# The Sharing Game: Benefits and Privacy Implications of (Co)-Location Sharing with Interdependences

ALEXANDRA-MIHAELA OLTEANU, School of Computer and Communication Sciences (IC), EPFL, Switzerland
KEVIN HUGUENIN, Faculty of Business and Economics (HEC), UNIL, Switzerland
MATHIAS HUMBERT, Swiss Data Science Center, ETH Zurich and EPFL, Switzerland
JEAN-PIERRE HUBAUX, School of Computer and Communication Sciences (IC), EPFL, Switzerland

Most popular location-based social networks, such as Facebook and Foursquare, let their (mobile) users post location and co-location (involving other users) information. Such posts bring social benefits to the users who post them but also to their friends who view them. Yet, they also represent a severe threat to the users' privacy, as co-location information introduces interdependences between users. We propose the first game-theoretic framework to analyze and predict the strategic behaviors, in terms of information sharing, of users of OSNs. In addition, in order to design parametric utility functions that are representative of the users' actual preferences, we conduct a survey of 250 Facebook users and use conjoint analysis to quantify the users' benefits of sharing vs. viewing (co)-location information and their preference for privacy vs. benefits. We evaluate our framework through data-driven numerical simulations. We show how users' individual preferences influence each other's decisions, we determine several factors that significantly affect these decisions (specifically the considered adversary and the relative preference for privacy vs. benefits) for a better understanding of the users' behaviors and the interdependent privacy risks. Our findings are instrumental in the design of next-generation privacy protection systems.

Additional Key Words and Phrases: Location-based social networks; privacy; utility; game theory

**ACM Reference Format:**
Alexandra-Mihaela Olteanu, Kevin Huguenin, Mathias Humbert, and Jean-Pierre Hubaux. 0. The Sharing Game: Benefits and Privacy Implications of (Co)-Location Sharing with Interdependences. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 0, 0, Article 0 ( 0), 27 pages.
https://doi.org/0000001.0000001

## 1 INTRODUCTION

With the advent of mobile networking, mobile users can easily connect to the Internet and determine their actual locations with their smartphones, while on the go. Major online social network (OSN) providers, such as Facebook, understood early on the interest users have in sharing their location jointly with their posts, pictures, etc. This location-sharing feature has gained even more momentum as users increasingly access their favorite OSNs from their smartphones (most Facebook check-ins and photos are made from mobile devices). Another popular feature, currently implemented in many mobile location-based social networks, is the ability to mention other users, such as friends, in posts or to tag them on pictures. Ilia et al. [20] perform a user study that demonstrates that 84.7% of


2474-9567/0/0-ART0 $15.00
https://doi.org/0000001.0000001

posted pictures contain one or more face(s), whereas 87% contain one tag (users do not typically tag themselves) and 12.2% contain more than one tag. In many cases, such information indicates that the users mentioned in a post are co-located. As for location information, sharing co-location information–the fact that two users are together (the actual location might not be known)–brings social benefits (as also pointed out by Krasnova et al.[25]) to those sharing it but also to their friends who view it: Users enjoy knowing with whom their friends are and telling their friends with whom they are.

However, these features also raise privacy concerns. Although it has been known for years that location information leads to severe privacy issues–this has been extensively studied in the literature (e.g., [12, 26, 33]; see also FindYou [40], a location privacy auditing tool, available at https://find-you.herokuapp.com/), it was only recently that the effect of co-location information on users' location privacy was studied [34]. A critical aspect of co-locations is that they relate to all the involved users (such information is co-owned by the involved users [15, 45]) and introduce interdependences between the users' location privacy, as the location information disclosed by users affects the privacy of their friends. As such, users lose partial control over their privacy and it becomes complex to predict the optimal sharing behavior. Such interdependent privacy risks are quite problematic if users have different, possibly opposite, views about sharing and privacy: it creates so-called multi-party privacy conflicts [45, 46].

Awareness about the interdependent nature of privacy is increasing, yet this is not explicitly addressed by current laws, because of its complexity. Although opinion 5/2009 on online social networking produced by the Working Party on Data Protection – which is an advisory board set up by the EU for the reform of the data protection laws – raises awareness about the case of users uploading data about other individuals, even in the General Data Protection Regulation (GDPR) (Regulation EU 2016/679), recently adopted by the EU Parliament, the case where individuals share data about individuals online is not directly mentioned, and the problem remains unsolved. Therefore, from a legal perspective, there are little regulations that apply to sharing on OSNs (except for the extreme case of sharing sexually explicit content) and this important problem deserves further study.

We propose the first unified framework for modeling the direct and indirect benefits, and the privacy implications of location and co-location sharing, in addition to the resulting strategic behaviors of the users. Such a framework enables us to analyze and predict the behavior of users regarding location and co-location sharing on OSNs. To this end, we build our framework using two well-established modeling and analytical tools: game theory [14, 32, 48] and conjoint analysis [16]. Game theory enables us to model and formalize the users' sharing rationale and behavior. Such models include a number of parameters, typically in the expression of the users' utility, that characterize the users' behaviors. Conjoint analysis enables us to rigorously quantify, based on a personalized user survey, the relative benefits of sharing and viewing location and co-location information, and the associated relative costs in terms of location privacy. The values obtained through conjoint analysis are used to derive the different parameters of the game-theoretic model. Although several works [7, 36] have investigated interdependent privacy risks from a game-theoretic perspective (especially in the context of Facebook applications), this is the first work that investigates the strategic aspects of (co)-location sharing in the presence of interdependent privacy risks. Our framework could typically be used to gain insight into users' sharing behavior but also to design appropriate incentive mechanisms and location sharing features in order to influence the behavior of OSN users, eventually optimizing the overall privacy-sharing trade-off.

Our contributions are as follows. We identify the important problem of location sharing with interdependent privacy risks (introduced by co-location), namely the Sharing Game, and we propose the first game-theoretic framework to formalize it. Following a conjoint analysis approach, we design and conduct a user survey of Facebook users (N=250) to quantify users' preferences of (1) sharing or viewing posts, (2) location or co-location information, and (3) location privacy or sharing benefits. Our survey results indicate that, interestingly, there is no consensus regarding users' preferences: For instance, some users prefer sharing location information and others prefer sharing co-location information. We evaluate our analytical framework through simulations, in a

number of key experimental setups and scenarios. We use values of the parameters derived from the empirical data, avoiding the pitfalls of purely theoretical results, for a better understanding of realistic human behaviors. Our simulations notably unravel situations in which users can be forced into a vicious circle of sharing their information or encouraged to over-share.

The rest of the paper is organized as follows. In Section 2, we survey the related work. In Section 3, we describe the considered setting and the system model, including the users and the adversary, as well as the proposed framework for studying users' sharing behaviors. In Section 4, we describe the methodology and the results of the survey of Facebook users in order to estimate the key parameters of our model. In Section 5, we evaluate our framework in a number of scenarios. In Section 6, we discuss directions for improvement and extension of our work. Finally, in Section 7, we conclude the paper and we discuss future work.

## 2 RELATED WORK

Our work is related to two broad research areas: information sharing on OSN and interdependent privacy with game theory.

### 2.1 Information Sharing on OSNs: Privacy & Utility

Users share large amounts of information, including location, co-location and photos, with their friends on OSNs; this comes with privacy risks. Deciding whether to share information (and the precision at which the information is shared), is a complex process. It involves many factors including the users' contexts, the visibility of the shared information (i.e., who has access to it and the relationship between the user who shares the information and the users who can access it [47, 51]), the shared information itself, and the benefits and privacy risks [50] associated with sharing. In some cases, the happiness of a user's friends also becomes part of the decision process; this is usually captured through a so-called *altruistic factor*, as introduced in [30] and experimentally measured using techniques based on conjoint analysis in [37, 38]. Conjoint analysis studies were also used to quantify the value which users attribute to their friends' information in the context of app adoption (e.g., in [39]). In practice, deciding whether to share information often comes down to finding a sweet spot between privacy and benefits [56]. The decision process can be automated by (1) maximizing privacy under benefits (service quality) constraints [43] (or conversely), (2) taking a game-theoretic approach for modeling the interplay between the users and the adversaries [42], or (3) by mimicking the users' sharing decisions using machine-learning techniques, after a training phase [8]. In our work, we model decision making as the optimization of a utility function that incorporates both benefits and privacy. One of our contributions is to parametrize this function by applying conjoint analysis on user data collected through a targeted survey. Also, as users' decisions affect those of other users, we follow a game-theoretic approach for modeling the interplay between users and, ultimately, their decisions.

### 2.2 Interdependent Privacy & Game Theory

The notion of interdependent privacy, i.e., how actions performed by one user affect the privacy of another, was first formalized by Biczók and Chia [7]. Interdependent privacy raises the following concern: Users' privacy is no longer under their sole control. Numerous real-life examples of interdependent privacy risks were studied in the literature, including information about users' friends accessed by Facebook apps [7, 36], sensitive attributes inferred from those of a users' friends on OSNs [4, 13, 31], demographic information inferred from a user's interests [10], genomic data inferred from that of a user's relatives [18, 19], location leaked from geo-tagged pictures that friends upload online [17], relationships inferred from pictures [44], and co-locations detected from the users' IP address at hotspots [49] or reported on OSNs [34]. From a social perspective, a large body of work has been devoted to the study of users' individual and collaborative coping mechanisms for multi-party privacy conflicts

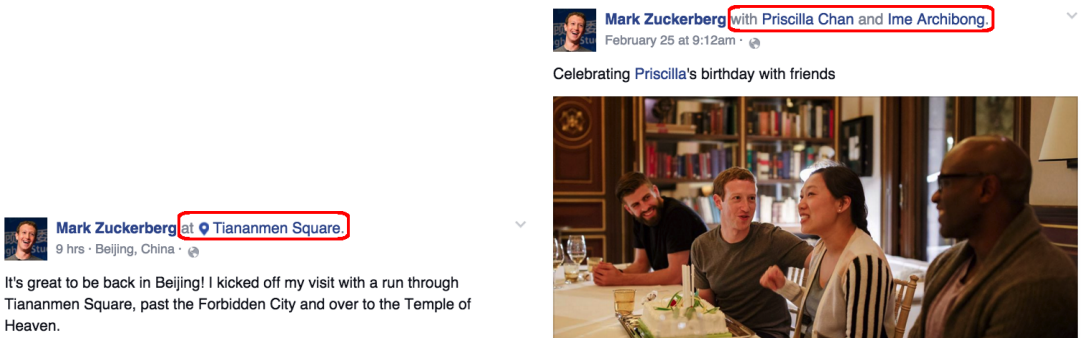
related to co-owned data (also referred to as regulation of interpersonal boundaries) [6, 11, 21, 27, 45, 46, 52, 55]. These works focus mostly on the case of photos sharing on online social platforms and take an experimental and empirical approach to the problem–i.e., they rely on interviews and surveys. Game theory is a first class candidate tool for studying the interactions between users who are subject to interdependent privacy risks, as it enables the modeling of the effect of users' strategies on other users' utility, as well as the users' decision making process. It was successfully used to analyze users' application adoption behaviors [7, 36], privacy decision-making [1], such as sharing genomic data [19]. The study of interdependent privacy risks from an economic perspective follows the long line of research on interdependent security games surveyed in [28].

Our work is the first to study the interactions between OSN users in the case of (co-)location sharing, where shared co-locations create interdependent privacy risks. Unlike the game-theoretic approaches surveyed above, we take into account the time dimension by considering a repeated game, which introduces singular behaviors. In addition, we rely on a rigorous approach, based on user surveys, to determine realistic values of the different parameters of our model.

## 3 SYSTEM MODEL & FORMALIZATION

We consider a mobile location-based online social network (OSN) with standard sharing features. Users are mobile and located within a given geographical region of interest (typically the same city) and time is discrete. At some point in time, $t$, by checking-in at a given location, a user can post information about her location on her OSN profile. She can also post co-location information by tagging a close friend in a picture, or in a status update, thus making this information available to the OSN provider, all her friends and all her tagged friend's friends. Figure 1 illustrates an example of this behavior. In turn, a tagged user can "un-tag" herself from a post in which she is tagged, making this information unavailable to all users but not to the OSN provider. Sharing brings not only social benefits, but also *location privacy* implications, for both the user who shared the information and her tagged friend.

