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

Alexandra-Mihaela Olteanu¹, Kévin Huguenin², Mathias Humbert³, and Jean-Pierre Hubaux¹

¹EPFL, Lausanne, Switzerland
²LAAS-CNRS, Université de Toulouse, CNRS, France
³CISPA, Saarland University, Saarbrücken, Germany

Abstract—Most popular location-based social networks, such as Facebook and Foursquare, let their 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’ (location) privacy, as co-location information introduces interdependences between users. We propose a game theoretical framework for analyzing and predicting the strategic behaviors, in terms of (co)-location information sharing, of users of social networks. In addition, in order to design parametric utility functions that are representative of the users’ actual preferences, we conduct a survey of Facebook users (N=250) and rely on conjoint analysis for quantifying 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 others’ 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. We also identify interesting insights that could be instrumental in the design of the next-generation privacy protection systems.

1. Introduction

Major online social network (OSN) providers, such as Facebook, understood early 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, while on the go, their favorite OSNs from their smartphones. 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. [1] perform a user study that demonstrates that 84.7% of 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 in fact co-located. As for location information, sharing such co-location information—the fact that two users are together (the actual location might not be known)—brings social benefits 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 concerns regarding privacy. Although it has been known for years that location information raises severe privacy issues—this has been extensively studied in the literature (e.g., [2]–[4]; see also FindYou [5], a personal 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 [6]. A singular and critical aspect of co-location is that they 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 own privacy and it becomes complex to predict the optimal behavior. Such interdependent privacy risks¹ are particularly problematic if users have different, possibly opposite, views about location sharing and privacy.

We propose a 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 and conjoint analysis. 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]–[10] 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

1. The term was coined by Biczók and Chia in [7].
We survey the related work. Finally, in Section 7, we conclude that we describe the considered setting and the system model, and we design and conduct a user survey of Facebook users (N=250) to quantify users’ preferences of (1) sharing or viewing posts, (2) sharing 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 theoretical 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 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 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 on social networks. In Section 3, 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 4, we evaluate our framework in a number of different scenarios. In Section 5, we list the limitations of our approach, model and evaluation and we give leads for improvement. In Section 6, we survey the related work. Finally, in Section 7, we conclude the paper and we discuss future work.

2. System Model and Formalization

In our study, 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, in a post 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 or picture 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 online 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, as well as the decision to allow a friend to post co-location information, is made strategically by both the users involved.

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) sharing 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 theoretical 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 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 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 on social networks. In Section 3, 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 4, we evaluate our framework in a number of different scenarios. In Section 5, we list the limitations of our approach, model and evaluation and we give leads for improvement. In Section 6, we survey the related work. Finally, in Section 7, we conclude the paper and we discuss future work.

2.1. Game Theory 101

We briefly introduce the relevant concepts of game theory in this paragraph, which can be skipped by the knowledgeable reader. Game theory is the study of the strategic interaction between multiple rational decision-makers who aim to maximize their own utility [11]–[13]. 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 enables us to predict 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 used in dynamic games 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.
2.2. User Model

We are aware of the irrationality of people, especially when they have to make privacy-related decisions [14]. Nevertheless, we believe that privacy awareness will increase, notably because (unfortunately) a growing number of people suffer the consequences of their (and others’) privacy carelessness. Moreover, it is very likely that smartphones will be more and more involved in the sharing decisions we make, as it is already demonstrated by the growing sophistication of the applications’ permission systems. A privacy-protection software run for this purpose can be “rational” and strictly follow the parametrisation model provided by its user. It is therefore of high interest to see what happens under the assumption of rationality.

With this assumption in mind, 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, …, T\}, which we call the Sharing Game. 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 representing whether user \(i\) shares location and co-location, respectively. For alternate more compact notations, we use \(\hat{L}\) for \(sl_i(t) = 0\), \(L\) for \(sl_i(t) = 1\), \(\hat{C}\) 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: \(\hat{L}\hat{C}\) – sharing nothing, \(\hat{L}C\) – sharing only the co-location information, \(\hat{L}\hat{C}\) – 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: \(\hat{L}\hat{C}\) – sharing nothing and \(\hat{L}C\) – sharing location information.

