Interdependent and Multi-Subject Privacy: Threats, Analysis and Protection

Thèse N° 9373

Présentée le 20 mars 2019

à la Faculté informatique et communications
Laboratoire pour les communications informatiques et leurs applications 1
Programme doctoral en informatique et communications

pour l'obtention du grade de Docteur ès Sciences

par

Alexandra-Mihaela OLTEANU

Acceptée sur proposition du jury

Prof. R. Urbanke, président du jury
Prof. J.-P. Hubaux, Prof. K. C. Huguenin, directeurs de thèse
Dr A. Lehmann, rapporteuse
Dr I. Ion, rapporteuse
Prof. C. Troncoso, rapporteuse

2019

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Education is the most powerful weapon which you can use to change the world.
Nelson Mandela

To my family

Abstract

In Alan Westin's generally accepted definition of privacy [1], he describes it as an individual's right "to control, edit, manage, and delete information about them[selves] and decide when, how, and to what extent information is communicated to others." Therefore, privacy is an individual and independent human right. The great Mahatma Gandhi once said that "interdependence is and ought to be as much the ideal of man as self-sufficiency. Man is a social being." To ensure this independent right to inherently social beings, it will be difficult, if not impossible. This is especially true as today's world is highly interconnected, technology evolves rapidly, data sharing is increasingly abundant, and regulations do not provide sufficient guidance in the realm of interdependency. In this thesis, we explore the topic of interdependent privacy from an adversarial point of view by exposing threats, as well as from an end-user point of view, by exploring awareness, preferences and privacy protection needs.

First, we quantify the effect of co-locations on location privacy, considering an adversary such as a social-network operator that has access to this information: Not only can a user be localized due to her reported locations and mobility patterns, but also due to those of her friends (and the friends of her friends and so on). We formalize this problem and propose effective inference algorithms that substantially reduce the complexity of localization attacks that make use of co-locations. Our results show that an adversary can effectively incorporate co-locations in attacks to substantially reduce users' location privacy; this exposes a real and severe threat.

Second, we investigate the interplay between the privacy risks and the social benefits of users when sharing (co-)locations on OSNs. We propose a game-theoretic framework for analyzing users' strategic behaviors. We conduct a survey of Facebook users and quantify their benefits of sharing vs. viewing information and their preference for privacy vs. benefits. Our survey exposes deficits in users' awareness of privacy risks in OSNs. Our results further show how users' individual preferences influence, sometimes in a negative way, each other's decisions.

Third, we consider various types of interdependent and multi-subject data (photo, co-location, genome, etc.) that often have privacy implications for data subjects other than the uploader, yet can be shared without their consent or awareness. We propose a system for sharing such data in a consensual and privacy-preserving manner. We implement it in the case of photos, by relying on image-processing and cryptographic techniques, as well as on a two-tier architecture. We conduct a survey of Facebook users; it indicates that there is interest in such a system, and that users have increasing privacy concerns due to prejudice or discrimination that they have been or could still easily be exposed to.

In conclusion, this thesis provides new insights on users' privacy in the context of interdependence and constitutes a step towards the design of novel privacy-protection mechanisms. It should be seen as a warning message for service providers and regulatory institutions: Unless the interdependent aspects of privacy are considered, this fundamental human right can never be guaranteed.

Keywords: privacy protection, interdependent privacy, multi-subject data, location privacy, co-location, genomic privacy, online social networks, Bayesian inference, game theory, consensual sharing

Résumé

Dans son ouvrage *Privacy and Freedom* [1], Alan Westin décrit la protection de la sphère privée comme un droit humain, individuel et indépendant: celui de pouvoir contrôler, modifier et effacer les informations à propos de soi et d'être capable de décider quand, comment et dans quelle mesure celles-ci peuvent être transmises à des tiers. D'autre part, les paroles de Mahatma Gandhi rappellent le caractère profondément social de l'être humain et portent l'interdépendance au niveau d'idéal pour l'Homme, au même titre que l'autosuffisance. Il semble difficile, voire impossible, de garantir un droit individuel et indépendant à un être fondamentalement social, en particulier dans le monde hautement interconnecté d'aujourd'hui. Cette thèse explore l'impact de la nature interdépendante des données personnelles sur la protection de ces dernières.

Premièrement, nous quantifions l'impact des données de co-localisation sur la protection de la sphère privée. Dans le contexte des médias sociaux par exemple, un utilisateur peut être localisé non-seulement à partir des données qu'il a directement partagées, mais également à partir de celles partagées par ses amis, amis d'amis et ainsi de suite. Nous formalisons ce problème et montrons comment les données de co-localisation peuvent être utilisées pour (mieux) localiser un individu. Pour ce faire, nous proposons un algorithme d'inférence de localisation et mesurons son impact, exposant ainsi une menace sérieuse et réelle.

Deuxièmement, nous investiguons les interactions entre les risques (en termes de sphère privée) et les bénéfices qui découlent du partage de données de (co-)localisation sur les médias sociaux. Nous proposons un modèle fondé sur la théorie des jeux pour analyser le comportement des utilisateurs. Nous en estimons les différents paramètres à l'aide d'une enquête menée auprès d'utilisateurs de Facebook. Nos résultats montrent comment les préférences d'un individu peuvent influencer les décisions d'autres utilisateurs, les poussant parfois à divulguer plus d'information.