At any time $t$, an adversary–either the service provider or the friends of one or both of these two users–has access to previously reported locations and co-locations and can use this information to infer the users' locations at time $t$. We propose a framework in which, at any time, the decision to post location and co-location information, and the decision to allow a friend to post co-location information, is made strategically by both the users involved.




(a) (b)

Fig. 1. Illustrative screenshots of location (a) and co-location (b) sharing on an online social network (Facebook). These are pictures from the public Facebook profile of Mark Zuckerberg.

| | |
|---|---|
| $s_i(t) = (sl_i(t), sc_i(t))$ | Strategy of user $i$ at time $t$ |
| $\bar{L}$ or $sl.(t) = 0$ (False) | Hide location |
| $L$ or $sl.(t) = 1$ (True) | Share location |
| $\bar{C}$ or $sc.(t) = 0$ (False) | Hide co-location |
| $C$ or $sc.(t) = 1$ (True) | Share co-location |
| $s^*(t) = \left(s_i^*(t), s_j^*(t)\right)$ | Equilibrium strategy profiles (decisions) at time $t$ |
| $B_i\left(t, s_i(t), s_j(t)\right)$ | User $i$'s benefits at time $t$ for strategies $\left(s_i(t), s_j(t)\right)$ |
| $b_{sl}^i$ | User $i$'s benefit of sharing her actual location |
| $b_{vl}^i$ | User $i$'s benefit of viewing her friend's location |
| $b_{sc}^i$ | User $i$'s benefit of sharing co-location with a friend |
| $b_{vc}^i$ | User $i$'s benefit of viewing co-loc. shared by a friend |
| $f_{sv}^i$ | User $i$'s preference factor: sharing vs. viewing |
| $f_{lc}^i$ | User $i$'s preference factor: location vs. co-location |
| $f_{pb}^i$ or $\alpha_i$ | User $i$'s preference factor: privacy vs. benefits |
| $s_i(\cdot) = \{s_i, s_i^*(t-1), \dots\}$ | User $i$'s past decisions and a possible strategy, $s_i$, at $t$ |
| $P_i\left(t, s_i(\cdot), s_j(\cdot), \mathcal{B}\right)$ | User $i$'s privacy at time $t$ for strategies $\left(s_i(\cdot), s_j(\cdot)\right)$ |
| $u_i\left(t, s_i(\cdot), s_j(\cdot)\right)$ | User $i$'s utility at time $t$ for strategies $\left(s_i(\cdot), s_j(\cdot)\right)$ |
| $\alpha_i$ | Weight with which user $i$ values privacy over benefits |
| $SW(t, s_i(\cdot), s_j(\cdot))$ | Social welfare at time $t$ for strategies $\left(s_i(\cdot), s_j(\cdot)\right)$ |

Table 1. Table of notations.

### 3.1 Game Theory 101

Game theory is the study of the strategic interaction between multiple rational decision-makers who aim to maximize their own utility [14, 32, 48]. This mathematical theory enables us to derive more than the optimal strategy that a rational agent would adopt given various parameters: It enables the modeling and prediction of stable states, called *equilibria*, in which none of the agents can improve his utility given all other agents' utility functions and strategies. It has been notably used in economics, biology, political science, psychology, and computer science. Game theory is especially relevant for our work as it enables us to model and analyze users' preferences and interactions, and to predict their resulting rational behaviors. A core concept of game theory is the Nash equilibrium (NE), which represents the stable state in which no agent (a so-called player), by taking into account other players' strategies (so-called opponents), has incentive to deviate from his strategy. A refinement of the NE is the subgame perfect Nash equilibrium (SPNE). This refers to an equilibrium derived by considering a smaller part of the whole game tree, by eliminating incredible threats (strategies that would not rationally be chosen). A common method for finding a SPNE is called backward induction; it first considers the last actions of the game and derives the best decision of the last player, given all other previous possible decisions in the game. Social welfare is defined as the sum of the utilities of all players. A strategy profile (set of players' strategies) is called *social optimum* if it maximizes the social welfare. Note that a NE is not necessarily a social optimum, but that finding a socially-optimal NE is highly desirable.

### 3.2 User Model

We model the interactions between a user and one of her close friends (also called players) as a game over a time window of interest ($\{1, \ldots, T\}$), called the *Sharing Game*.[1] The strategy of a user $i$ at time $t$, denoted by $s_i(t)$, is chosen from any possible combinations of decisions to share or not to share her own location and her possible co-location with her friend. We denote $s_i(t) \triangleq (sl_i(t), sc_i(t))$, where $sl_i(t)$ and $sc_i(t)$ are binary variables that represent whether user $i$ shares location and co-location, respectively. For alternate more compact notations, we use $\bar{L}$ for $sl_i(t) = 0$, $L$ for $sl_i(t) = 1$, $\bar{C}$ for $sc_i(t) = 0$ and $C$ for $sc_i(t) = 1$. When the two players are co-located, each of them can choose any combination of the four possible strategies: $\bar{L}\bar{C}$ – sharing nothing, $\bar{L}C$ – sharing only the co-location information, $L\bar{C}$ – sharing only the location information or $LC$ – sharing both. However, when the users are not co-located they can only choose whether to share their own location, choosing between two possible strategies: $\bar{L}\bar{C}$ – sharing nothing and $L\bar{C}$ – sharing location information.

The social benefits of user $i$, which stem from a decision to share information at some time $t$ are denoted by $B_i(t, s_i(t), s_j(t))$, where $j$ denotes the other user. Her privacy at $t$, denoted by $P_i(t, s_i(\cdot), s_j(\cdot), \mathcal{B})$, is a function of both users' strategies at times $\{t-k, ..., t\}$, where $k \in \{0, \ldots, t-1\}$ denotes the number of previous time instants the adversary uses to gather reported information. The privacy function can incorporate specific background user information (denoted by $\mathcal{B}$), e.g., her mobility profile. Although we consider that a user's benefits at some time only depend on the users' strategies at that time, we emphasize that the privacy function takes into account previous time instants as well. In other words, a decision made at time $t$ has privacy implications at later time instants.

The utility function of a player at time instant $t$ captures both her social benefits and her privacy and we assume a player is only interested in the privacy at the current time she has to make a decision, $t$.

$$u_i(t, s_i(\cdot), s_j(\cdot)) \quad = \quad (1-\alpha_i)\, B_i\left(t, s_i(t), s_j(t)\right) + \alpha_i P_i(t, s_i(\cdot), s_j(\cdot), \mathcal{B}) \tag{1}$$

where $\cdot$ denotes the times $\{t-k, ..., t\}$ and $\alpha_i \in [0, 1]$ denotes the weight with which user $i$ values her privacy over her social benefits. Note that, in the decision making process, *players can be assisted by a tool to evaluate the privacy implications*, namely the value $P_\cdot(\cdot)$, of each of their decision regarding sharing–for instance, a Facebook extension would suggest this, thus complete information would be available to the players in the decision making process.

At any $t$, a player's social benefits are computed as a normalized sum of the benefits of sharing information (i.e., location and co-location) and viewing information shared by her friend, specifically,

$$B_i(t, s_i(t), s_j(t)) = \frac{b^i_{sl} sl_i(t) + b^i_{sc} sc_i(t) + b^i_{vl} sl_j(t) + b^i_{vc} sc_j(t)}{b^i_{sl} + b^i_{sc} + b^i_{vl} + b^i_{vc}} \tag{2}$$

where $b^i_{sl}$ and $b^i_{sc}$ denote user i's benefit of sharing location and co-location and $b^i_{vl}$ and $b^i_{vc}$ her benefit of viewing location and co-location. Note that these parameters are specific to a user.

The game is played repeatedly, at successive time instants, from 1 to $T$. At every time instant, we model the interactions as a perfect and complete information, non-cooperative, extensive form game. This type of game corresponds to the interactions in a typical OSN, where the players' actions at some instant are inherently ordered: The second player (or his application implementing the decision model) knows the choice of the first one and decides (or suggests the player) her strategy accordingly. Therefore, without loss of generality, we consider that the players' actions are ordered at every time instant. In reality, players would also play such a game successively over time (for each sharing action), hence our choice of a repeated game.

We list the following assumptions, that properly model the existing OSNs' interfaces (such as Facebook's): (1) Location posts of a player are visible to all her friends *and* to the service provider. (2) Co-location posts

[1]Note that the adversary, with respect to whom the users' privacy is evaluated (typically the service provider), is *not* a player of the game.

initiated by either of the two players are always visible to the service provider and cannot be removed (even if the second player removes them, the service provider still has this information). (3) For a co-location post to be visible to friends of the two players, both of them have to agree to share it, in which case it is visible to the union of their friends. (4) If a player un-shares a co-location shared by the first player (by un-tagging or even asking it to be removed), the first player cannot share that co-location again. (5) Decisions made by the players are considered fixed: Once they strategically choose the best decisions at time $t$, they will not revisit them at later time instants. Table 1 summarizes the notations used in our formalism.

We are aware of the fact that some people might act irrationally, especially when it comes to privacy-related decisions [3]. Yet, we believe that privacy protection demand will increase, notably because a growing number of people suffer the consequences of their (and others') privacy carelessness. Moreover, smartphones are increasingly involved in the sharing decisions users make, as demonstrated by the growing sophistication of the apps' permission systems. A privacy-protection software run for this purpose *can* be "rational" and strictly follow the parametrization model provided by its user to aid him in decision making. It is therefore of high interest to see what happens under the assumption of *rationality*.

## 3.3 Adversarial Models

Although the adversary is not a player in our game, the privacy of the players depends on who the adversary is: For the same strategy profile, different adversaries have access to all or only some of the shared information. We consider four possible adversaries, specifically the service provider and three different sets of users, essentially subsets of the players' friends. Note that these are all adversaries that our survey participants report being concerned about and we considered the adversaries and the information that is available to them for the typical default privacy settings for OSN posts.

*3.3.1 Service Provider adversarial model (SP).* The service provider adversary has access to all location and co-location posts made by the players. The specificity of this adversary is that, once either of the players shares information, this information is always known to him. In other words, the second player cannot un-share co-location information with respect to the service provider. We assume that the SP does not gather location information about its users (i.e., the players) through other channels, such as from their IP address.[2]

*3.3.2 Friends adversarial models (MF, FF, CF).* In these adversarial models, privacy is computed from the perspective of the players' friends. The common point of these models is that, unlike the SP model, co-location information potentially shared by the first player can be removed by the second one (e.g., by un-tagging). Figure 2 illustrates the valid set of players' strategies in this case. We consider three different subsets of the friends, based on the information available to each of them, as illustrated in Figure 3: (i) "My other friends model" (MF)–this adversary has access to all the location posts made by the player and to co-location posts made by both players; (ii) "My friend's other friends model" (FF)–this adversary has access to all the location and co-location posts made by the other player and to co-location posts made by the player; and (iii) "Our friends in common model" (CF)–this adversary has access to all location and co-location posts made by both players. Note that the FF adversary can also be representative (with a possibly higher value for $\alpha$.) for a public adversary–while this is not the default visibility for posts, users can post information with public visibility.

## 3.4 Analysis Methodology

At each time instant, $t$, we use backward induction, a typical method for finding a subgame perfect Nash equilibrium (SPNE) that dictates the players' decisions. Equilibria decisions made at time instants prior to $t$, denoted by $s^*(t') = \left(s_i^*(t'), s_j^*(t')\right)$, where $t - k \leq t' < t$, are used when computing the privacy of the players.

[2]This could be achieved by simply incorporating such side information in the privacy evaluation function.

