the strategy pair approaches the optimal strategy pair. It is provable that the iteratively updated π^t converges to the optimal strategy pair. Due to page limitation, we omit the detailed proof.

V. Evaluation

In this section, we conduct trace-driven simulations to evaluate the smartphone user's payoffs under the privacy attack. First, we show the proposed algorithm largely improves the convergence speed compared with the traditional learning algorithm. Then, we demonstrate the effectiveness of the proposed algorithm by comparing the sum of discounted payoffs when the user adopts different strategies. We also study how the user's utility and strategies are affected by some system parameters.

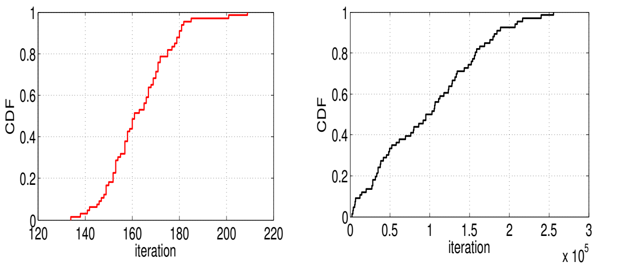
A. Setup

The user model, system parameters, and baselines used for evaluation are described as follows.

- **User Model**. We evaluate the performance of our proposed algorithm using the Reality Mining dataset [3], which was collected by the MIT Media Laboratory from September 2004 to June 2005 [20]. The Reality Mining dataset records the continuous activities of 94 students and staff at MIT equipped with Nokia 6600 smartphones, which are pre-installed with several pieces of software that collects data about call logs, Bluetooth devices in proximity of approximately five meters, location at granularity of cell tower, application usage, transportation model (e.g., driving, walking, stationary), etc. The total length of all subjects' traces combined is 266,200 hours, with average, minimum, and maximum length being 122 days, 30 days, and 269 days, respectively. As location is the most complete and fine-grained context in the dataset, we select location traces as the user's contexts in our evaluation. The average, minimum, and maximum numbers of locations per user is 19, 7, and 40, respectively. Based on the location traces, we train a Markov chain for each user. Then, we simulate user's behaviors based on the trained Markov chain. For each user, a certain percentage p of contexts are selected as sensitive contexts.
- **System Parameters**. Unless explicitly otherwise stated, we use the following system parameters in our simulations. For each user, the the percentage of sensitive contexts p is set to 0.5, satisfaction threshold eta set to 0.7, QoS steepness θ set to 10, the discount factor γ set to 0.8. According to [8], there are three sensors (i.e., GPS, WiFi, and Bluetooth) used to identify user's location contexts. Thus, we set the number of sensors needed to identify the context to 3, and the power limitation of the adversary L is set to 2. The weights of sensors $\{\kappa_i : i = 1, ..., K\}$ are set to the normalized values drawing from a uniform distribution.
- **Baselines**. We compare the convergence speed of the proposed algorithm and that of the traditional learning algorithm that learns state values directly according to (13). We also compare the performance of users adopting different strategies. We compare the optimal policies obtained by the proposed algorithm (denoted by *proposed*)

[3] http://realitycommons.media.mit.edu/realitymining.html

with *fixed* strategy and *myopic* strategy. The fixed strategy draws an action that uniformly sets the granularity of each sensor to $\frac{1}{K}$. And the myopic strategy is the optimal strategy obtained by myopic learning, where the effects of current actions on the future payoffs are ignored, i.e., γ is considered to be 0 in the myopic learning.



(a) Convergence speed of the proposed algorithm

(b) Convergence speed of the traditional learning algorithm

Fig. 3. CDF of convergence speed

B. Results

1) Convergence Speed: We first show the convergence speed improved by the proposed algorithm in Figure 3. Figure 3 depicts the cumulative distribution function (CDF) of iterations needed to learn the optimal policies for all users in the Reality Mining dataset. We can see that the convergence speed of the proposed algorithm for all users are less than 220 iterations, while for more than half of the users, the convergence speed of the traditional algorithm are more than 10^5 iterations, which demonstrates that the proposed algorithm largely improves the convergence speed compared with traditional learning algorithm. The improvement of the proposed algorithm comes from the smaller cardinality of the equivalent state value, which eliminate the context dimension in the learning process.