Troisièmement, nous nous concentrons sur le partage de différents types de données interdépendantes ou multi-sujets (photos, co-localisations, génomes). Bien souvent, la décision de partager ces données est prise de manière unilatérale par un individu, sans consentement unanime parmi les sujets concernés, voire à l'insu de ces derniers. Nous proposons un système permettant de partager ces données de manière consensuelle, tout en protégeant la sphère privée des participants. Afin d'évaluer sa faisabilité technique, nous l'implémentons dans le cadre du partage de photos, en utilisant des techniques cryptographiques et de traitement d'image. Enfin, via une enquête ciblée, nous montrons un intérêt pour notre système de la part des utilisateurs de médias sociaux.

En conclusion, cette thèse permet de mieux comprendre les enjeux liés à la protection de la sphère privée dans un contexte d'interdépendance des données et constitue une avancée dans la conception de mécanismes de protection adaptés. Ce travail constitue également un message aux fournisseurs de services de partage en ligne et aux institutions de régulation: tant que les aspects interdépendants de la protection de la sphère privée ne sont pas pris en compte, ce droit humain ne pourra pas être pleinement garanti.

Mots-Clés: protection de la sphère privée, interdépendance, données multi-sujets, données de (co-)localisation, médias sociaux, inférence Bayésienne, théorie des jeux, partage consensuel, données génomiques

Acknowledgments

First and foremost, I am extremely grateful to my advisor, Prof. Jean-Pierre Hubaux, and to my co-advisor, Prof. Kévin Huguenin. Jean-Pierre, thank you for the opportunity to work on interesting research topics, for allowing me the freedom that I needed, for your understanding, your trust, and for always caring about my career. I learned a lot from you. Kévin, thank you for your immense academic and moral support, the ongoing inspiration, the humour, and the priceless memories brainstorming on whiteboards. You both have my immense gratitude and admiration.

I would like to express my gratitude to my thesis committee members Dr. Iulia Ion, Dr. Anja Lehmann, Prof. Carmela Troncoso, and Prof. Rüdiger Urbanke for their efforts spent reviewing my dissertation. It was a privilege to have them sit on my committee.

I am very thankful to my co-authors for our fruitful collaborations and their contributions to this thesis: Dr. Mathias Humbert, thank you for your dedication; it motivated me from the very beginning. Dr. Italo Dacosta, it was always a pleasure working with you. I also want to extend my gratitude to Dr. Elisa Celis, Dr. Konstantinos Chatzikokolakis, Prof. Virgil Gligor, Prof. Yves Pigneur, Prof. Reza Shokri, and Dr. Marco Stronati for their helpful input and ideas. I want to mention also the ICSIL system administrators and lab secretaries for their ongoing help and assistance. I am grateful to Holly for her help in improving my English, and to Patricia and Angela for their friendly help.

I am thankful to all my colleagues and friends at LCA1 for sharing this journey with me: Kévin, Mathias, Reza, Erman, Jean Louis, Zhicong, Anh, Italo, Juan, Joao, David, Christian, Ludovic, Mickaël, and Apostolos. Special thanks to Anh for growing wise together with me, to Italo for his friendly guidance and to Jean Louis for never being too busy for a chat.

My PhD years would not have been the same without the support of my friends and family. Stefan, thank you for inspiring me with the courage to jump and for helping me remember how a sunny day looks. Adish, your help was crucial and I am glad it led me to discover the great place that is EPFL. Denzil, I am grateful for your support in my decision to continue my journey here and for remaining a good friend. My dear, old friends from overseas, who are too many to enumerate, thank you for forgiving me for my lack of time during these past years and for your continued friendship. Ioana and Cristi, thank you for never forgetting Christmas and for welcoming me back. Monica, I appreciate that you are always curious about my research and I treasure the memories from each of our trips. AleJi, I promise I will visit soon. To my friends from EPFL, Sonia, Damian, Bogdan, Alex, and George, thank you for your friendship. Benoît, I am grateful for the peaceful moments at high altitudes and for the fact that you always found the time for a kind gesture, all the while tolerating and improving my limited French. Amit, those philosophical discussions that we shared will be a great memory. Nils, the ukulele lessons were much needed and appreciated. Erwan, your presence was like the priceless drops of water they give an athlete along the last miles of a marathon; thank you for your advice, the dinners delivered to my door step, and for all the adventures we shared. Anca, you will have to forgive me as I do not know how to even begin to thank you properly for being there unconditionally, persistently, wholeheartedly, and selfishly every step along the way; to you, Her Majesty, and the royal subjects, danke schön.

To my family, my most heartfelt appreciation for their encouragement and endless support. Thank you for staying closely by my side, for helping me anytime I needed it and for your unconditional love, unlimited sacrifice, kind understanding, and valuable life lessons. You have my eternal love, gratitude, and respect.

This thesis is dedicated to the family that I am proud to call my own, and to the one I was blessed to find along the way. *Mamă, tată, Anca, teza mea vă este dedicată.*

Contents

Chapter 1

Introduction

The defence of privacy will be the saviour of the future, essentially.