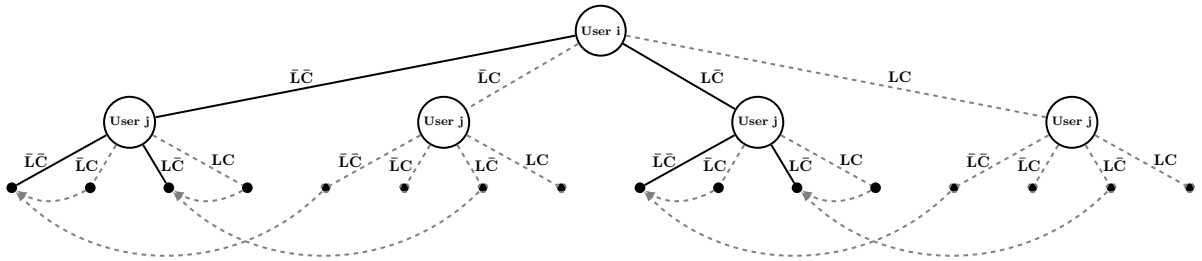
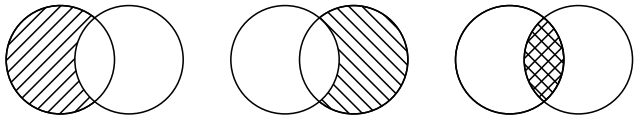



Fig. 2. Possible strategies for one time instant of the Sharing Game in the friends adversarial models (depicted in Figure 3). User $i$ is the first player (he chooses a strategy first) and user $j$ the second player (he reacts to $i$'s choice). Only the black solid strategies are valid when the two users are *not* co-located. All strategies (including the gray dashed ones) are valid when the two players are co-located. Horizontal arrows indicate the fact that the second player can revert a co-location shared by the first player (e.g., by un-tagging herself or asking that the post be removed), hence choosing not to share the co-location. Therefore, when co-located, only strategy profiles in which the players agree whether or not to share their co-location are valid.



(a) MF (b) FF (c) CF

Fig. 3. Friends adversarial models (hashed area) for user $i$: (a) My other friends model (MF); (b) My friend's other friends model (FF); (c) Our friends in common model (CF). The social circle of user $i$ (resp. $j$) is represented by the left (resp. right) circle. Their intersection represents the common friends of $i$ and $j$.

The first player, player $i$, anticipates the second player's (player $j$'s) best response, as a function of her possible strategies $s_i$, essentially

$$\forall s_i,\ s_j^*(s_i) = \arg\max_s u_j\left(t, s_i, s, s^*(t-1), ..., s^*(t-k)\right) \tag{3}$$

where $s^*(t-1), ..., s^*(t-k)$ denote the users' past decisions.

This eliminates incredible outcomes that player $j$ would never rationally choose. Player $i$ chooses her best strategy out of the remaining outcomes, as follows

$$s_i^* = \arg\max_s u_i\left(t, s, s_j^*(s), s^*(t-1), ..., s^*(t-k)\right) \tag{4}$$

The equilibrium decisions at time $t$ are then given by

$$s^*(t) = \left(s_i^*(t), s_j^*(t)\right) = \left(s_i^*, s_j^*(s_i^*)\right) \tag{5}$$

We define social welfare, at some time $t$, as the sum of the players' utilities, for any strategy profile, specifically

$$SW(t, s_i(\cdot), s_j(\cdot)) = u_i(t, s_i(\cdot), s_j(\cdot)) + u_j(t, s_i(\cdot), s_j(\cdot)) \tag{6}$$

In case of multiple equilibria at time $t$, the players coordinate and choose the one that maximizes their social welfare. The game is *repeated* in a similar way at successive time instants, each time taking into account the players' decisions from previous time instants.

### 3.5 Equilibria Properties

We are interested in different properties for the players' equilibria decisions.

*3.5.1 Social-optimality at equilibrium.* We say that the social welfare is maximized for the equilibrium decisions at time $t$ (or, equivalently, that the equilibrium at time $t$ is socially-optimal) if the following property holds:

$$\forall \left(s_i(t), s_j(t)\right) \neq \left(s_i^*(t), s_j^*(t)\right) : \\ SW\left(t, s_i^*(t), s_j^*(t), s^*(t-1), ..., s^*(t-k)\right) \geq SW\left(t, s_i(t), s_j(t), s^*(t-1), ..., s^*(t-k)\right) \quad (7)$$

*3.5.2 Individual utility maximization at equilibrium.* A player's $i$ utility is maximized for the equilibrium decisions at time $t$ if the following property holds:

$$\forall \left(s_i(t), s_j(t)\right) \neq \left(s_i^*(t), s_j^*(t)\right) : \\ u_i\left(t, s_i^*(t), s_j^*(t), s^*(t-1), ..., s^*(t-k)\right) \geq u_i\left(t, s_i(t), s_j(t), s^*(t-1), ..., s^*(t-k)\right) \quad (8)$$

We consider the proportion of time instants, across {1,…,T} for which the equilibria decisions are socially-optimal and the proportion of time instants for which the equilibria decisions maximize each player's utility.

## 4 SURVEY

The model presented in the previous section includes a number of parameters that appear in the expression of the utility function that drives the users' strategic behaviors. As such, these parameters characterize the users' sharing behaviors; in practice, they vary from one user to another. In order to obtain realistic values of these parameters, as well as to study the general trend and the variability across users, we conducted a survey of Facebook users in 2016.

### 4.1 Conjoint Analysis 101

We briefly introduce the conjoint analysis technique in this paragraph, which can be skipped by the knowledgeable reader. Conjoint analysis [16] is an experimental approach used to detect the hidden rules users rely on to make decisions (involving trade-offs) between services. In this approach, a service is viewed as a combination of attributes, each of which has different levels (values). Users are asked to rank multiple versions of the service (each being a different combination of attribute levels). The combination of attributes and levels can lead to a large number of versions to be ranked. In order to keep the complexity of this task manageable for the users, the number of proposed versions can be reduced, in an optimal way, to a reasonable yet meaningful number, through *fractional factorial design* [29]. The hidden value users place on each of the attribute levels is then quantified through statistical analysis, as *part-worth utilities* and *importance values*. The importance values represent how much difference each attribute makes in the total utility of the service.

### 4.2 Methodology

We recruited participants through the Amazon Mechanical Turk platform. To be eligible, they were required to have a minimum Human Intelligence Task (HIT) approval rate of 95% with at least 100 past approved HITs and an active Facebook account. We checked this last criterion by using the "Log-in with Facebook" feature. We only use the information about the participants' Facebook account for screening purposes and we did not store any such information; we made this point very clear in the advertisement page of our survey (in order to not discourage privacy-concerned potential participants).

After the standard demographic questions (part I), we polled the survey participants about their preferences regarding the posts they share or view on social networks (part II). The second part of the survey was composed of three questions to assess the participants' preferences regarding, respectively, (1) sharing vs. viewing posts with location information (i.e., check-in posts), (2) sharing posts with location information vs. sharing posts with co-location information, and (3) location privacy vs. benefits of sharing location information. We designed these three survey questions by following a rigorous *full-profile conjoint analysis* approach [16] and making use of a dedicated tool, namely XLSTAT [53]

This approach enables us to quantify individual values for each of the participants' preferences factors.

**Sharing vs. Viewing ($f_{sv}$).** After a brief reminder about what a check-in post is (illustrated with a screenshot of a Facebook timeline), the participants were told that, for technical reasons, some of their own two most-recent check-in posts and some of their friends' two most-recent check-in posts might be removed from Facebook. Then, the participants were asked to rank by preference a number of scenarios corresponding to different combinations of the numbers of posts kept (e.g., "two of *your* recent posts are kept and one of *your friend's* recent posts is kept", "none of your recent posts is kept and one of your friend's recent posts is kept"). The participants were asked to take into account only benefit considerations (i.e., not privacy). In order to limit the bias coming from the content of the posts, we explicitly mentioned that these are all the posts that they once shared and they would like to keep, and we did not include the content of the participants' actual recent posts in the survey page. The initial ordering of these options was randomized. For this question, two attributes were used: the number of the participant's *own* kept check-in *posts* and the number of the participant's *friends'* kept check-in *posts*. Each attribute had three possible values (i.e., none, one or two). This yielded an optimal number of five options to rank (out of a total of nine). In order to detect sloppy answers, we included in the list of options to be ordered a sixth option in which no posts are removed, and we explicitly stated in the text of the question that this should be the preferred option. The ranking provided by the users enabled us to compute their preference factors $0 \leq f_{sv} \leq 1$, from the importance values attributed to each attribute: $f_{sv}$ is the normalized importance value of the attribute *own posts*, whereas $1 - f_{sv}$ is the normalized importance value of the attribute *friends' posts*. A value greater than 0.5 denotes a preference for *sharing* information over *viewing* information.

**Location vs. Co-location ($f_{lc}$).** This question was designed by following the same methodology as for the first question: After a brief reminder about what a co-location post is (illustrated with screenshots), the participants were asked to order, according to their preferences, six options in which a number of their own recent posts with *location* information and a number of their own recent posts with *co-location* information would be removed (e.g., "two of your recent check-in posts are kept and one of your recent co-location posts is kept."). The ranking provided by the users enabled us to compute their preference factors $f_{lc}$, similarly to $f_{sv}$.

**Location privacy vs. Sharing benefits ($f_{pb}$).** After a brief reminder about location privacy, the participants were asked to order, according to preference, six options with different numbers of check-in posts and the corresponding levels of location privacy, in terms of the average precision with which their location can be inferred during a day (e.g., "12 location posts for an average location privacy of 400 m"). These numbers were extracted from the experimental results presented in [34]. The ranking provided by the users enabled us to compute their preference factors $f_{pb}$, similarly to $f_{sv}$.

Finally (part III), we polled the participants about their usage of Facebook, their privacy concerns, and about their knowledge of the privacy threats related to (co)-location information. The transcript of the questionnaire is shown at the end of the manuscript and aggregated results are available at https://infoscience.epfl.ch/record/218755.

It took approximately ten minutes to complete the survey; the participants were paid $2 for their work. We ruled out participants with inconsistent responses in part II. More specifically, we considered as inconsistent a ranking that violates the natural order, i.e., considering that removing some of the existing posts is preferable to keeping them all. In the end, we obtained a sample of $N = 250$ valid participants; the sample was diverse and balanced in terms of the participants' demographics: 46% of the participants were female, the participants had

various primary areas of employments, and their ages ranged from 19 to 68 years old, with an average of 33 and a standard deviation of 9.48. The participants were active Facebook users: 70% of the participants declared that they use Facebook multiple times per day (93% do so multiple times per week), 30% of them make at least one post with location information per week, and 37% of them make at least one post with co-location information (in statuses, in posts or in pictures) per week.

### 4.3 Results

We extracted the aforementioned three preference factors from the survey data by using XLSTAT. Note that, due to the fact that only a limited number of scenarios can be presented to the participants for ordering, the preference factors can take only a limited number of values. We illustrate the relevant statistics in Table 2 and in Figure 4. Note that these results should be taken with a grain of salt as previous works (e.g., [2, 22]) have shown that (reported) privacy attitudes do not always correspond to actual behaviors. We observe that the average of the factors is close (yet slightly higher) than 0.5 (specifically, .57 ± .15, .56 ± .15 and .60 ± .39 for $f_{sv}$, $f_{lc}$ and $f_{pb}$, respectively); this means that there is no strong consensus among the participants regarding their preferences. In fact, the distributions of the factor values are bi-modal: Users tend to have a clear preference for one of the two options (e.g., location vs. co-location). This phenomenon appears clearly for $f_{pb}$ (i.e., privacy vs. benefits) that has a high standard deviation (0.39). In the case of $f_{sv}$, for instance, the proportion of indifferent users (for whom $f_{sv} = 0.5$) is substantial (16.8%) and almost as large as the proportion of users who prefer viewing over sharing (23.2%). These results are in line with those of previous studies that showed that there exist multiple usage profiles on social networks: Some users connect to social networks mostly to share news with their friends whereas others do so mostly to view news about their friends [5, 35]. 54% of the users prefer location to co-location information ($f_{lc} > 0.5$) and 20% do not have a preference ($f_{lc} = 0.5$), whereas 63.2% favor privacy over social benefits ($f_{pb} > 0.5$).

| | $f_{sv}$ | $f_{lc}$ | $f_{pb}$ |
|---|---|---|---|
| avg. ± stddev. | .57 ± .15 | .56 ± .15 | .60 ± .39 |
| proportion of users with $f_* > 0.5$ (prefer sharing/ location/ privacy) | 60% | 54% | 63.2% |
| proportion of users with $f_* = 0.5$ (indifferent) | 16.8% | 20% | N/A |
| proportion of users with $f_* < 0.5$ (prefer viewing/ co-location/ benefits) | 23.2% | 26% | 36.8% |

Table 2. User preference factors extracted from the survey data by using a conjoint-analysis approach. $f_*$ denotes, depending on the column, $f_{sv}$, $f_{lc}$, or $f_{pb}$.