The social benefits of user \(i\) that stem from a decision to share information at some time \(t\) are denoted by \(B_i(t, s_i(t), s_j(t))\). Her privacy at \(t\), denoted by \(P_i(t, s_i(t), s_j(t), \mathcal{B})\), is a function of both users’ strategies at times \(\{t - k, \ldots, t\}\), 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. While 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 a certain time instant \(t\) captures both her social benefits and her privacy. We assume a player is only interested in the privacy at the current time she has to make a decision, \(t\). Specifically,

\[
u_i(t, s_i(\cdot), s_j(\cdot)) = (1 - \alpha_i) B_i(t, s_i(t), s_j(t)) + \alpha_i P_i(t, s_i(\cdot), s_j(\cdot), \mathcal{B}) \tag{1}\]

where \(\cdot\) denotes the times \(\{t - k, \ldots, t\}\) 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 \(P(\cdot)\), of their decision regarding sharing.

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,

\[
B_i(t, s_i(t), s_j(t)) = \frac{b_{is}^l sl_i(t) + b_{is}^p sc_i(t) + b_{is}^v sl_j(t) + b_{is}^v sc_j(t)}{b_{is}^l + b_{is}^p + b_{is}^v + b_{is}^v} \tag{2}
\]

We present our methodology for estimating these parameters in the next section. Note that the values of these parameters are specific to a user.

The game is played repeatedly by the two players, 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 of a typical OSN. The players’ actions are ordered: The second player knows the choice of the first player and decides her strategy accordingly.

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

<table>
<thead>
<tr>
<th>t_i(t) = (t_i(t), s_i(t))</th>
<th>Strategy of user i at time t</th>
</tr>
</thead>
<tbody>
<tr>
<td>L or s_i(t) = 1 (True)</td>
<td>Share location</td>
</tr>
<tr>
<td>C or s_i(t) = 0 (False)</td>
<td>Hide location</td>
</tr>
<tr>
<td>C or s_i(t) = 1 (True)</td>
<td>Share co-location</td>
</tr>
</tbody>
</table>

2.3. Adversarial Models

Privacy 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, namely 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.

2.3.1. 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 adversarial model is that, once either of the players shares their co-location, this information is always known to the adversary. In other words, the second player cannot un-share co-location information with respect to the service provider.

2.3.2. 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.

2.4. Analysis Methodology

At each time instant, t, we use backward induction to find a SPNE that dictates the players’ decisions. Equilibria decisions made at time instants prior to t, denoted by s∗(t') = (s∗i(t'), s∗j(t')), where t' > t, are used when computing the privacy of the players. The first player, player i, anticipates the second player’s (player j’s) best response, as a function of her possible strategies s_j, essentially

∀s_i, s∗_j(s_j) = arg max_s_j u_j(t, s_i, s_j, s∗_i(t-1), ..., s∗_j(t-k)) (3)

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(t) = arg max_s_i u_i(t, s_i, s∗_j(s), s∗_i(t-1), ..., s∗_i(t-k)) (4)

The equilibrium decisions at time t are then given by

s∗(t) = (s∗_i(t), s∗_j(t)) = (s∗_i(t), s∗_j(s∗_j(t))) (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(s), s_j(s)) = u_i(t, s_i(s), s_j(s)) + u_j(t, s_i(s), s_j(s)) (6)

In case of multiple equilibria at time t, we assume 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’ previous decisions.

2.5. Equilibria Properties

We are interested in different properties for the players’ equilibria decisions.
2.5.1. 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:

$$\forall (s_i(t), s_j(t)) \neq (s'_i(t), s'_j(t)) :$$

$$SW(t, s'_i(t), s'_j(t), s(t−1), ..., s(t−k)) \geq SW(t, s_i(t), s_j(t), s(t−1), ..., s(t−k))$$ (7)

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

$$\forall (s_i(t), s_j(t)) \neq (s'_i(t), s'_j(t)) :$$

$$u_i(t, s'_i(t), s'_j(t), s(t−1), ..., s(t−k)) \geq u_i(t, s_i(t), s_j(t), s(t−1), ..., s(t−k))$$ (8)

In our evaluation, we consider the proportion of time instants, across $\{1, \ldots, T\}$, for which the equilibria decisions are socially-optimal and the proportion of time instants for which the equilibria decisions maximize each player’s utility.

3. 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 user survey of Facebook users in early 2016.

3.1. Conjoint Analysis 101

We briefly introduce the conjoint analysis technique in this paragraph, which can be skipped by the knowledgeable reader. Conjoint analysis [15] 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 [16]. 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.