Svea Eckert

History provides abundant examples of technology having unanticipated, underestimated and lasting consequences on humanity. Privacy is no exception to this, despite having been one of the main concerns in our society for more than a century. With each new wave of the industrial revolution, we have found ourselves in an increasingly challenging situation in terms of ensuring the right to individual privacy, long recognized to be a fundamental human right[1]. The situation is becoming dire. With the advent of GPS and camera-equipped smartphones and the growing popularity of online social networks (OSNs), we interact with an increasing number of devices, services, and individuals, and in an unprecedented number of ways: We track and share our fitness activities; we immortalize our every move and social interaction through photos and posts on OSNs; we search for directions everywhere we go; we hunt for jobs; we chat, shop, date, read books, watch movies, and listen to music; and we can even determine our ancestry – all in the digital environment. The cost? Our data and, consequently, our privacy. Be it through the Web, mobile phones, or social media, in our daily activities, we produce increasing amounts of data, such as traces of our location over time, our social ties, our preferences, even our personal genome. But we trade these data in exchange for the many services that we enjoy.

Sadly, grasping the extent of where and how this data, which is often held by corporations, is and could be used is a massive challenge; and one that we, as a society, have so far failed. To name only a few examples, the research community has showed how, from simple location data, a user's social ties [2, 3], her interests [4], her personal locations (*e.g.,* home and workplace) [5], and her identity [6] can be inferred; even aggregate location data can be exploited for membership inference attacks (*i.e.,* determining whether

[1] "No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks." Article 12, Universal declaration of human rights, United Nations, 1984

or not a target user is part of the aggregate), as shown by Pyrgelis et al. [7]. Seemingly harmless information that a user shares, such as her music interests, can ultimately be exploited in order to leak private information (*e.g.,* her gender, relationship status and age) by making use of information from other people (virtual strangers), whose harmless interests are similar to hers [8]. Moreover, from our connections on OSNs or the data that they share, information about us can be retrieved, such as our age [9] and our (potentially hidden) OSN profile – representing our entire online identity – can be exposed [10, 11]. From the genomic data that our relatives share on dedicated online platforms, our own genomic code can be inferred [12]. This code represents our entire biological identity, from which extremely sensitive information can be extracted, such as our predisposition to certain diseases.

Undoubtedly, much of the data that is made available very often involves (and has privacy implications for) data subjects other than the individual who shares it online; and these individuals often have no control over the sharing decision, or might not even be aware of the fact that the data was made available. The consequences of this sharing can often be very dramatic, even when there is no ill intent from the individual who shared the data. An example of this is the case of an individual, whose decision to share his genome online made possible the discovery of illegitimate siblings and ultimately lead to his parents divorce [13]. With malicious intent, the damage can be catastrophic to a person's reputation. A sadly popular example of this is revenge pornography [14–16] (*i.e.,* the disclosure, by a former partner, of photos or videos portraying sexually explicit activity after the end of the relationship). This phenomenon has even led its victims to commit suicide [17, 18].

Although we are naturally linked as human beings and, as a result, our data is naturally correlated and our actions inherently have consequences on others, our opinions on the topic of privacy differ [19]; this is another aspect of the problem. Furthermore, our awareness regarding privacy threats is also variable and often insufficient. This is partly due to the fact that, whenever a user opts to use a service, its benefits are advertized but the associated privacy risks are not; these risks are often impossible to grasp, due to the complexity of the interactions and to the lack of transparency on how the data is handled. Yet, it is well known that privacy and utility have been and, most likely, will always be conflictual. At present, there are little to no mechanisms in place to prohibit the disclosure, by one individual, of information that could affect another individual. Hence, there is no incentive for service providers to change their behavior: Current regulations do not enforce it (most likely due to the complexity of such a task), and users do not enforce it (partially due to their lack of awareness).

The limits of how far data can be exploited, by the service provider holding it, or by third parties, to the detriment of an individual are extremely fuzzy. Government agencies already use data shared on OSNs for surveillance; and companies have and will find new ways to use such available data to discriminate against people (for insurance purposes, for hiring purposes, etc.) [20]. It is, therefore, paramount that we do not stop asking ourselves about how far surveillance, discrimination, prejudice, blackmail and reputation damage could extend, if left untamed. It is critical, perhaps now more than ever, for the future of our society that we continually fight to ensure the fundamental right to privacy, equally, for each and every individual.

In this thesis, we strive to push the balance of power towards the users and the legislative authorities, by exposing new situations where individual privacy can be affected in unforeseen ways, typically due to data interdependencies, and by designing technical solutions to keep users informed and to mitigate the privacy risks introduced by others.