As for the questions related to privacy issues on Facebook, 24.8% of the participants declared being "very concerned" about privacy, 50% declared being "moderately concerned" and 25.2% not concerned (as illustrated in Figure 5a). When the participants report being co-located with a friend (say Bob), their feared adversaries are Bob's friends who are not friends with the participant (i.e., the FF model, 44% of the participants), the common friends of Bob and the participant (i.e., the CF model, 24.4%), Facebook (i.e., the SP model, 24.4%) and the participants' friends who are not friends with Bob (i.e., the MF model, 21.2%), as illustrated in Figure 5b; 26% of the participants reported not being concerned by any of these adversaries. 42.4% of the participants were not aware that their friends' posts that include location or co-location information can decrease their own location privacy. Only 50% of the participants declared being aware that their posts have privacy implications for themselves and for their friends, whereas 30.8% of the participants were not aware that their posts have any effect on privacy (as illustrated in Figure 6). Finally, we asked the participants whether the survey would affect their future sharing behavior on

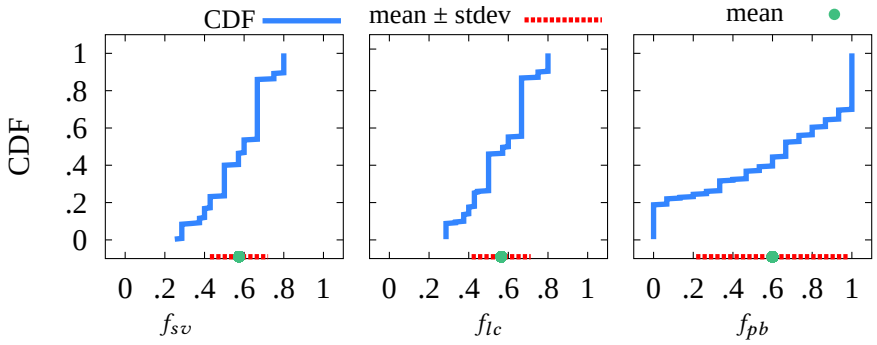



Fig. 4. CDFs of the preference factors of our survey participants.

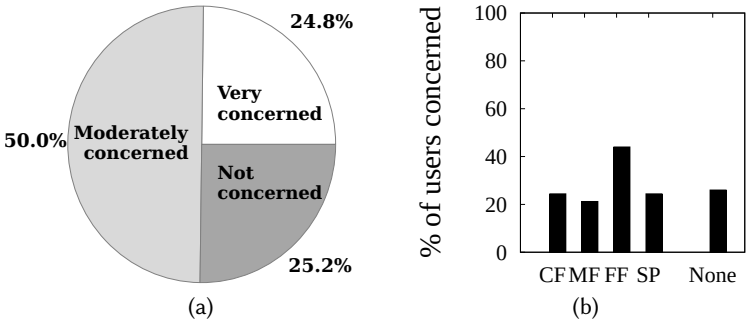



Fig. 5. (a) Users' concern about location privacy; (b) The adversaries that users are concerned about: Our friends in common (CF), My other friends (MF), My friend's other friends (FF), The service provider (SP).

Facebook: A substantial fraction of the participants (around 35%) declared they would be more careful, especially for co-location information, by, for instance, preventing their friends from tagging them in posts:

> "I may remove tags or ask friends not to tag me with locations in the future." (female, 35 y/o)
>
> "I may think twice before checking in, or at least consider the impact tagging others has on their privacy." (male, 31 y/o)
>
> "Yes because I was unaware of this issue and it now makes me a little scared." (male, 19 y/o)

Of the participants who stated that their behavior would not change, 31% declared already being careful with their posts and tags.

An anonymized and sanitized version of the answers (part II and some of part III) is available (password *FbS250*) at https://www.dropbox.com/s/4ukfg5nbhd1x0p7/data.zip?dl=1.

## 5 EVALUATION

We evaluate our framework by simulating and analyzing the users' sharing decisions in different experimental setups.

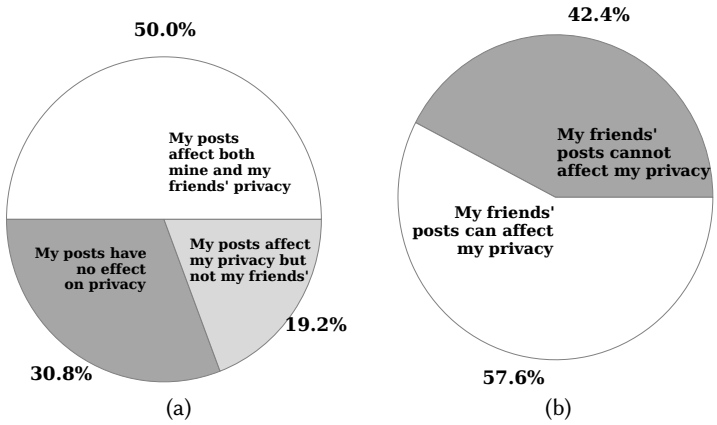



Fig. 6. Users' awareness about (a) privacy risks stemming from their own posts; (b) own privacy risks stemming from friends' posts.

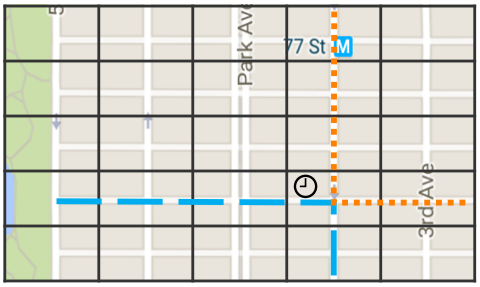



Fig. 7. Scenario considered in the evaluation: Two users, Alice (dotted) and Bob (dashed), coming from distinct directions, meet for some time, and later separate in distinct directions.

## 5.1 Experimental Setup

In this section, we describe the experimental setup of the different building blocks of our framework.

*5.1.1 Quantification of the users' privacy.* We quantify users' privacy ($P_.$) as their location privacy, by relying on the inference framework proposed by Olteanu et al. [34][3]; we re-use the corresponding formalism and software library. In short, we assume discrete locations (i.e., the geographical area of interest is partitioned into cells by using a regular square grid; when reporting their locations, users report the cells in which their actual locations fall; and the adversary has access to the users' mobility profiles in the form of transition probabilities between cells). Privacy is computed as the adversary's expected error when localizing users, using a junction tree exact inference algorithm on the Bayesian network [23] that models the probabilistic dependencies between all the users' locations over the time period of interest. The location and co-location disclosures available to the adversary depend on the considered adversary, among those presented in Section 3.3, namely the SP, MF, FF, and CF models, and on the users' strategic decisions. For the sake of simplicity, we consider the same adversary for both users:

[3]Note that our model is flexible enough to allow the use of other frameworks for inferring location privacy, for instance, that proposed by Xu et al. [54].

For example, if the first user's location privacy is computed with respect to the OSN service provider, so is that of the other user. At each time instant $t$, the adversary considers *all* past location and co-location posts from the users when inferring their locations. Note that, in practice, users are not yet able to evaluate their privacy accurately, but this aspect could be provided to them by the OSN or by a software module (e.g., in a mobile app).

*5.1.2 Users' parameters.* We rely on the results of our user survey to parametrize the users' utility function. More specifically, we derive the value of the parameters $\alpha$, $b_{sl}$, $b_{sc}$, $b_{vl}$ and $b_{vc}$ from the values of the preference factors $f_{pb}$, $f_{lc}$ and $f_{sv}$. To do so in a coherent way and keep the number of parameters low, we make a few evaluation assumptions: We assume that (1) the users' preferences between sharing and viewing is the same for posts with location information as for posts with co-location information, (2) the users' preferences between posts with location information and posts with co-location information is the same for the users' own posts as for their friends' posts. Using these assumptions, we derive the values of the parameters from the preference factors (of which we consider different values in each experiment) as follows: $\alpha = f_{pb}$, $b_{sc} = \frac{f_{sv}}{1-f_{sv}} b_{vc}$, $b_{vl} = \frac{f_{lc}}{1-f_{lc}} b_{vc}$, $b_{sl} = \frac{f_{sv}}{1-f_{sv}} \cdot \frac{f_{lc}}{1-f_{lc}} b_{vc}$ where $b_{vc}$ is a free variable (we set it to 1).

*5.1.3 Scenario.* In order to evaluate our framework and to gain insight about the effects of the different parameters, we consider the canonical meeting scenario, illustrated in Figure 7: Two users, Alice and Bob, coming from distinct locations ($t = 1$), meet for some time (one time unit, $t = 2$), and later separate in distinct directions ($t > 2$). We consider $T = 5$ time instants in total. At each time instant, both Alice and Bob can either report or hide their actual location. Additionally, at $t = 2$, either of them can choose to report being co-located with the other. The adversary uses, in the inference process, all the users' reports and the same basic mobility profile for Alice and for Bob: In one time unit, Alice/Bob either stays in the cell she/he is in (with probability .5) or moves to one of the neighboring cells (with the remaining equal probabilities). The rationale behind this choice is to understand the basics of the interplay between the users, independently from the specifics and the singularities of their individual data.

## 5.2 Experimental Results

In order to understand the effect of each of our model's parameters, we study through simulations the different strategic decisions players choose in several situations[4].

*5.2.1 The effect of the considered privacy adversary.* We study how the adversary that is considered by the players when assessing their privacy influences their decisions. In a first experiment, we consider a *homogeneous* scenario, in which the parameters in both the users' utility are set using the average values of $f_{sv}$, $f_{lc}$ and $f_{pb}$ obtained in our survey, as presented in Figure 4. Figure 8 illustrates the different game outcomes, for the four adversarial models we presented in Section 3.3. A first observation is that the players' decisions are quite diverse, thus demonstrating that the adversarial model can influence what players share.