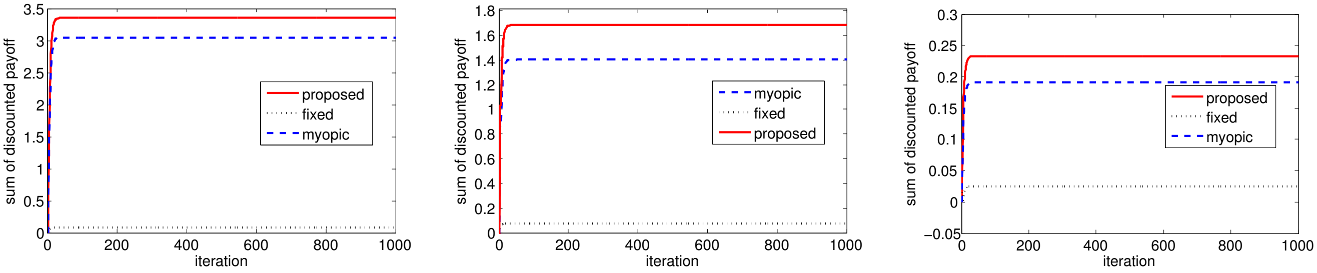
2) Comparison of Different Strategies: Figure 4 compares the performance of the smartphone user when it adopts different strategies to evaluate the proposed context privacy stochastic game and the proposed algorithm. It is assumed that the adversaries use their optimal stationary strategy learned by the minimax algorithm. As shown in Figure 4, the proposed and the myopic strategies achieve higher sum of discounted payoff than the fixed strategy against the adversaries with different power limitations, since the former two strategies maximize the worst-case performance, while the fixed strategy takes actions without considering the adversary's action. Moreover, the proposed strategy achieves highest sum of discounted payoff. This is because the proposed strategy also takes the future payoff into consideration when optimizing the current strategy. Therefore, when smartphone users are under attack from adversaries that are capable of dynamically changing their strategies, the best choice is to adopt the strategy learned from the proposed algorithm that considers future payoff and the dynamics of the adversaries.

Moreover, comparing Figure 4(a), Figure 4(b), and Figure 4(c), we can see that sum of discounted payoff achieved by the proposed strategy goes down as the power limitation of the adversaries L increases. This is because as L increases, the adversaries are able to access more sensing data, it is more likely for the adversaries to successfully attack the user. In such situation, the user may take more conservative actions (i.e., releasing data with less granularity), which results in lower service quality, or the user take the same action to preserve service quality, which, however, causes more privacy loss. As such, either case leads to lower payoff.

3) Impacts of System Parameters: In the following, we show how the percentage of sensitive contexts and satisfaction threshold affect the sum of discounted payoff, and we also depict the optimal policies in different contexts. These evaluation results can provide some guidance in the design of the context privacy preserving schemes.

The average sums of discounted payoff of all users are reported in Figure 5 and Figure 6. From Figure 5, we can see that the sums of discounted payoff achieved by the proposed and myopic strategies get lower as the percentage of sensitive contexts increases, since it would cost more privacy loss to release the same amount of data when the users have more sensitive contexts. The sum of discounted payoff achieved by the fixed strategy stays relatively the same over different percentage of sensitive contexts, because the service quality is invariant and dominates the payoff when adopting the fixed strategy. Moreover, the gap between the sums of discounted payoff obtained by adopting the proposed and myopic strategies approaches to zero when the percentage of sensitive contexts goes down. This is because the consideration of future payoff only affect the weights of privacy loss in the sum of discounted payoff, and both strategies pay more attention to the service quality part when there are fewer sensitive contexts, which reduces the difference between with (the proposed strategy) and without (the myopic strategy) consideration of future payoff. This observation can provide some guidance for the context privacy preserving schemes that for the users with a small faction of sensitive contexts, the impact of current actions on the future payoff can be neglected so as to design more efficient algorithm.