Contributions

In this thesis, we address privacy issues related to various types of interdependent and multi-subject personal data. Our main contributions are as follows:

1. We identify the user-localization problem when co-location information is used. This kind of information about users is increasingly available online: For instance, by tagging the names of the friends with whom they are, mobile users increasingly frequently report their co-locations with other users in the messages and the pictures they post on OSNs. The users' IP addresses also constitute a source of co-location information. Combined with (possibly obfuscated) location information, such co-locations can be used to improve the inference of the users' locations, yet at the cost of high complexity. We formalize this problem and derive an optimal inference algorithm that also exploits co-location information, by incorporating probabilistic knowledge of the users' mobility and their disclosed locations. We analyze its complexity and show that, in practice, this kind of inference algorithm is intractable due to the explosion of the state-space size. We further suggest several approximate inference algorithms, including a solution that relies on the belief propagation algorithm executed on a Bayesian network model. We contribute implementations of the proposed algorithms and extensively evaluate their performance on a mobility dataset. One of our main findings is that the belief-propagation method (approximate inference that makes use of the data from all the users) converges to the solution of the optimal inference in polynomial time; thus, we show that an adversary can *effectively* incorporate co-locations to better localize users. Using the mobility dataset, we further quantify the effect of co-location information on location privacy, under different user mobility settings (*e.g.,* frequency of co-location disclosures, whether or not the user obfuscates her location before disclosing it, etc.). Our experimental results show that, even in the case where the adversary considers co-locations of the targeted user with a single friend, the median location privacy of the user is substantially decreased. Furthermore, in the case where a user does not disclose any location information herself, her privacy can decrease by up to 21% simply due to the information reported by her friends. This finding is paramount, as it brings to light the fact that users do not have full control over their location privacy. Therefore, we reveal a new threat to location privacy and suggest that protection mechanisms must not ignore the social aspects of location privacy. For a first attempt at mitigating the privacy risks stemming from co-location information, we propose and evaluate some simple countermeasures, including reporting fake co-locations and coordinated location disclosure; the latter is entirely within users' control and can reduce the privacy loss by up to 50%. Finally, we propose other generalization-based methods that socially aware location privacy-protection mechanisms could employ, such as generalizing the identity of the co-located users or the time of a shared co-location.

2. We study one of the most popular features in location-based OSNs (such as Facebook and Foursquare); this feature permits users to 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. They also constitute a severe threat to the privacy of all users involved and can have long-lasting and unanticipated effects for other users. This is due to the interdependences that co-location information introduces: Many users become connected and the collection of the data that is shared can have privacy implications for all of them. We identify the need to evaluate such benefits and privacy consequences, for a first step to understanding users' reasons for sharing such information and to providing the missing appropriate features that inform the affected users and solicit their consent. We propose the first game-theoretic framework for formalizing these complex interdependences and for analyzing the strategic behaviors of users; we take into account different attitudes and preferences regarding the sharing of location and co-location data. In order to design parametric utility functions that are representative of the users' actual preferences and to avoid the pitfalls of purely theoretical results, we also 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 preferences for privacy vs. benefits. We extensively evaluate our framework through data-driven numerical simulations on a mobility dataset: We study the resulting equilibria and their properties, using values of the parameters derived from the empirical data. We analyze how users' individual preferences influence each others' decisions, and we determine several factors that significantly influence these decisions. Our simulations unravel situations in which one of the users can be forced into a situation that she does not desire, and we demonstrate that sharing co-location information can also encourage users to over-share their locations; this is a fact that service providers could exploit by promoting co-location sharing features with the purpose of actually gathering more location data from their users. The survey also reveals the high diversity of opinions in terms of social benefits and location privacy. Our findings demonstrate the need to deploy, in practice, appropriate mechanisms for assisting users in their sharing decisions. Our generic framework represents the basis of such an effort.

3. We propose the first system that addresses the critical problem of *consensual and privacy-preserving sharing* of multi-subject personal and interdependent data in the online environment: *Before* such data is published, our system determines the affected parties, in a private manner, and asks for their consent. The key difference from the related works is that our proposed system is privacy-preserving with respect to *both other individuals using the system **and** the involved service providers*. We identify the many different challenges inherent to the design of such a system, we propose the main building blocks, and we discuss the incentives for adoption of all the parties involved. We implement and evaluate our system, ConsenShare, in the case of photos, by relying on image processing and cryptographic techniques, as well as on a two-tier architecture: one entity for detecting the data subjects and contacting them and another entity for hosting the data and for collecting consent. We benchmark the performance (CPU and bandwidth) of ConsenShare by using a dataset of 20k photos that we collected from Flickr. We conduct a survey of Facebook users (N = 321). Our experimental results demonstrate the feasibility of our approach (in terms of the acceptability of the overheads), and the survey results

demonstrate a potential desire from the users. We conclude that it is technically possible, without much overhead, to ensure users' privacy by giving them control in the sharing of photos in which they appear, and to preserve the main features of existing online platforms. We further discuss how each building block can be adapted to other types of data (*e.g.,* co-locations, videos, genomic data, etc.). In doing so, our work lays the foundation for the design of systems that will enable the sharing of data in a respectful manner, for the privacy of all the users involved, and *without* the need to trust these systems with the sensitive data.

Thesis Outline

This thesis contains three parts. In Chapter 2, we show how an adversary can efficiently exploit co-location information to better localize users by identifying a novel threat to users' location privacy. In Chapter 3, we analyze the interplay between friends on location-based OSNs and study how their individual preferences and shared data affects their respective sharing decisions over time. In Chapter 4, we describe a system that supports the sharing of interdependent and multi-subject personal data in a consensual and privacy-preserving way.