In the *SP* and *CF* models (Figures 8a and 8b), at $t = 1$ (when no co-location has yet been reported and thus there is no correlation between the users' locations or their privacy), the equilibrium decisions are that nothing be shared–the first blue rectangle and red circle pair. Note that for all time instants where users are not co-located ($t \neq 2$), the equilibrium decisions can only be "share nothing" or "share location". The equilibrium at $t = 1$ maximizes social welfare (there is a green triangle for $t = 1$), but either of the players would have a higher utility (both the blue rectangle and the red circle are empty) if the other one shared his own location (because, in the current absence of correlation, they would enjoy viewing where their friend is without any privacy cost to themselves). However, such an outcome is not an equilibrium because neither of them wants to share their location at this time (mainly due to the fact that the social benefit gained by sharing location would be less than

[4]More situations are described in our technical report [reference will be available for the final version]

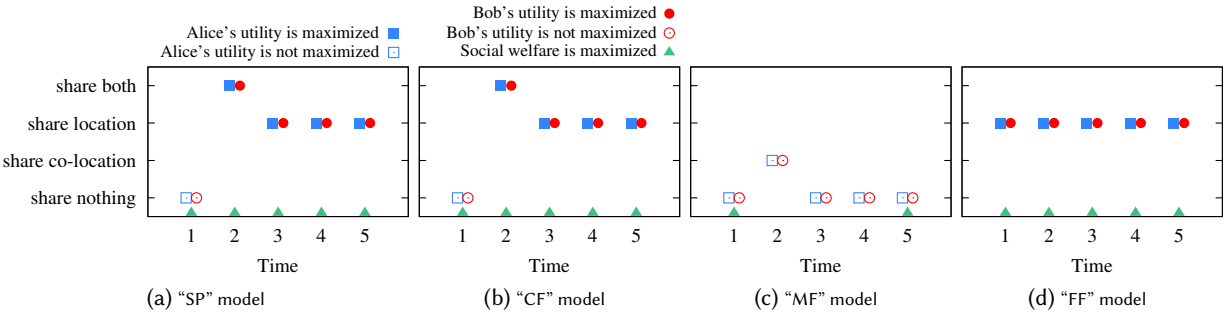



Fig. 8. Players' decisions at equilibrium, $\left(s^*_{Alice}(t), s^*_{Bob}(t)\right)$, for $f_{sv} = 0.57$, $f_{lc} = 0.56$, $f_{pb} = 0.60$ and different adversarial models: (a) Service Provider (SP), (b) Our friends in common (CF), (c) My other friends (MF), (d) My friend's other friends (FF) models. The $x$ axis shows the time window of interest. On the $y$ axis, for every time instant, Alice's decision is represented by a blue rectangle and Bob's decision by a red circle. A player's corresponding shape is full if its utility at equilibrium is maximized, and empty otherwise. Additionally, each time instant is marked by a green triangle, if the equilibrium decisions maximize social welfare.

their incurred privacy loss, weighted by $1-\alpha$ and $\alpha$, respectively). At time $t = 2$, when the players are co-located, the additional benefit of sharing a co-location along with the benefit of sharing a location, overcomes the privacy loss; and the players' equilibrium decisions are that everything be shared $(LC, LC)$. This equilibrium not only maximizes social welfare, but also gives the best utility for both of the players at this time. Once these decisions to share have been made at $t = 2$, the privacy at $t = 3$ is already substantially compromised; hence the benefit of sharing location overcomes the (now) small relative privacy loss and both players choose to share everything, that is, their own locations. Similarly, the decision to share a location at $t = 3$ affects a player's privacy at $t = 4$ severely enough that they again decide to share their location (for the social benefits) and this effect propagates at successive time instants.

In the *MF* model (Figure 8c), there is a different equilibrium at the time of co-location, $t = 2$. The outcome where both players share everything, $(LC, LC)$ is still the one that maximizes social welfare, but it is no longer an equilibrium because each of the players can now deviate from it by not sharing their own location to achieve better privacy, hence utility (e.g., outcome $(LC, \bar{L}C)$ would be better for Bob than outcome $(LC, LC)$, because his adversary–his friends who are not Alice's friends–cannot see that Alice also shares her location). This was not the case in the *SP* model, where information shared by either player is automatically seen by the provider). In this case, the equilibrium is outcome $(\bar{L}C, \bar{L}C)$: Sharing only a co-location does come with a small privacy cost (privacy can decrease even when only co-location and no location information is available due to the mobility profiles, as demonstrated in [34]), but this loss is smaller than the benefit gained by sharing. This equilibrium maximizes neither the social welfare nor a player's utility (either of them would have a better utility if the other would share their location, because they enjoy viewing where their friend is, at no privacy cost to themselves). At time $t = 3$, the players' privacy is higher than it was in the *SP* and *CF* models–for any strategy profile–because the decisions made at $t = 2$ provide the adversary with less information. Sharing the location is not justified because, in this case, the privacy cost this would bring is higher than the benefit gain, hence the equilibrium decisions are that nothing be shared. This equilibrium does not maximize players' utilities (each would still prefer to see the other's location at no privacy cost) or the social welfare. This effect is propagated over time, at successive time instants, and the equilibria decisions are the same, that nothing be shared. Furthermore, as the effect of

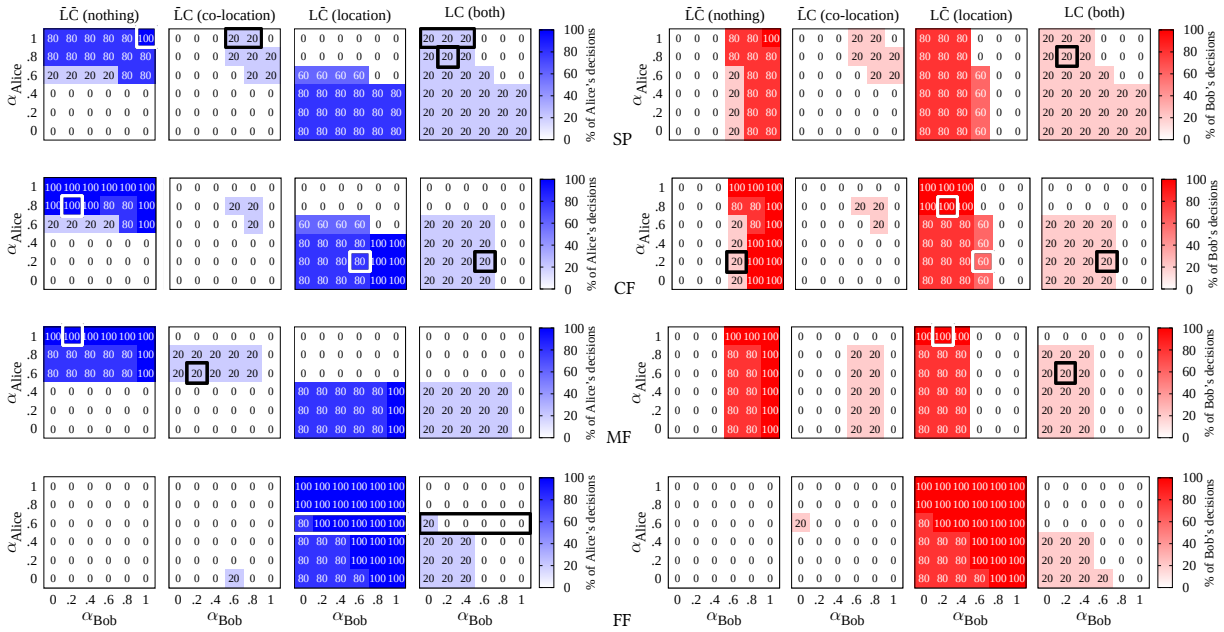



Fig. 9. Players' decisions at equilibrium, aggregated over time for $f_{sv} = 0.57$, $f_{lc} = 0.56$, and different adversarial models: Service Provider (SP)–first row, Our friends in common (CF)–second row, My other friends (MF)–third row, My friend's other friends (FF)–forth row models. For each adversarial model and each possible combination of values for $\alpha_{Alice}$ and $\alpha_{Bob}$, eight heatmaps (left four for Alice, right four for Bob) indicate the percentage of times, aggregated over the number of time instants, that a player made one of the four possible decisions: "share nothing", "share location", "share co-location" or "share both" (in all combinations $\alpha_{Alice}$-$\alpha_{Bob}$, the values of the four cells for a player sum to 100). We highlight with rectangles the cases that we discuss in Section 5.2.2.

the reported co-location at time $t = 2$ *fades away* over time, privacy increases, and at $t = 5$ the equilibrium also maximizes social welfare.

Finally, in the *FF* model (Figure 8d), the equilibrium at times when the players are not co-located is always $(L\bar{C}, L\bar{C})$: In this case, sharing their own location brings them some social benefits without any privacy costs (this adversary cannot see if they share location). When players are co-located, the equilibrium is $(L\bar{C}, L\bar{C})$ (sharing also the co-location would result in minimal privacy), and it maximizes both the social welfare and both the players' utilities.

*5.2.2 The effect of privacy vs. benefits preferences.* We present a *heterogeneous* scenario, where players place different importance on privacy and social benefits. We consider the average values for $f_{sv}$ and $f_{lc}$ and vary $f_{pb}$ in $[0, 1]$. Figure 9 illustrates our results (see caption for details). Obviously, when players have different values for $f_{pb}$ (recall that $\alpha = f_{pb}$), their interests can be in conflict and decisions at equilibrium might differ: When co-located ($t = 2$), one player might share only co-location, whereas the other shares both (e.g., in the *MF* model when $\alpha_{Alice} = 0.6$ and $\alpha_{Bob} = 0.2$ Bob shares both, while Alice shares only co-location[5]) or one shares his location, whereas the other shares nothing (e.g., in the *MF* model when $\alpha_{Alice} = 1$ and $\alpha_{Bob} = 0.2$ Alice shares nothing, whereas Bob only shares his location).

[5]Recall that "share co-location" and "share both" decisions can only occur when the players are co-located (i.e., 20% of the times).

An interesting observation is that, in the *SP* model, when the two players are co-located, the equilibria strategies are always in the form of $(\bar{LC}, \bar{LC})$, $(\bar{LC}, \bar{LC})$ or $(LC, LC)$. This stems from the fact that if *one* player wants to share the co-location information, as the service provider automatically has access to it, the privacy of the other player is already compromised and he is *forced into sharing* as well and at least obtains the associated social benefits. This leads to equilibria in which one player's utility, or even the social welfare, are not maximized. Such outcomes can be avoided in the other models, where a player can undo the co-location shared by the other, and only equilibria with strategies where both players share or do not share the co-location information are allowed. An example can be observed in Figure 9, for $\alpha_{Alice} = 0.8$ and $\alpha_{Bob} = 0.2$: In the *SP* model, Alice is forced into sharing her location *and* co-location information at $t = 2$ because Bob, who places little importance on privacy, shares both, and the equilibrium is $(LC, LC)$; in the *CF* model, Alice does not allow Bob to post co-location information about her and the equilibrium in this case becomes $(\bar{LC}, L\bar{C})$–Alice shares nothing while Bob only shares his location.

Another observation is that, **in all adversarial models, *both* players tend to share more as *one or both* their $\alpha$ decreases** (i.e., as one or both value privacy less). Notably, a player's strategy can change, even when only his friend's preferences change. Let us look, for example, at the average case of $\alpha_{Alice} = 0.6$: As $\alpha_{Bob}$ decreases from 1 to 0, the amount of sharing Alice does increases (e.g., in the FF model, Alice only shares her location when $\alpha_{Bob} \in [0.2, 1]$, but she also shares the co-location when $\alpha_{Bob} = 0$). The same observation holds for the other values of $\alpha_{Alice}$. For the *SP* model, in particular, when Alice is very privacy conscious ($\alpha_{Alice} = 1$), her preferred outcome when co-located would be to share nothing, but she can only do this when $\alpha_{Bob} = 1$. She can gradually be forced into sharing her co-location with Bob (when $\alpha_{Bob} \in [0.6, 0.8]$) or even their co-location and her location (when $\alpha_{Bob} \leq 0.4$). Furthermore, the propagation of this effect can be observed not only at times where the players are co-located. Let us look, for example, at the case where $\alpha_{Alice} = 0.2$ and $\alpha_{Bob} = 0.6$: In the *CF* model, before his co-location with Alice (at $t = 1$)[6], Bob decides to not share anything (20% of the times). Once co-located, Bob and Alice have enough incentive to share both their co-location and location (20% of the times). After their co-location, Alice still has incentive to share her location. Their previously reported co-location, as well as Alice's successive reports of her location, continue to damage Bob's privacy, and he counteracts these losses by also sharing his location for the benefits (60% of the times).

*5.2.3 The effects of multiple users' preferences.* We present a more realistic setup, where each of the two players' parameters are assigned from the individual *preference profiles* of the survey participants. A preference profile represents the values of all preference factors ($f_{sv}, f_{lc}, f_{pb}$), for a specific survey participant; there are 250 such preference profiles. Analyzing the players' behaviors is substantially more complicated, due to the multiple influences present in such a complex setup. In order to find a meaningful interpretation, we alternatively split the 250 preference profiles into two subsets, based on the value of one of the preference factors.