3.2. Methodology

We recruited the survey participants through the Amazon Mechanical Turk platform. To be eligible, participants 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.²

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 [15] and making use of a dedicated tool, namely the XLSTAT statistical software for Microsoft Excel [17]. Such an 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”) (see Appendix). The participants were asked to take into account only benefit considerations (i.e., not privacy). 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 level 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

² We used the information from the participants’ Facebook account only for screening purposes and to prevent participants from taking the survey multiple times. We did not store any such information.
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 $f_{vc}$, 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 their preferences, 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 [6]. 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 full transcript of the survey questionnaire is available in the Appendix.

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 of the survey. 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.

### 3.3. Results

We extracted the aforementioned three preference factors from the survey data by using XLSTAT. Note that, due to the employed methodology, the preference factors can take only a limited number of values; this number is determined by the possible ordering of the scenarios presented and their number. We illustrate the relevant statistics in Table 2 and in Figure 4. Note that these results should be taken with a pinch of salt as previous works (e.g., [18], [19]) 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; 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 is substantial and almost as large as the proportion of users who prefer viewing over sharing. 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 [20], [21].

<table>
<thead>
<tr>
<th>Average &amp; Standard Deviation</th>
<th>$f_{vc}$</th>
<th>$f_{sv}$</th>
<th>$f_{pb}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_{sv}$ &gt; 0.5 (prefer sharing/location/privacy)</td>
<td>60%</td>
<td>54%</td>
<td>63.2%</td>
</tr>
<tr>
<td>$p_{sv}$ = 0.5 (indifferent)</td>
<td>16.8%</td>
<td>20%</td>
<td>N/A</td>
</tr>
<tr>
<td>$p_{sv}$ &lt; 0.5 (prefer viewing/co-location/benefits)</td>
<td>23.2%</td>
<td>26%</td>
<td>36.8%</td>
</tr>
</tbody>
</table>

Table 2. User Preference Factors Extracted from the Survey Data by Using a Conjoint-Analysis Approach. $f_x$ denotes, depending on the column, $f_{sv}$, $f_{vc}$, or $f_{pb}$.

Figure 4. CDFs of participants’ preference factors extracted from our survey, using conjoint analysis.

As for the questions related to privacy issues on Facebook, 24.8% of the participants declared being “very concerned” about privacy, and 50% declared being “moderately concerned” (as illustrated in Figure 5a). When the participants report being co-located with a friend (say Bob), their most 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%), and Facebook (i.e., the SP model, 24.4%), as illustrated in Figure 5b. 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 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 at https://www.dropbox.com/s/4ukfg5nbhd1x0p7/data.zip?dl=1 (password Fb$250).

4. Evaluation

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

4.1. Experimental Setup

In this section, we describe the experimental setups of the different building blocks of our framework, including the location privacy function and the parameters of the utility function.

4.1.1. 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. [6]; 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 [22] 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 the “Model” section, 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: 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-locations 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).

4.1.2. 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 $\alpha$, $b_{dl}$, $b_{sc}$, $b_{sl}$ and $b_{vc}$ from the values of the preference factors $f_{pb}$, $f_{lc}$ and $f_{sv}$. To do so in a coherent way, while keeping 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:

$$\alpha = f_{pb}$$

$$b_{sl} = \frac{f_{sv}}{1-f_{sv}} \cdot \frac{f_{lc}}{1-f_{lc}} b_{vc}$$

$$b_{sc} = \frac{f_{sv}}{1-f_{sv}} b_{vc}$$

$$b_{vl} = \frac{f_{sv}}{1-f_{sv}} b_{vc}$$

(9)

where $b_{vc}$ is a free variable that can be set to 1.
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, 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 four (or less) neighboring cells (with the remaining equal probabilities).

4.2. Experimental Results

In this section, 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.