Publications

Chapter 2 is a combination of the results from [21] and [22]. Chapter 3 contains the findings of [23]. Chapter 4 relies on the results of [24].

Chapter 2

Location Inference Attacks Using Co-location Data

I think the greatest freedom that I have gained is that I no longer have to worry about what happens tomorrow, because I'm happy with what I've done today.

Edward Snowden

2.1 Introduction

Increasingly popular GPS-equipped mobile devices with Internet connectivity enable users to enjoy on-the-go a wide range of online location-based services. For instance, mobile users can search for nearby points of interest and get directions, possibly in real time, to their destinations. Yet, these location-based services raise serious privacy concerns because a large amount of personal information can be inferred from a user's whereabouts. The research community has extensively studied the problem of location privacy; more specifically, location-privacy protection mechanisms (so-called LPPMs), which can anonymize and obfuscate the users' locations before sending them to online location-based services, have been proposed [16]. In addition, formal frameworks for quantify location privacy in the case where users disclose their (possibly obfuscated) locations have been proposed [19, 20]. In such frameworks, the mobility profiles of the users play an important role in the inference of the users' locations, namely in a localization attack.

In parallel, social networks, and in particular location-based social networks, have become immensely popular. Every day, millions of users post information, including their locations, about themselves, but also about their friends. An emerging trend is to report co-locations with other users on social networks, *e.g.,* by tagging friends on

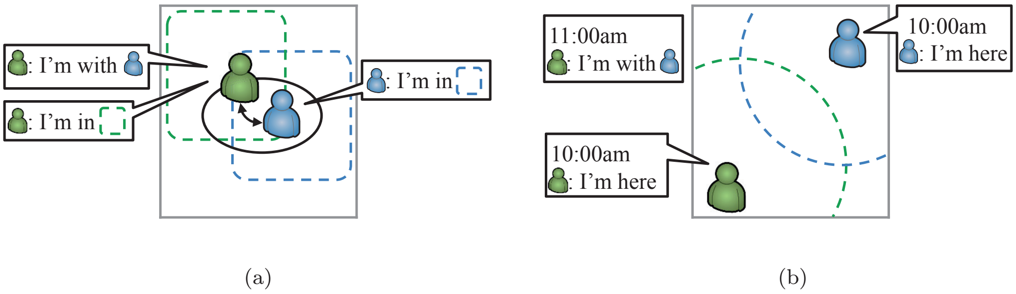


Figure 2.1: Examples showing how co-location information can be detrimental to privacy. (a) A user reports being in a given area, and a second user reports being in another (overlapping) area and that she is co-located with the first user. By combining these pieces of information, an adversary can deduce that both users are located in the intersection of the two areas, thus narrowing down the set of possible locations for both of them. (b) Two users (initially apart from each other, at 10am) declare their exact individual location. Later (at 11am), they meet and report their co-location without mentioning where they are. By combining these pieces of information, the adversary can infer that they are at a place that is reachable from both of the initially reported locations in the amount of time elapsed between the two reports.

pictures they upload or in the messages they post.[1] For instance, our preliminary survey involving 132 Foursquare users, recruited through Amazon Mechanical Turk, reveals that 55.3% of the participants report co-locations in their check-ins and that for the users who do so, on average, 2.84%±0.06 of their check-ins contain co-location information. In fact, co-location information can be obtained in many different ways, such as automatic face recognition on pictures (which contains the time and location at which the picture was taken in their EXIF data, *e.g.,* Facebook's Photo Magic [25]), Bluetooth-enabled device sniffing and reporting neighboring devices. Similarly, users who connect from the same IP address are likely to be attached to the same Internet access point, thus providing evidence of their co-location. Such data falls into the category of multiple-subject personal data [26].

Attacks exploiting both location and co-location information (as mentioned in [27]) can be quite powerful, as we show in this chapter. Figure 2.1 depicts and describes two instances in which co-location can improve the performance of a localization attack, thus degrading the location privacy of the users involved. It is clear that the proper exploitation of such information by an attacker can be complex because he has to consider jointly the (co-)location information collected about a potentially large number of users. This is due to the fact that, in the presence of co-location information, a user's location is correlated with that of her friends, which is in turn correlated to that of their own friends, and so on.

This family of attacks and their complexity is the focus of this chapter. More specifically, we make the following four contributions: (1) We identify and formalize the localization problem with co-location information, we propose an optimal inference algorithm and analyze its complexity. We show that, in practice, the optimal inference algorithm is intractable due to the explosion of the state space size. (2) We describe how an at-

[1] Note that the fact that a users tags one of her friends in a post does not necessarily mean that they are co-located; our formalism takes this fact into account.