**The case of *sharer* / *viewer* players**. We study how the fact that the players have different values for the $f_{sv}$ preference factor affects their decisions. We select two subsets of preference profiles from our survey data: the *sharers* (150 profiles)–for which $f_{sv} > 0.5$–and the *viewers* (58 profiles)–for which $f_{sv} < 0.5$. We evaluate the outcome of the Sharing Game in three cases, for each possible pairs of preference profiles: when Alice has a *sharer*'s preference profile and Bob a *viewer*'s, when both have *sharers* profiles and when both have *viewers* profiles.

Figure 10 shows our aggregated results (see caption for details). We note that the interplay between the various parameters of the preference profiles (e.g., a *sharer* profile encourages sharing because $f_{sv} > 0.5$, but it could also discourage sharing if $f_{pb} > 0.5$) results in a large variety in the distribution of players' equilibria decisions. Despite this variability, a few trends are still distinguishable. First, in general, a *sharer* shares more information than a *viewer* and the most information is shared when both Alice and Bob are *sharers*, whereas

[6]This detail is not directly readable form Figure 9, as it presents statistics aggregated over time instants.

the least information is shared when both are *viewers*. Second, regardless of the players' types (*sharer*/*viewer*), and due to the forcing effect, the largest amount of co-location is shared in the SP model (e.g., 17% of all time instants when both players are *sharers*); the smallest amount of co-location is shared in the FF model (e.g., 3.6% of all time instants when both players are *viewers*), when players find it most beneficial to report few co-locations and report their location most often (at no privacy cost). Furthermore, the equilibria decisions are frequently socially-optimal: From 52% of the times (in the FF model, when both Alice and Bob are *viewers*) to 85% of the times (in the CF model, when both Alice and Bob are *sharers*). Regardless of the adversary, the most socially-optimal equilibria are reached when both players are *sharers* and the least when both players are *viewers* (due to the fact that a *viewer* player shares less than a *sharer* player and, consequently, their opponent benefits less from their posts).

**The case of *benefits-oriented* / *privacy-oriented* players**. We present the case where the players have different values for the $f_{pb}$ factor. We select two subsets of preference profiles from our survey data: the *privacy-oriented* (158 profiles)–for which $f_{pb} > 0.5$–and the *benefits-oriented* (92 profiles)–for which $f_{pb} < 0.5$. We evaluate the outcome of the Sharing Game in three cases–when Alice is *privacy-oriented* and Bob is *benefits-oriented*, when both are *privacy-oriented* and when both are *benefits-oriented*–for each possible pairs of preference profiles.

Figure 11 illustrates our aggregated results (see caption for details). It is interesting that, when both players are *benefits-oriented*, the amount of shared co-location is substantial: It is *always shared* in the SP, MF and CF adversarial models (20% of all time instants), and is shared approximately 19.7% of all time instants in the FF model.[7] When one player is *benefits-oriented* and the other is *privacy-oriented*, **the amount of shared co-location varies significantly, with respect to the considered adversary**: It is always shared in the SP case, shared 5.4% of all time instants (27% of the time instants when the players are co-located) in the CF case, 10% of all time instants in the MF case and only 2% of all time instants in the FF case. One reason for this behavior is that the CF adversary has access to location information shared by both players, whereas the MF adversary only has access to location shared by one of them, so privacy losses stemming from shared co-locations are higher in the CF case, and thus less co-location information is shared. Interestingly, this also causes ***both* players** to **share their location *more frequently* in the CF case than in the MF case** (in the CF case, it is enough that one player share his location after a shared co-location, for both players' privacy to be damaged, so the other player would be *forced* to also share his location for some benefit). When both players are *privacy-oriented*, location sharing is substantially reduced, but co-location is still shared 15% of all time instants in the SP case. The FF case illustrates a naturally emerging countermeasure: In all the cases, players find it most beneficial to report few co-locations (unlinking themselves from their friend makes the information unavailable to the FF adversary) and report their location most often (at no privacy cost). The equilibria decisions are frequently socially-optimal: From 45% of the times (in the FF model, when Alice is *benefits-oriented* and Bob is *privacy-oriented*) to 99% of the times (in the FF model, when Alice and Bob are *benefits-oriented*). We notice that the case of players having opposite views regarding $f_{pb}$ is particularly problematic: Regardless of the considered adversary, this case presents the least amount of socially-optimal equilibria decisions; furthermore, the utility of the *benefits-oriented* player is rarely maximized because his opponent would seldom share or allow sharing; finally, misaligned preferences can lead to different decisions for the players–they only make the same decision 24% of the times in the SP model, 19.2% in the CF model and 11.6% in the MF model.

[7] To infer these numbers from Figure 11, we sum the values for "co-location" and "both". As discussed in Section 5.2.2, in any adversarial model, both players share the same amount of co-location.

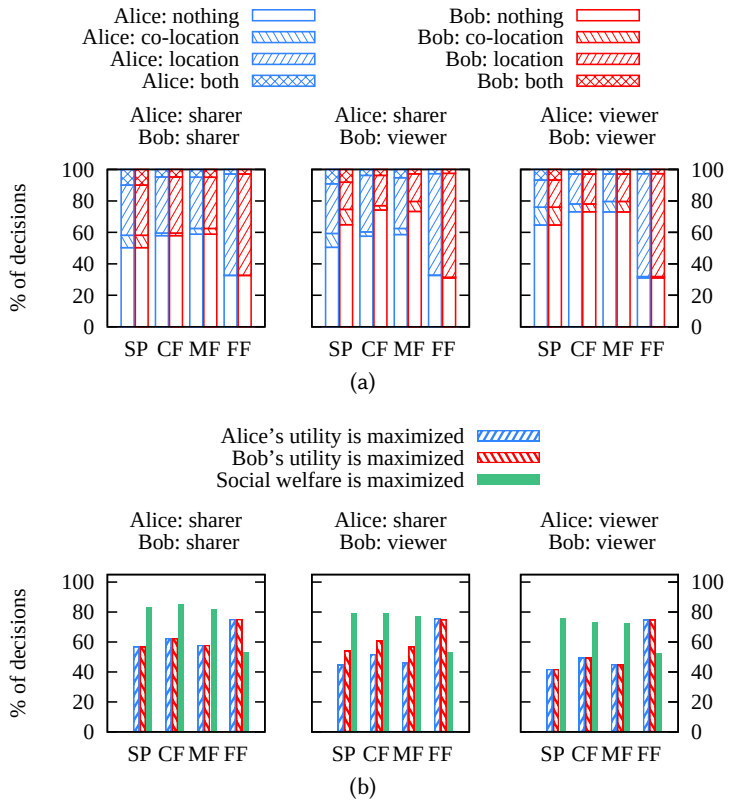



Fig. 10. Equilibria decisions (a) and their properties (b), when Alice and Bob have different preference profiles, corresponding to real survey data: 150 *sharers*' profiles ($f_{sv} > 0.5$) and 58 *viewers*' profiles ($f_{sv} < 0.5$). We present three scenarios: both Alice and Bob are *sharers* (left plots), Alice is a *sharer* and Bob is a *viewer* (middle plots) and both are *viewers* (right plots). Note that, due to the symmetry of the trajectories in the meeting scenario, the case where Alice is a *viewer* and Bob is a *sharer* is symmetric to the case where Alice is a *sharer* and Bob is a *viewer*. Different adversarial models (SP, CF, MF, FF) are illustrated on the $x$ axis. In each of the top three plots, for each adversarial model, two bars (blue on the left for Alice and red on the right for Bob) indicate–on the $y$ axis–the proportion of times (aggregated over time instants and the number of preference profile pairs considered in that scenario) a player made one of the four possible decisions: share nothing (empty pattern), share only location (hash right pattern), share only co-location (hash left pattern) or share both (hash right-left pattern). Each of the three bottom plots show, on the $y$ axis, for each adversarial model, the proportion of times social welfare and individual utilities are maximized.

## 5.3 Experimental Conclusions

We conclude that the considered adversary has a strong influence on the users' decisions, the value of $\alpha$ ($f_{pb}$, the preference for privacy versus social benefits of *one* user) can also influence *both* users' decisions, whereas other model parameters have a more moderate effect. We observed some interesting patterns of behavior, such as the fact that a *vicious-circle effect* can occur in the SP adversarial model: When a player (say Alice) has a strong incentive to share, it is enough that she share one co-location information and, with respect to the service provider, her friend (Bob)–who might not be willing to share at all–will continue to have his privacy affected and

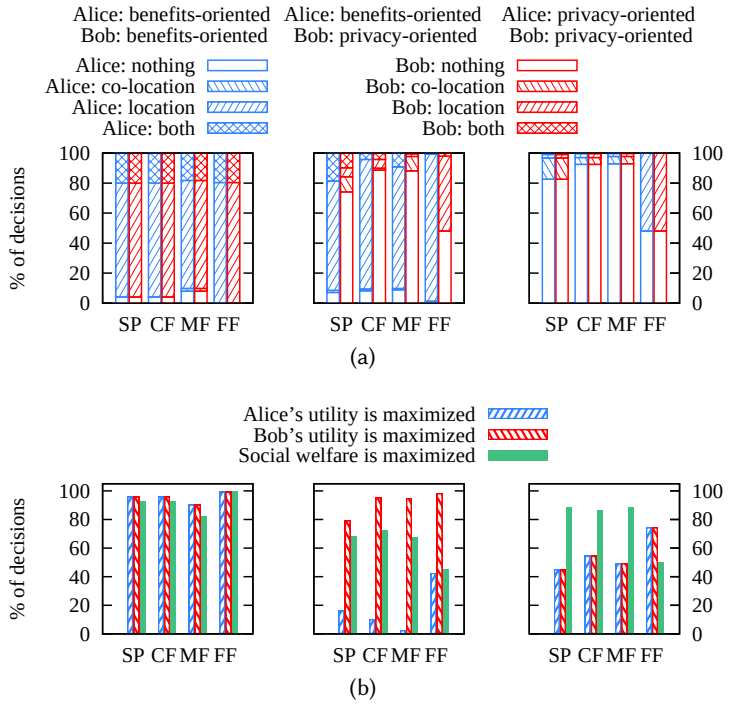



Fig. 11. Equilibria decisions (a) and their properties (b), when Alice and Bob have different preference profiles, corresponding to real survey data: 92 *benefits-oriented* profiles ($f_{pb} < 0.5$) and 158 *privacy-oriented* profiles ($f_{pb} > 0.5$). We present three scenarios: both Alice and Bob are *benefits-oriented* (left plots), Alice is *benefits-oriented* and Bob is *privacy-oriented* (middle plots) and both are *privacy-oriented* (right plots). Note that, due to the symmetry of the trajectories in the meeting scenario, the case where Alice is *privacy-oriented* and Bob is *benefits-oriented* is symmetric to the case where Alice is *benefits-oriented* and Bob is *privacy-oriented*.

be forced into sharing his location at later times. This effect is made even worse if Alice still wants to share her own location at other time instants, further damaging Bob's privacy. It is evident that a sequence of meeting scenarios between Alice and Bob, where Alice always shares their co-locations and sometimes shares her location, can force Bob into a sharing behavior as well; this is because the effect propagates not only in space (influences from a friend), but also in time. When the adversary is represented by the OSN friends, we observed that the effect of a shared co-location can eventually fade away: If Bob does not want Alice to share a co-location, he can un-tag himself and, assuming he does not want to share his location at later times, his privacy will be protected. However, we showed that it is possible (e.g., in the SP and CF models) that a (common) decision to share co-location create the incentive to *over-share locations after the time of the co-location*. This is an interesting finding from a design perspective for the OSN service providers: Building and advertising features that allow the sharing of co-location information, would also encourage users to share their locations more often. Finally, we noticed that, in the FF model, a natural tendency for privacy concerned players is to share few co-locations but they still share a significant amount of location information.

## 6 DISCUSSION

This work represents the first step towards modeling the interplay between users in the context of (co-)location sharing and the idea of combining a game-theoretic model with real user parameters in this setting is also novel. As the first attempt to tackle the problem of understanding and predicting users' interaction in such a complex context, we focused on a number of specific scenarios and assumptions, which open several interesting directions for future work.