4.2.1. 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 \( f_{AV}, f_L \) and \( f_{PB} \) obtained in our survey, as presented in Table 2. Figure 8 illustrates the different game outcomes, for the four adversarial models we presented in Section 2.3 (see caption for details). 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 in Figures 8a and 8b. 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 corresponding green triangle for \( t = 1 \)), but either of the two players would have a higher utility (both the blue rectangle and the red circle are empty) if the other player shared their location (because they would enjoy viewing where their friend is without any privacy cost to themselves due to the current absence of correlation). 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 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 \((LC, LC)\) 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 \((LC, LC)\): Sharing only a co-location does come with a small
privacy cost, 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 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 $(L\bar{C}, ar{L}C)$: In this case, sharing their own location brings them some social benefits without any privacy costs (because this adversary, not being in their friends list, cannot see if they share location). When players are co-located, the equilibrium could be either $(L\bar{C}, L\bar{C})$ or $(\bar{L}C, L\bar{C})^4$, depending on how highly players value privacy over benefits. In this homogeneous scenario, where $f_{pb} = 0.6$, the equilibrium is $(L\bar{C}, L\bar{C})$ (sharing both location and co-location—$(L\bar{C}, L\bar{C})$—would yield minimal privacy for both players), and it maximizes both the social welfare and the players’ individual utilities.

4.2.2. The effect of privacy vs. benefits preferences ($f_{pb}$). We present a heterogeneous scenario, where players place different importance on privacy and social benefits. We consider the average values for $f_{sa}$ 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) 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, while Bob only shares his location).

An interesting observation is that, in the SP model, when the two players are co-located, the equilibrium strategies are always in the form of $(\bar{L}C, L\bar{C})$, $(L\bar{C}, L\bar{C})$ or $(L\bar{C}, L\bar{C})$. 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 $(LC, LC)$—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 $t = 1^4$), 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).

4.2.3. 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 simplify this task, 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.

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 $t = 1^4$), 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).

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 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

6. This detail is not directly readable form Figure 9, as it presents statistics aggregated over time instants. We omit a detailed illustration of the equilibrium decisions at each time instant due to lack of space.
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 privacy-oriented / benefits-oriented players. We study how the fact that the players have different values for the $f_{pb}$ preference factor affects their decisions. 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.

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 (we do not illustrate this due to lack of space).

Figure 11. Equilibria decisions (top) and their properties (bottom), 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.

4.3. Discussion

We conclude that the considered adversary has a strong influence on the users’ decisions, the value of $\alpha$ ($f_{sv}$, the preference for privacy versus social benefits of one user) can also influence both users’ decisions, whereas other model parameters such as $f_{sv}$ (preference of sharing versus viewing) 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 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

In order to infer these numbers from Figure 11, we sum the values for “co-location” and “both”. Note that, in any adversarial model, both players share the same amount of co-location (details discussed in Section 4.2.2).
privacy. It is easy to imagine how a sequence of meeting scenarios between Alice and Bob, where Alice always shares their co-location and sometimes shares her location, can force Bob into a sharing behavior as well. 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 frequently share their locations.

5. Limitations

In this section, we discuss the limitations of our work; it should be noted that this work is a first step towards modeling the interplay between users in the context of (co-)location sharing. Overall, we made a number of simplifying assumptions to better understand the effect of a limited number of parameters. First, in order to validate our framework and better understand the effect of the parameters on the interaction between users, we considered a game with only two players. The framework can be extended to more than two players in a straightforward way; we plan to carry it out in future work. 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) will also occur with more players, with an even greater impact.

Second, in our study, the privacy implications that stem from sharing co-locations are limited to location privacy, which is not the case in practice: Beyond location, co-locations are by themselves sensitive information and can be used, for instance, to infer private information about the users. These types of privacy risks could also be considered.

Third, users are not always fully rational (especially in terms of privacy decisions [14]) or selfish. In addition, when it comes to privacy, users’ attitudes often differ from their behaviors [18], [19]. Considering a cooperative game theoretical model or including altruism into the users’ utility functions would be one step towards a more realistic modeling of human behavior. 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: (1) 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 [23]; (2) 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 [10].

Fourth, in the evaluation, we made a number of assumptions regarding the relationships between the different parameters of our model. We did so to keep the number of free parameters low—which facilitates our evaluation—and to avoid the questionnaire fatigue effect that would have decreased the quality of the participants’ responses in our survey. In future work, we will evaluate more preference factors through similar user surveys (e.g., different values for $f_{pb}$ for the different adversarial models). Another limitation of our study is the fact that our sample of participants is not necessarily representative (Mechanical Turk workers are mostly US-based [24]), especially given the fact that social benefits and privacy are highly dependent on the culture. To obtain more significant statistics, we also intend to run user surveys with a more diverse sample of participants.