tacker can drastically reduce the computational complexity of the attack by means of well-chosen approximations. We present a polynomial-time heuristic based on a limited set of considered users (i.e., optimal inference with the data of only two or three users) and an approximation that is based on the belief propagation (BP) algorithm executed on a general Bayesian network model of the problem (approximate inference with the data of all the users). (3) Using a mobility dataset, we evaluate and compare the performance of the different solutions in different scenarios, with different settings. The belief propagation-based solution gives results significantly better (in terms of the performance of the inference) than the heuristic. (4) We propose and evaluate some countermeasures (i.e., social-aware location-privacy protection mechanisms), including fake co-locations reporting and coordinated location disclosure. Our experimental results show that, even in the case where the adversary considers co-locations with only a single friend of the targeted user, the median location privacy of the user is decreased by up to 62% in a typical setting. Even in the case where a user does not disclose any location information, her privacy can decrease by up to 21%, due to the information reported by other users. A paramount finding of our work is that users partially lose control over their location privacy because co-locations and individual location information disclosed by other users substantially affect their own location privacy. Our experimental results also show that a simple countermeasure (i.e., coordinated location disclosure) can reduce the privacy loss by up to 50%. To the best of our knowledge, this is the first attempt to quantify the effects of co-location information that stems from social relationships, on location privacy; thus making a connection between OSNs and location privacy.

The remainder of the chapter is organized as follows. In Section 2.2, we define and formalize the system model. In Section 2.3, we present the optimal localization attack for N users and assess its complexity. In Section 2.4, we show how this complexity can be reduced by means of approximations. In Section 2.5, we propose and evaluate some countermeasures. In Section 2.6, we briefly analyze the co-location problem from a differential privacy perspective. In Section 2.7, we report on the experimental evaluation of the localization attack with co-locations. In Section 2.8, we survey the related work. In Section 2.9, we conclude the chapter and suggest directions for future work.

2.2 System Model and Formalization

We consider a set of mobile users who move within a given geographical area. While on the go, users make use of some online services to which they communicate potentially obfuscated locations (*i.e.,* where they are) and co-location information (*i.e.,* who they are with). Note that such information could be communicated unintentionally by the users (e.g., leaked from their IP addresses) without their knowing it. We consider that a curious service provider (referred to as the adversary) wants to infer the location of the users from this information hence to track them over time. In order to carry out the inference attack based on which the location privacy of the users is evaluated, the adversary would model the users as described below. Our model is built upon [28] and uses similar notations. Figure 2.2 on page 10 gives an overview of the considered scenario and Table 2.1 on page 11 summarizes the main notations used in our formalization throughout the chapter.

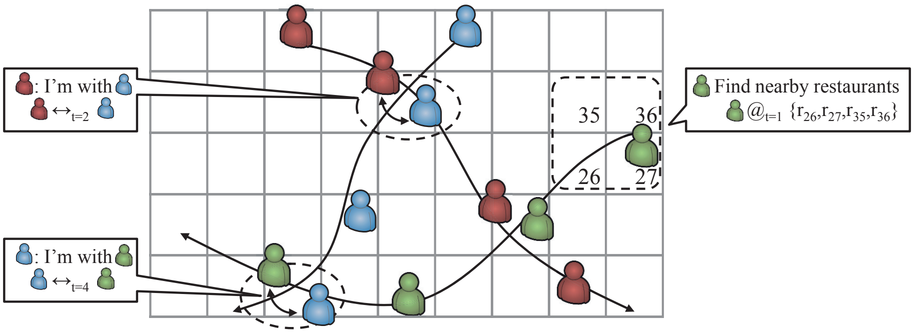


Figure 2.2: Scenario of (co-)location exposure. Three users move in a given geographical area. They communicate their potentially obfuscated locations and accurate co-location information to a service provider (*i.e.,* the adversary) who wants to infer their locations.

2.2.1 Users

We consider a set $\mathcal{U} = \{u_1, \ldots, u_N\}$ of N mobile users who move within a given geographical area that is partitioned into M regions (locations) $\mathcal{R} = \{R_1, \ldots, R_M\}$. Time is discrete, and we consider the state of the system (including the locations of the users) at the successive time instants $\{1, \ldots, T\}$. The region in which a user $u \in \mathcal{U}$ is at time instant $t \in \{1, \ldots, T\}$ is called the *actual location* of the user and is denoted by $a_u(t)$. The mobility of the users is modeled by a first order time-homogeneous Markov chain. We denote by $p_u(\rho, r)$ the probability that user u moves from region ρ to region r during one time instant, and by $\pi_u(r)$ the probability that user u is in region r at time t (*i.e.,* the stationary distribution of p_u). We use the term "co-location" for when two users are at the same location at some point in time. The fact that users u and v are co-located at time t means that $a_u(t) = a_v(t)$; we denote by $u \leftrightarrow_t v$ the fact that a co-location between users u and v at time t is reported (by either of them), and we consider an associated binary variable $c_{u,v}(t)$; specifically, $c_{u,v}(t) = c_{v,u}(t) = 1$ if $u \leftrightarrow_t v$ and $c_{u,v}(t) = c_{v,u}(t) = 0$ otherwise. Note, however, that a co-location being reported does not necessarily mean that the users are really co-located (for instance, on OSNs, users sometimes tag a friend in a picture just to get their attention, although they are not actually together). We consider the process of users reporting co-location information to be probabilistic. Specifically, for any pair of users u and v, the probability of reporting a co-location, knowing both their actual locations is denoted by

$$g_{u,v}(r, r') \triangleq \Pr\left(u \leftrightarrow_t v \,|\, a_u(t) = r, a_v(t) = r'\right) \tag{2.1}$$