Figure 6 reports the sums of discounted payoff achieved by different strategies over the applications with different satisfaction threshold η. It can be seen that for the applications with higher satisfaction threshold, the sums of discounted payoff achieved by all strategies go down, since the service quality is lower as satisfaction threshold increases. We can also see that when the satisfaction threshold is very low, say 0.2, the differences in the sums of discounted payoff are achieved by different strategies are quite small. It can be seen according to (3) (4) (6) that with low satisfaction threshold, high service quality is easily achieved with only slight privacy loss by contributing a small amount of data, which are the cases of adopting the proposed and the myopic strategies. As such, the service quality dominates the payoff and stays relatively the same over different strategies. Thus, the privacy leaked by the applications that require high accuracy is hard to preserve, and the privacy preserving schemes need to be



(a) $L = 1$ (b) $L = 2$ (c) $L = 3$

Fig. 4. Sum of discounted payoff of different strategies

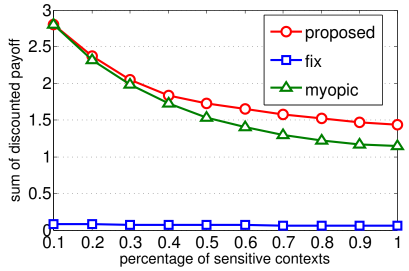


Fig. 5. Sum of discounted payoff vs. percentage of sensitive context

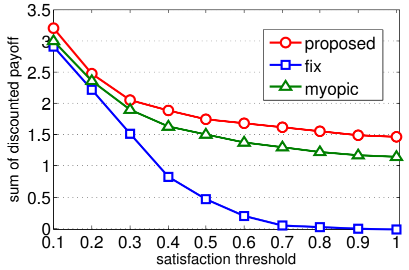


Fig. 6. Sum of discounted payoff vs. satisfaction threshold

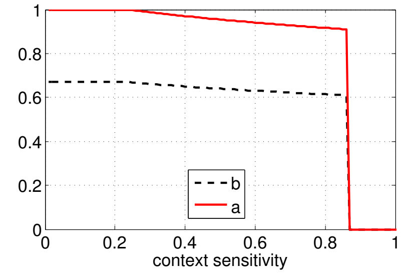


Fig. 7. Optimal policies in contexts of different sensitivities

carefully designed to find a good tradeoff between privacy and utility since different strategies have significant impact on the user's total payoff.

Next, we study the optimal strategy in contexts with different sensitivities. To control the value of context sensitivity, we use the average values of the state values of all users as the state values. We denote the total amount of released data $a = \sum_i \kappa_i a^t_{u,i}$ and the amount of information leaked to the adversary $b = \sum_i \kappa_i a^t_{u,i} a^t_{a,i}$, which represent the optimal strategies. Figure 7 depicts the variance of optimal a, b obtained by the proposed algorithm when the users are in different contexts. It can be seen that when the sensitivity of current context is smaller than 0.25 or larger than 0.87, the optimal a equals to 1 or 0, respectively. In such cases, either the variance of the service quality or the variance of the privacy loss dominates. While when the context sensitivity falls between 0.25 and 0.87, a and b slightly go down with the increment of the context sensitivity, since the users choose a more conservative strategy (releasing less data) as the privacy values more to the users. An interesting observation is that a stays larger than the satisfaction threshold (set to 0.7 by default) when the context sensitivity falls between 0.25 and 0.87. The reason is that below the satisfaction threshold the user enjoys only very limited service quality. Therefore, it is very important to identify the satisfaction threshold when designing the privacy preserving schemes.

VI. Related Work

Numerous techniques have been proposed for preserving privacy in LBSs and participatory sensing on mobile phone. Spatial cloaking and anonymization are widely adopted [4], [5], [21], [22], where a value provided by a user is indistinguishable from those of $k-1$ other users to provide privacy guarantee, known as *k-anonymity*. [4] devises a framework which provides k-anonymity with different context-sensitive personalized privacy requirements. Several clique-cloak algorithms are proposed in [4] to implement the framework by constructing a constraint graph. In [5], locality-sensitive hashing is utilized to partition user locations into groups that contain at least k users. A form of generalization based on the division of a geographic area is adopted by *Anonysense* [21], where a map of wireless LAN access points is partitioned. *KIPDA* [22] enables k-anonymity for data aggregation with a maximum or minimum aggregation function in wireless sensor networks. However, these privacy techniques focus on the single shot scenario, which do not protect user's privacy against adversaries knowing temporal correlations.