(1) We considered players do not report fake co-locations. Including *fake co-locations* into the model is straightforward, and simply increases the number of possible strategies at times where players are not co-located. However, the benefit of sharing fake co-locations is likely different than that of sharing true co-locations; we plan to carry out surveys to evaluate the benefit users gain for sharing fake co-locations and include this kind of information into our model.
(2) We included in our evaluation location privacy evaluated with respect to one adversary at a time. As adversaries are not easily separable and a user is likely sensitive to more than one type of the adversaries that we mentioned, a *combination of* these *different adversaries* (with appropriate weights estimated through similar conjoint analysis surveys) can be included in the utility function.
(3) We considered players are somewhat short-sighted in that, in their utility function, only social benefits and privacy effects at the current time are considered. Having understood the basics of interplay at this level, the model can now be extended to include a combination of *benefits and privacy implications at all time instants* into the utility function–typically weighted by a discount factor, which makes the influence of a player's valuation of the game diminish with time (in other words, immediate benefits and privacy implications weigh more than those in the distant past). However, doing this would also require further surveys for quantification of users' discount factors. We plan to carry this out in future work.
(4) We considered a game with two players. In doing so, we illustrated interdependence effects that work both in time (actions at previous times influence a user's future decisions for sharing) and in space (actions of a friend influence a user's decisions). To better illustrate the spacial effects, the framework can be extended to *more players and non-default visibility settings for posts*; intuitively, we expect that the cascading effect (one user's behavior affecting that of her friends', that of the friends of her friends, and so on and so forth) would occur, with an even greater impact, with more players. However, the complicated dependencies introduced by co-location information makes defining appropriate utility functions for the general N-player T-time Sharing Game more challenging. In order to represent this game, as well as reduce the complexity of the computation of its Nash equilibria, we propose to use multi-agent influence diagrams (MAIDs), which were introduced by Koller et al. [24]. MAIDs were proposed for efficiently solving games, by a divide and conquer approach: first splitting the problem into sub-problems which are independent of each other (in the sense that they are strategically irrelevant to each other), and then combining their local equilibria to obtain a global game equilibrium. We plan to explore this in future work.
(5) We considered location check-ins are similar for a user, in the sense of including only the associated location privacy loss in the utility function. However, in practice (especially when combined with co-location), a check-in at a hotel vs. a check-in in a park could bare different meanings. To this end, solutions for predicting the utility/benefit[8] of check-ins (such as that proposed by Bilogrevic et al. [9]) can be used to associate a sensitivity factor with each check-in, in addition to the location privacy loss. Such a factor would typically be inversely proportional to the predicted utility/benefit of the check-in.
(6) Beyond location, co-locations are by themselves sensitive information and can be used, for instance, to infer private information about the users or expose social ties; our framework can be extended to support such

[8]Not to be confused with utility in the game theory sense. Utility here means benefit only.

*other quantifiable aspects of privacy* and incorporate them into the utility function, with an appropriate weight that can also be estimated through a conjoint analysis study.

(7) Our evaluation focused on the most interesting cases, while keeping the number of free parameters low. We only quantified a limited number of preference factors in order to avoid the questionnaire fatigue effect that would have decreased the quality of the participants' responses. *More preference factors can be evaluated through similar user surveys* (e.g., different values for $f_{pb}$ for the different adversarial models).

(8) Given the fact that social benefits and privacy are highly dependent on the culture, our sample of participants is not necessarily representative of the global population as the vast majority of Mechanical Turk workers are US-based [41]. To obtain more significant statistics, we intend to run user *surveys with a more diverse and targeted sample of participants*.

(9) Having studied the basic interplay of human behavior, the next step towards a more complex and realistic model would be to consider a cooperative/colaborative game-theoretic model (one can envision than a cooperative mentality applies to friends [11, 21, 27, 45, 46, 52]) or to include *altruism* into the users' utility functions. To some extent, altruism is already implicitly included in our framework: The benefits of sharing information include both the fact that users enjoy sharing, but also the fact that they are happy that their friends enjoy viewing their posts. We could extend this in two ways: (i) Users care about their friends' global utility; the utility of user $j$ should be included in the expression of user $i$'s, weighted by an altruistic factor as proposed in [30]; (ii) Users care about their friends' privacy (their friends' benefits are already included in the sharing benefits); the privacy of user $j$ should be included in the expression of user $i$'s, weighted by an altruistic factor that can be estimated as in [38]. We plan to carry it out in future work.

Ultimately, our extensible model with quantifiable parameters can serve as the first building block to assist and finally automate user decision making in an informed manner. Specifically, we envision that a software tool (e.g., a Facebook extension) would use our framework to take these interactions into account and assist users to improve their awareness and decision making process.

## 7 CONCLUSION

It is well-known that other people's behaviors affect our own privacy, in particular in the case of interdependent data. Yet, formalizing these complex interdependences and their implications is non-trivial, especially because human decisions play a dominant role. To address this issue, we focused on the (co-)location sharing features provided by major OSNs. We proposed a coarse-grained game-theoretic model and provided a first framework to study the interplay between two friends. A major challenge in such approaches is to assign meaningful values to the parameters that characterize user preferences. For this purpose, we carried out a survey of Facebook users, which also confirmed the anticipated high diversity of opinions in terms of social benefits and location privacy. We studied the resulting equilibria and their properties, in different settings. In particular, we showed how, because of conflicting preferences, one of the users can be forced into a situation that she does not desire and we demonstrated that sharing co-location information can additionally encourage users to over-share their locations. In addition to already mentioned future work, we intend to develop appropriate warning mechanisms to be run on smartphones; these would help users better understand and anticipate the consequences of their (co-)location sharing decisions.

## SURVEY TRANSCRIPT

In Figures 12 and 13, we provide the transcript of our survey.

## ACKNOWLEDGMENTS

The authors are thankful to Elisa Celis and Italo Dacosta for their useful insights. Parts of this work were carried out while Kévin Huguenin was with LAAS-CNRS/Université de Toulouse, France and Mathias Humbert was with Saarland University, Saarbrücken, Germany.

## REFERENCES

[1] Alessandro Acquisti. 2009. Nudging Privacy: The Behavioral Economics of Personal Information. *IEEE Security Privacy* 7, 6 (2009), 82–85. https://doi.org/10.1109/MSP.2009.163

[2] Alessandro Acquisti and Jens Grossklags. 2004. *Privacy Attitudes and Privacy Behavior*. Springer, 165–178. https://doi.org/10.1007/1-4020-8090-5_13

[3] A. Acquisti and J. Grossklags. 2005. Privacy and rationality in individual decision making. *IEEE Security Privacy* 3, 1 (Jan 2005), 26–33. https://doi.org/10.1109/MSP.2005.22

[4] Lars Backstrom, Eric Sun, and Cameron Marlow. 2010. Find me if you can: improving geographical prediction with social and spatial proximity. In *WWW*. ACM, 61. https://doi.org/10.1145/1772690.1772698

[5] Fabrício Benevenuto, Tiago Rodrigues, Meeyoung Cha, and Virgílio Almeida. 2009. Characterizing User Behavior in Online Social Networks. In *IMC*. ACM, 49–62. https://doi.org/10.1145/1644893.1644900

[6] Andrew Besmer and Heather Richter Lipford. 2010. Moving beyond untagging: photo privacy in a tagged world. In *Proc. of SIGCHI*. ACM.

[7] Gergely Biczók and Pern Hui Chia. 2013. Interdependent Privacy: Let Me Share Your Data. In *FC*. Springer, 338–353. https://doi.org/10.1007/978-3-642-39884-1_29

[8] Igor Bilogrevic, Kévin Huguenin, Berker Ağır, Murtuza Jadliwala, Maria Gazaki, and Jean-Pierre Hubaux. 2015. A Machine-Learning Based Approach to Privacy-Aware Information-Sharing in Mobile Social Networks. *Pervasive and Mobile Computing (PMC)* (Nov. 2015), 1–12.

[9] Igor Bilogrevic, Kévin Huguenin, Stefan Mihaila, Reza Shokri, and Jean-Pierre Hubaux. 2015. Predicting users' motivations behind location check-ins and utility implications of privacy protection mechanisms. In *22nd Network and Distributed System Security Symposium (NDSS)*.

[10] A. Chaabane, G. Acs, and M.A. Kaafar. 2012. You Are What You Like! Information Leakage Through Users' Interests. In *NDSS*.

[11] Hichang Cho and Anna Filippova. 2016. Networked privacy management in Facebook: A mixed-methods and multinational study. In *Proc. of CSCW*. ACM.

[12] D.J. Crandall, L. Backstrom, D. Cosley, S. Suri, D. Huttenlocher, and J. Kleinberg. 2010. Inferring social ties from geographic coincidences. *Proc. of PNAS* 107 (2010).

[13] R. Dey, Cong Tang, K. Ross, and N. Saxena. 2012. Estimating age privacy leakage in online social networks. In *INFOCOM*. IEEE, 2836–2840. https://doi.org/10.1109/INFCOM.2012.6195711

[14] Drew Fudenberg and Jean Tirole. 1991. *Game theory*. MIT press.

[15] Lorena González-Manzano, Ana I. González-Tablas, José M. de Fuentes, and Arturo Ribagorda. 2014. CooPeD: Co-owned Personal Data management. *Computers & Security* 47 (2014).

[16] Paul E. Green and V. Srinivasan. 1978. Conjoint Analysis in Consumer Research: Issues and Outlook. *Journal of Consumer Research* 5, 2 (1978), 103–123.

[17] Benjamin Henne, Christian Szongott, and Matthew Smith. 2013. SnapMe if You Can: Privacy Threats of Other Peoples' Geo-tagged Media and What We Can Do About It. In *WiSec*. ACM, 95–106. https://doi.org/10.1145/2462096.2462113

[18] Mathias Humbert, Erman Ayday, Jean-Pierre Hubaux, and Amalio Telenti. 2013. Addressing the Concerns of the Lacks Family: Quantification of Kin Genomic Privacy. In *CCS'13: Proc. of the 20th ACM Conf. on Computer and Communications Security*. ACM, 1141–1152. https://doi.org/10.1145/2508859.2516707

[19] Mathias Humbert, Erman Ayday, Jean-Pierre Hubaux, and Amalio Telenti. 2015. On non-cooperative genomic privacy. In *FC*. Springer, 407–426. https://doi.org/10.1007/978-3-662-47854-7_24

[20] Panagiotis Ilia, Iasonas Polakis, Elias Athanasopoulos, Federico Maggi, and Sotiris Ioannidis. 2015. Face/Off: Preventing Privacy Leakage From Photos in Social Networks. In *CCS*. ACM, 781–792. https://doi.org/10.1145/2810103.2813603

[21] Haiyan Jia and Heng Xu. 2016. Autonomous and interdependent: Collaborative privacy management on social networking sites. In *Proc. of CHI*. ACM.

[22] Spyros Kokolakis. 2015. Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers Security* (2015). https://doi.org/10.1016/j.cose.2015.07.002

[23] Daphne Koller and Nir Friedman. 2009. *Probabilistic graphical models: principles and techniques*. MIT press.