We assume that co-locations reported by a user at different time instants are being reported independently of each other and of those reported by other users. We also assume that the reporting process for any user does not depend on time. Intuitively, this co-location reporting function can incorporate social ties (users report co-locations on social networks only with their friends), selective reporting of co-location (not every time that Alice is with Bob in the same location does she report that on their favorite social network), as well as erroneous co-locations (Alice might tag Bob in a picture, even though he is not really in that picture). Examples of erroneous co-locations also include the case where Alice and Bob have the same IP address but they are not together (e.g., they make use of the same proxy). More generally, this case includes the false positives

$\mathcal{U}$	Set of mobile users
$\mathcal{R}$	Set of regions that partition the whole area
N	Number of users ($N = \|\mathcal{U}\|$)
M	Number of regions ($M = \|\mathcal{R}\|$)
T	Number of time instants
$p_u(\cdot, \cdot)$	Mobility profile of user u
$\pi_u(\cdot)$	The stationary distribution of p_u
$f_u(\cdot)$	Obfuscation function employed by user u
$g_{u,v}(\cdot, \cdot)$	Co-location reporting function for users u and v
$\mathcal{K}$	Adversary's background knowledge
$a_u(t)$	Actual location of user u at time t
$\mathbf{a}(t)$	Actual locations of all the users at time t
$u \,@_t\, r$	User u reports being in r at time t
$o_u(t)$	Obfuscated location of user u at time t
$\mathbf{o}(t)$	Obfuscated locations of all the users at time t
$u \leftrightarrow_t v$	A co-location was reported between u and v at time t
$c_{u,v}(t)$	Binary variable incorporating whether $u \leftrightarrow_t v$
C_t	Set of all reported co-locations at time t
C	Set of all reported co-locations

Table 2.1: Table of notations.

of the underlying co-location detection technique used by the adversary, and possibly fake co-locations reported by users to protect their privacy. The probabilistic co-location reporting function and its parameters are assumed to be known to the adversary; in practice, it could be learned from models of the users' behavior, or from ground-truth data or, when applicable, from theoretical models of the underlying technical co-location detection method. Concrete examples of co-location reporting functions are given in Section 2.7. We assume all user-reported co-locations are observed by an adversary.

2.2.2 Location-Privacy Protection Mechanisms

In order to protect their privacy, we assume that users rely on location-privacy protection mechanisms (LPPM) for obfuscating their individual location information before this is communicated to an online service provider. We denote by $u \,@_t\, r'$ the fact that user u reports being at location r' at time t to the online service. The online service observes only the obfuscated location of the users; we denote this by $o_u(t)$ for user u at time t. We denote by $\mathcal{R}'$ the set of obfuscated locations; typically $\mathcal{R}'$ is the power set of $\mathcal{R}$, as LPPMs can return a set of locations instead of only one location. Typical LPPMs replace the actual location of a user with another location (*i.e.,* adding noise to the actual location) or merge several regions (*i.e.,* reducing the granularity of the reported location). We model an LPPM with a function that maps a user's actual location to a random variable that takes values in $\mathcal{R}'$, *i.e.,* , the user's obfuscated location. This means that the locations of a user at different time instants are obfuscated independently of each other and of those of other users. This also means that the way a user's locations are obfuscated does not depend on time. Formally, an LPPM is defined by the function $f_u(r, r')$ that denotes the probability that the LPPM used by u obfuscates location r to r', *i.e.,* $\Pr(o_u(t) = r' \mid a_u(t) = r)$. Excluding the co-location information, our model corresponds to a hidden Markov model (HMM) [29]. We assume that co-location information is not obfuscated and users do not rely on pseudonyms.[2] We denote by $\mathbf{o}(t)$ the vector of the observed locations of all the users at time t. More generally, we use bold notations to denote a vector of values of all users. We define $C_t = \{c_{u,v}(t)\}_{u,v\in\mathcal{U}}$ and $C = \bigcup_{t=1..T} C_t$.

2.2.3 Adversary

The adversary, typically an online service provider (or an external observer who has access to this information, *e.g.,* another user of the social network), has access to the observed locations and co-locations of one or several users and seeks to locate users, at a given time instant, specifically, carry out a *localization attack*.[3] Because of the co-location information, the locations of the users are not independent (they are correlated), thus when attacking the location of a given user, the adversary takes into account information potentially about all the users. The attack is performed *a posteriori*, meaning that the adversary has access to the observed traces over the complete period, namely $\{\mathbf{o}(t)\}_{t=1..T}$ and C, at the time of the attack. In addition to the observations during the time period of interest (*i.e.,* $\{1,\ldots,T\}$), the adversary has access to some of the users' past location traces, from which he builds individual mobility profiles for these users, under the form of transition probabilities $\{p_u\}_{u\in\mathcal{U}}$. See [28] for more details about the knowledge construction, in particular, on how the mobility profiles can be built from obfuscated traces with missing locations. The mobility and co-location reporting profiles constitute, together with the knowledge of the LPPMs used by the users (including their parameters), the adversary's *background knowledge* $\mathcal{K} = \{p_u(\cdot,\cdot)\}_{u\in\mathcal{U}}, \{f_u(\cdot)\}_{u\in\mathcal{U}}, \{g_{u,v}(\cdot,\cdot)\}_{u,v\in\mathcal{U}}$.