Finally, we only considered a simple, synthetic mobility trace in our evaluation. The rationale behind this choice is to study the problem on a canonical example (i.e., a “meeting scenario”) in order to understand the basics of the interplay between the users, independently from the specifics and the singularities of their individual data.

6. Related Work

Our work is related to two broad research areas: (1) information sharing on OSN–privacy and benefits, and (2) interdependent privacy and game theory. We survey related existing works in these areas and position our work.

6.1. Information Sharing on OSNs: Privacy and Benefits

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 [25], [26]), the shared information itself, and the benefits and privacy risks [27] associated with sharing. In some cases, the happiness of a user’s friends also come into the picture; this is usually captured through a so-called altruistic factor, as introduced in [23] and experimentally measured in [9]. In practice, deciding whether to share information often comes down to finding a sweet spot between privacy and benefits [28]. The decision process can be automated by (1) maximizing privacy under benefits (service quality) constraints [29] (or conversely), (2) taking a game-theoretic approach for modeling the interplay between the users and the adversaries [29], or (3) by mimicking the users’ sharing decisions using machine-learning techniques, after a training phase [30]. This latter approach has the advantage of relieving the controversial assumption on the users’ rationality and understanding of the problem [14].
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-theoretical approach for modeling the interplay between users and, ultimately, their decisions.

6.2. Interdependent Privacy and 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], [8], sensitive attributes inferred from those of a users’ friends on OSNs [31]–[33], demographic information inferred from a user’s interests [34], genomic data inferred from that of a user’s relatives [35], location leaked from geo-tagged pictures that friends upload online [36], relationships inferred from pictures [37], and co-locations detected from the users’ IP address at hotspots or reported on OSNs [6].

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], [8], privacy decision-making [38], such as sharing genomic data [35]. The study of interdependent privacy risks from an economic perspective follows the long line of research on interdependent security games surveyed in [39].

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.

7. Conclusion

It is a well-known fact that other people’s behaviors affect our own privacy, in particular in the case of (co-)location data. Yet, formalizing these interdependences and their implications is non-trivial, especially because human decisions play a dominant role in such settings. To address this issue, we focused on the (co-)location sharing features provided by major OSNs. We proposed a simple, coarse-grained, game-theoretic model and provided a first framework to study the interplay between two friends on OSNs. A major challenge in such approaches is to assign appropriate values to the various 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 and preferences in terms of social benefits and location privacy. We made use of this framework in different settings and studied the resulting equilibria, as well as their properties. 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 rational users to over-share their locations, even at later times. As part of 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.

References


Appendix

In this appendix, we provide the transcript of our survey questionnaire (Figure 12).

![Survey Questionnaire Transcript](Figure 12)
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. Image 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 item.

<table>
<thead>
<tr>
<th>Your choices</th>
<th>Your ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 of your recent check-in posts are kept and 0 of your recent co-location posts are kept</td>
<td></td>
</tr>
<tr>
<td>0 of your recent check-in posts are kept and 1 of your recent co-location posts are kept</td>
<td></td>
</tr>
<tr>
<td>1 of your recent check-in posts is kept and 1 of your recent co-location posts are kept</td>
<td></td>
</tr>
<tr>
<td>1 of your recent check-in posts is kept and 2 of your recent co-location posts are kept</td>
<td></td>
</tr>
<tr>
<td>2 of your recent check-in posts are kept and 1 of your recent co-location posts are kept</td>
<td></td>
</tr>
<tr>
<td>2 of your recent check-in posts are kept and 2 of your recent co-location posts are kept</td>
<td></td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Your choices*</th>
<th>Your ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>19 posts for an average privacy of 200m</td>
<td></td>
</tr>
<tr>
<td>12 posts for an average privacy of 400m</td>
<td></td>
</tr>
<tr>
<td>5 posts for an average privacy of 830m</td>
<td></td>
</tr>
<tr>
<td>24 posts for an average privacy of 0m</td>
<td></td>
</tr>
<tr>
<td>0 posts for an average privacy of 1100m (1,1km)</td>
<td></td>
</tr>
<tr>
<td>10 posts for an average privacy of 610m</td>
<td></td>
</tr>
</tbody>
</table>

* These numbers were extracted from the experimental results presented in [6].

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?

Figure 12. Transcript of our survey questionnaire (2/2).