[24] Daphne Koller and Brian Milch. 2003. Multi-agent influence diagrams for representing and solving games. *Games and economic behavior* 45, 1 (2003), 181–221.
[25] Hanna Krasnova, Sarah Spiekermann, Ksenia Koroleva, and Thomas Hildebrand. 2010. Online social networks: Why we disclose. *Journal of information technology* 25, 2 (2010), 109–125.
[26] John Krumm. 2009. A survey of computational location privacy. *Personal and Ubiquitous Computing* 13, 6 (Aug. 2009), 391–399. https://doi.org/10.1007/s00779-008-0212-5
[27] Airi Lampinen, Vilma Lehtinen, Asko Lehmuskallio, and Sakari Tamminen. 2011. We're in it together: interpersonal management of disclosure in social network services. In *Proc. of CHI*. ACM.
[28] Aron Laszka, Mark Felegyhazi, and Levente Buttyan. 2015. A survey of interdependent information security games. *Comput. Surveys* 47, 2 (2015), 23. https://doi.org/10.1145/2635673
[29] R.L. Mason, R.F. Gunst, and J.L. Hess. 2003. *Statistical design and analysis of experiments with applications to engineering and science*. J. Wiley.
[30] Dominic Meier, Yvonne Anne Oswald, Stefan Schmid, and Roger Wattenhofer. 2008. On the Windfall of Friendship: Inoculation Strategies on Social Networks. In *EC*. ACM, 294–301. https://doi.org/10.1145/1386790.1386836
[31] Alan Mislove, Bimal Viswanath, Krishna P. Gummadi, and Peter Druschel. 2010. You Are Who You Know: Inferring User Profiles in Online Social Networks. In *WSDM*. ACM, 251–260. https://doi.org/10.1145/1718487.1718519
[32] Roger B Myerson. 2013. *Game theory*. Harvard university press.
[33] A. Noulas, M. Musolesi, M. Pontil, and C. Mascolo. 2009. Inferring interests from mobility and social interactions. In *Proc. of NIPS Workshops*.
[34] Alexandra-Mihaela Olteanu, Kevin Huguenin, Reza Shokri, Mathias Humbert, and Jean-Pierre Hubaux. 2016. Quantifying Interdependent Privacy Risks with Location Data. In *IEEE Trans. Mobile Comput.* https://doi.org/10.1109/TMC.2016.2561281
[35] Xinru Page, Bart P. Knijnenburg, and Alfred Kobsa. 2013. FYI: Communication Style Preferences Underlie Differences in Location-sharing Adoption and Usage. In *UbiComp*. ACM, 153–162. https://doi.org/10.1145/2493432.2493487
[36] Yu Pu and Jens Grossklags. 2014. An economic model and simulation results of app adoption decisions on networks with interdependent privacy consequences. In *GameSec*. Springer, 246–265. https://doi.org/10.1007/978-3-319-12601-2_14
[37] Yu Pu and Jens Grossklags. 2015. Towards a Model on the Factors Influencing Social App Users' Valuation of Interdependent Privacy. *PoPETS* 2016, 2 (2015), 61–81. https://doi.org/10.1515/popets-2016-0005
[38] Yu Pu and Jens Grossklags. 2015. Using Conjoint Analysis to Investigate the Value of Interdependent Privacy in Social App Adoption Scenarios. In *ICIS*. Association for Information Systems, 20.
[39] Yu Pu and Jens Grossklags. 2017. Valuating Friends' Privacy: Does Anonymity of Sharing Personal Data Matter?. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*. USENIX Association.
[40] Christopher Riederer, Daniel Echickson, Stephanie Huang, and Augustin Chaintreau. 2016. FindYou: A Personal Location Privacy Auditing Tool. In *WWW*. 243–246. https://doi.org/10.1145/2872518.2890546
[41] Joel Ross, Lilly Irani, M. Six Silberman, Andrew Zaldivar, and Bill Tomlinson. 2010. Who Are the Crowdworkers?: Shifting Demographics in Mechanical Turk. In *CHI*. ACM, 2863–2872. https://doi.org/10.1145/1753846.1753873
[42] Reza Shokri. 2015. Privacy games: Optimal user-centric data obfuscation. *PoPETS* 2015, 2 (2015), 299–315.
[43] Reza Shokri, George Theodorakopoulos, Carmela Troncoso, Jean-Pierre Hubaux, and Jean-Yves Le Boudec. 2012. Protecting location privacy: optimal strategy against localization attacks. In *CCS'12: Proc. of the 19th ACM Conf. on Computer and Communications Security*. ACM, 617–627. https://doi.org/10.1145/2382196.2382261
[44] Yan Shoshitaishvili, Christopher Kruegel, and Giovanni Vigna. 2015. Portrait of a Privacy Invasion; Detecting Relationships Through Large-scale Photo Analysis. *PoPETS* 2015, 1 (2015), 41–60. https://doi.org/10.1515/popets-2015-0004
[45] J. M. Such and N. Criado. 2016. Resolving Multi-Party Privacy Conflicts in Social Media. *IEEE KDE* (2016).
[46] Jose M. Such, Joel Porter, Sören Preibusch, and Adam Joinson. 2017. Photo Privacy Conflicts in Social Media: A Large-scale Empirical Study. In *Proc. of CHI*.
[47] Eran Toch, Justin Cranshaw, Paul Hankes Drielsma, Janice Y. Tsai, Patrick Gage Kelley, James Springfield, Lorrie Cranor, Jason Hong, and Norman Sadeh. 2010. Empirical Models of Privacy in Location Sharing. In *UbiComp*. ACM, 129–138. https://doi.org/10.1145/1864349.1864364
[48] John Von Neumann and Oskar Morgenstern. 2007. *Theory of games and economic behavior*. Princeton university press.
[49] Nevena Vratonjic, Kévin Huguenin, Vincent Bindschaedler, and Jean-Pierre Hubaux. 2014. A Location-Privacy Threat Stemming from the Use of Shared Public IP. *IEEE Trans. on Mobile Computing (TMC)* 13, 11 (Nov. 2014), 2445–2457. https://doi.org/10.1109/TMC.2014.2309953
[50] Gang Wang, Sarita Y. Schoenebeck, Haitao Zheng, and Ben Y. Zhao. 2016. "Will Check-in for Badges": Understanding Bias and Misbehavior on Location-based Social Networks. In *ICWSM*. AAAI.
[51] Jason Wiese, Patrick Gage Kelley, Lorrie Faith Cranor, Laura Dabbish, Jason I. Hong, and John Zimmerman. 2011. Are You Close with Me? Are You Nearby?: Investigating Social Groups, Closeness, and Willingness to Share. In *UbiComp*. ACM, 197–206. https://doi.org/10.1145/2030112.2030140

[52] Pamela Wisniewski, Heather Lipford, and David Wilson. 2012. Fighting for my space: Coping mechanisms for SNS boundary regulation. In *Proc. of SIGCHI*. ACM.
[53] xlstat 2016. XLSTAT statistical software for Microsoft Excel. https://www.xlstat.com/en/. (2016). last visited: Aug. 2016.
[54] Fengli Xu, Zhen Tu, Yong Li, Pengyu Zhang, Xiaoming Fu, and Depeng Jin. 2017. Trajectory Recovery From Ash: User Privacy Is NOT Preserved in Aggregated Mobility Data. In *Proceedings of the 26th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 1241–1250.
[55] Heng Xu. 2011. Reframing privacy 2.0 in online social network. *U. Pa. J. Const. L.* (2011).
[56] Mu Yang, Yijun Yu, Arosha K. Bandara, and Bashar Nuseibeh. 2014. Adaptive Sharing for Online Social Networks: A Trade-off between Privacy Risk and Social Benefit. In *TrustCom*. IEEE, 45–52. https://doi.org/10.1109/TrustCom.2014.10

Received August 2017

**Part I: Demographics**

(1) What is your gender?
◯ Female ◯ Male

(2) What is your age? [ ]

(3) What is your primary area of employment?
◯ Homemaker ◯ Retired ◯ Student (undergraduate) [...] ◯ Transportation
◯ Other: [ ]

**Part II: Preferences**

(4) A check-in post is a post in which location is disclosed, by checking-in at a point of interest like an airport, concert hall, square etc. Imagine that, due to technical constraints, Facebook may have to remove some or all of your 2 most recent check-in posts (your friends will not see these posts anymore) and/or some or all of your close friends' 2 most recent check-in posts (you will not see these posts anymore). Note that there are **posts you and your friends already shared** therefore you do not want Facebook to delete any of them! (choose option "2 of your recent posts are kept and 2 of your friends' recent posts are kept" as most preferred). Order the following scenarios in decreasing order of preference. Click on an item in the list on the left, starting with your highest ranking item, moving through to your lowest ranking item.

**Your choices** | **Your ranking**

- 2 of your recent posts are kept and 2 of your friends' recent posts are kept
- 1 of your recent posts are kept and 2 of your friends' recent posts are kept
- 2 of your recent posts are kept and 1 of your friends' recent posts are kept
- 2 of your recent posts are kept and 0 of your friends' recent posts are kept
- 0 of your recent posts are kept and 1 of your friends' recent posts are kept
- 1 of your recent posts are kept and 1 of your friends' recent posts are kept

(5) A check-in post is a post in which location is disclosed. A co-location post is a post in which you tag the friends you are with - either through a status message or a picture. Imagine that, due to technical constraints, Facebook may have to remove some or all of your 2 most recent check-in posts and/or some or all of your 2 most recent co-location posts (think of posts in which you either tag friends, or check-in, but not both). If removed, your friends will not see these posts anymore. Note that there are **posts you already shared** therefore you do not want Facebook to delete any of them! (choose option "2 of your recent check-in posts are kept and 2 of your recent co-location posts are kept" as most preferred). Order the following scenarios in decreasing order of preference. Click on an item in the list on the left, starting with your highest ranking item, moving through to your lowest ranking.

**Your choices** | **Your ranking**

- 2 of your recent check-in posts are kept and 0 of your recent co-location posts are kept
- 0 of your recent check-in posts are kept and 1 of your recent co-location posts are kept
- 1 of your recent check-in posts is kept and 1 of your recent co-location posts are kept
- 1 of your recent check-in posts is kept and 2 of your recent co-location posts are kept
- 2 of your recent check-in posts are kept and 1 of your recent co-location posts are kept
- 2 of your recent check-in posts are kept and 2 of your recent co-location posts are kept

Fig. 12. Transcript of our survey questionnaire (1/2).

(6) We define location privacy as the precision with which someone (Facebook, your friends, or public observers) can guess your location at any moment during the day. An average location privacy of 50 meters means that at any time during the day, your location can be guessed as close as 50 meters from your real location. With each of your check-in posts, your location privacy can change.


Order the following scenarios in decreasing order of preference. Click on an item in the list on the left, starting with your highest ranking item, moving through to your lowest ranking item.

| **Your choices*** | **Your ranking** |
|---|---|
| 19 posts for an average privacy of 200m | |
| 12 posts for an average privacy of 400m | |
| 5 posts for an average privacy of 830m | |
| 24 posts for an average privacy of 0m | |
| 0 posts for an average privacy of 1100m (1,1km) | |
| 10 posts for an average privacy of 610m | |

* These numbers were extracted from the experimental results presented in [34].

**Part III: Social Networks Usage**

(7) On average how many times per week do you use Facebook?
◯ Several times per day ◯ One time per day ◯ A few days per week ◯ One time per week ◯ Less than one time per week

(8) On average how many times per week do you check-in on Facebook? (A check-in post is a post in which location is disclosed.)
◯ More than one time per day ◯ One time per day ◯ Once every few days ◯ Once per week ◯ Less than one time per week

(9) On average how many times per week do you tag the friends that are with you on Facebook, in pictures or in statuses?
◯ More than one time per day ◯ One time per day ◯ Once every few days ◯ Once per week ◯ Less than one time per week

(10) How concerned are you about location privacy (i.e., the fact that someone can infer your more or less precise location at some points in time)?
◯ Very concerned ◯ Moderately concerned ◯ Not concerned

(11) Were you aware that check-ins or tagging your friends can decrease your location privacy and your friends' location privacy?
◯ I was aware they would impact my own privacy as well as my friends' privacy
◯ I was only aware they would impact my own privacy
◯ I was not aware they have any effect on privacy

(12) Were you aware that the check-ins and tags that your friends post can decrease your location privacy?
◯ Yes ◯ No

(13) Imagine that you are at a venue with a friend, who just checked-in at this venue and tagged you in his post. In terms of your location privacy, whom are you concerned about?
☐ The friends that you have in common on Facebook ☐ Your other friends on Facebook (these are not friends of your friend)
☐ Your friend's other friends on Facebook (these are not your friends) ☐ Facebook ☐ None of the above

(14) Will the information you learned through this survey change your behavior on Facebook in any way? If so how?

Fig. 13. Transcript of our survey questionnaire (2/2).