Differential privacy has been considered as a major axis in data publishing. Publishing different types of data has been studied, such as histogram [23], [24], set-valued data [25] and decision trees [26]. Among these studies, the data type related to our work is histogram. Blum et al. [23] divides the input counts into bins of roughly the same count to construct a one-dimensional histogram. By observing that the accuracy of a differential privacy compliant histogram depends heavily on its structure, Xu et al. [24] propose two algorithms with different priorities for information loss and noise scales. However, these techniques focus on data modifications but do not environmental dynamics and adversaries' adjustable strategies.

Another category preserves privacy via cryptographic techniques. Girao et al. [27] aggregate data based on homomorphic encryption, which preserves privacy by performing certain computations on ciphertext. The limitation of homomorphic encryption is that a server must know all the users that have reported data to compute the final aggregated results. Secure information aggregation frameworks are proposed in [28]. However, the cryptographic techniques fail to cope with context privacy since the adversaries can decode the true sensing data by compromising context-aware applications.

The only existing work that studies the context privacy issue is MaskIt [9]. MaskIt is a middleware that employs a privacy check to decide whether to release or suppress the current user context. As such, MaskIt limits the adversaries from knowing the user being in a sensitive context even the adversaries have knowledge about the temporal correlation between user's contexts. Nevertheless, MaskIt does not consider the adversaries' capability to adjust their attacking strategies.

Several game theoretic analyses on location privacy have also been discussed. Freudiger et al. [6] study the problem of selfishness in location privacy schemes based on pseudonym changes, and analyze the non-cooperative behavior of mobile nodes with an n-player complete information game. Shokri et al. [7] formulate the location privacy problem as Stackelberg Bayesian games with the consideration of user's service quality and adversary's cost. However, these location privacy problems are quite different from the context privacy discussed in this paper, where the stochastic dynamics and temporal correlation of user's behaviors and environments are considered.

VII. Conclusion

This paper studied the privacy problem of context-aware applications on smartphones. Considering the distinct features of the context privacy problem including the context dynamics and powerful adversaries with knowledge of temporal correlations between contexts and capabilities of adjusting their attacking strategies, we formulate the interactive competition between users and adversaries as a zero-sum stochastic game. To obtain the user's optimal defense strategy efficiently, we propose a minimax learning algorithm to solve an equivalent problem with reduced dimensions. Evaluations on real smartphone traces demonstrate the efficacy of the optimal defense strategy obtained by the proposed algorithm. The proposed stochastic game framework and evaluation results can provide some guidelines for the design of privacy preserving mechanisms for context privacy protection.

Acknowledgement

The research was supported in part by grants from 973 project 2013CB329006, China NSFC under Grant 61173156, RGC under the contracts CERG 622410, 622613 and HKUST6/CRF/12R, as well as the grant from Huawei-HKUST joint lab.

References

[1] Microsoft, "Location based services usage and perceptions survey," Apr. 2011. [Online]. Available: http://www.microsoft.com/en-hk/download/details.aspx?id=3250

[2] D. Weitzner, "Obama administration calls for a consumer privacy bill of rights for the digital age," Feb. 2012. [Online]. Available: http://www.whitehouse.gov/blog/2012/02/23/we-can-t-wait-obama-administration-calls-consumer-privacy-bill-rights-digital-age

[3] W. Enck and et al., "Taintdroid: an information-flow tracking system for realtime privacy monitoring on smartphones," in *Proc. OSDI*, Oct. 2010, pp. 1–6.

[4] B. Gedik and L. Liu, "Location privacy in mobile systems: A personalized anonymization model," in *Proc. IEEE ICDCS*, Jun. 2005.

[5] K. Vu, R. Zheng, and J. Gao, "Efficient algorithms for k-anonymous location privacy in participatory sensing," in *Proc. IEEE INFOCOM*, Mar. 2012.

[6] J. Freudiger, M. Manshaei, J. Hubaux, and D. Parkes, "On non-cooperative location privacy: a game-theoretic analysis," in *Proc. ACM CCS*, Nov. 2009, pp. 324–337.

[7] R. Shokri, G. Theodorakopoulos, C. Troncoso, J. Hubaux, and J. Le Boudec, "Protecting location privacy: Optimal strategy against localization attacks," in *Proc. ACM CCS*, Oct. 2012, pp. 617–627.

[8] S. Nath, "ACE: exploiting correlation for energy-efficient and continuous context sensing," in *Proc. ACM MobiSys*, 2012.

[9] M. Götz, S. Nath, and J. Gehrke, "MaskIt: Privately releasing user context streams for personalized mobile applications," in *Proc. ACM SIGMOD*, May 2012, pp. 289–300.

[10] A. Parate, M.-C. Chiu, D. Ganesan, and B. M. Marlin, "Leveraging graphical models to improve accuracy and reduce privacy risks of mobile sensing," in *Proc. ACM MobiSys*, 2013.

[11] E. Kim, S. Helal, and D. Cook, "Human activity recognition and pattern discovery," *IEEE Perv. Comp.*, vol. 9, no. 1, pp. 48–53, 2010.

[12] A. Mannini and A. Sabatini, "Accelerometry-based classification of human activities using markov modeling," *Computational Intelligence and Neuroscience*, 2011.

[13] E. Toch and et al., "Empirical models of privacy in location sharing," in *Proc. ACM Ubicomp*, Sep. 2010, pp. 129–138.

[14] H. Lin, M. Chatterjee, S. Das, and K. Basu, "Arc: an integrated admission and rate control framework for competitive wireless cdma data networks using noncooperative games," *IEEE Trans. Mobile Comput.*, vol. 4, no. 3, pp. 243–258, 2005.

[15] B. Wang, Y. Wu, K. Liu, and T. Clancy, "An anti-jamming stochastic game for cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 4, pp. 877–889, 2011.

[16] M. Littman, "Markov games as a framework for multi-agent reinforcement learning," in *Proc. ICML*, Jul. 1994, pp. 157–163.

[17] D. Bertsekas, *Dynamic programming and optimal control*. Athena Scientific, 1995.

[18] X. Cao, *Stochastic learning and optimization: a sensitivity-based approach*. New Yorks Springer, 2007.

[19] J. Hu, M. Wellman *et al.*, "Multiagent reinforcement learning: Theoretical framework and an algorithm," in *Proc. ICML*, 1998, p. 242C250.

[20] N. Eagle, A. S. Pentland, and D. Lazer, "Inferring friendship network structure by using mobile phone data," *Proceedings of the National Academy of Sciences (PNAS)*, vol. 106, no. 36, pp. 15 274–15 278, 2009.

[21] M. Shin, C. Cornelius, D. Peebles, A. Kapadia, D. Kotz, and N. Triandopoulos, "Anonysense: A system for anonymous opportunistic sensing," *J. Perv. Mobile Comput.*, vol. 7, no. 1, pp. 16–30, 2011.

[22] M. Groat, W. He, and S. Forrest, "KIPDA: k-indistinguishable privacy-preserving data aggregation in wireless sensor networks," in *Proc. IEEE INFOCOM*, Apr. 2011.

[23] A. Blum, K. Ligett, and A. Roth, "A learning theory approach to non-interactive database privacy," in *Proc. ACM STOC*, May 2008.

[24] J. Xu, Z. Zhang, X. Xiao, Y. Yang, and G. Yu, "Differentially private histogram publication," in *Proc. IEEE ICDE*, Apr. 2012.

[25] R. Chen, N. Mohammed, B. Fung, B. Desai, and L. Xiong, "Publishing set-valued data via differential privacy," *in VLDB*, vol. 4, no. 11, 2011.

[26] A. Friedman and A. Schuster, "Data mining with differential privacy," in *Proc. ACM SIGKDD*, Jul. 2010.

[27] J. Girao, D. Westhoff, and M. Schneider, "Cda: Concealed data aggregation for reverse multicast traffic in wireless sensor networks," in *Proc. IEEE ICC*, May 2005.

[28] B. Przydatek, D. Song, and A. Perrig, "Sia: Secure information aggregation in sensor networks," in *Proc. ACM SenSys*, Nov. 2003.