The output of a localization attack that targets user u at time t, is a *posterior probability distribution* over the set $\mathcal{R}$.

$$h_t^u(r) \triangleq \Pr\left(a_u(t) = r \mid \{\mathbf{o}(t)\}_{t=1..T}, C, \mathcal{K}\right) \quad . \tag{2.2}$$

2.2.4 Location-Privacy Metric

The location privacy $\mathrm{LP}_u(t)$ of user u at time t, with respect to a given adversary, is captured by the expected error of the adversary when performing a localization attack [28]. Given the output $h_t^u(\cdot)$ of the localization attack, the location privacy writes

$$\mathrm{LP}_u(t) \triangleq \sum_{r\in\mathcal{R}} h_t^u(r) \cdot d(r, a_u(t)) \quad , \tag{2.3}$$

where $d(\cdot\,,\cdot)$ denotes a distance function on the set $\mathcal{R}$ of regions, typically the Haversine distance between the centers of the two regions.

2.3 Optimal Localization Attack

Without co-location information (as in [28]) and under the assumptions described in the previous section, the localization problem translates to solving an HMM inference

[2] Note that even if pseudonyms are used, the identity of the users can be inferred by using their social network [30] or their locations [28]. We make this assumption because our main target scenario is users *posting* information attached to their real identities on social networks.

[3] This attack is somewhat similar to correlation attacks on continuous location-based queries with cloaking as presented in [31].

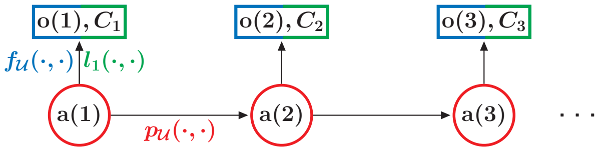


Figure 2.3: Sample HMM for $T = 3$ time instants. *States* are represented by red circles, and *observations* by blue-green rectangles. State transition probabilities are specified by the joint user mobility profiles $p_{\mathcal{U}}$ and output probabilities are specified by a combination of $f_{\mathcal{U}}$ (for individual observations) and l_t (for co-location observations).

problem, for which the *forward-backward* algorithm is a known solution. Essentially, the forward-backward algorithm defines forward and backward variables that take into account the observations before and after time t, respectively. The forward variable is the joint probability of the location of user at time t and all of the observations up to, and including, time t. The backward variable is the conditional probability of all observations after time t, given the actual location of user at that time instant. Then, the posterior probability distribution of the possible locations for the targeted user is obtained by combining (*i.e.,* multiplying and normalizing) the forward and backward variables. With co-location information, the locations of the users are not mutually independent: as soon as two users are co-located at some point in time t, their locations, before and after time t, become dependent. Actually, the fact that two users meet a same third user (even if they meet her at different time instants) suffices to create some dependencies between their locations; this means that, to perform the localization attack on a user, the adversary must take into account the locations (*i.e.,* the obfuscated location information and the co-location information) of *all* the users who are connected to u by a *chain* of co-location (*i.e.,* the connected component of u in the co-location graph). Formally speaking, this means that the adversary cannot rely only on the *marginal* distributions of the users' location; instead he must consider the *joint* distributions. In other words, co-locations turn N disjoint inference problems (*i.e.,* HMM problems solved by the forward-backward algorithm) into a joint inference problem.

To solve the localization problem, we consider all the users jointly and show that it translates to an HMM problem, as depicted in Figure 2.3. Note that more advanced learning techniques, such as neural networks, could also be used. We solve this problem by using the forward-backward algorithm [32, 33]. For a set $\mathcal{U}$ of users and time t, we define the following forward and backward variables:

$$\begin{aligned} \alpha_t^{\mathcal{U}}(\mathbf{r}) &\triangleq \Pr\left(\mathbf{o}(1)\ldots\mathbf{o}(t), C_1\ldots C_t, \mathbf{a}(t)=\mathbf{r}\,|\,\mathcal{K}\right) \\ \beta_t^{\mathcal{U}}(\mathbf{r}) &\triangleq \Pr\left(\mathbf{o}(t+1)\ldots\mathbf{o}(T), C_{t+1}\ldots C_T|\mathbf{a}(t)=\mathbf{r},\mathcal{K}\right) \end{aligned} \tag{2.4}$$

where $\mathbf{r}$ denotes a vector of size N, *i.e.,* $\mathbf{r} \in \mathcal{R}^N$, and represents the actual locations of all users at a single time instant. These variables can be defined recursively (over t) and, unlike in the case where no co-location observations are available, their expressions involve the co-location information. More specifically, we prove that for all $\mathbf{r} \in \mathcal{R}^N$, we have[4]

[4] For the sake of simplicity and clarity, we define the variables at $t = 0$ even though no observations are made at this time instant.

$$\alpha_t^{\mathcal{U}}(\mathbf{r}) = \begin{cases} \pi_{\mathcal{U}}(\mathbf{r}) & \text{if } t = 0 \\ l_t(\mathbf{r}, C) \cdot f_{\mathcal{U}}(\mathbf{r}, \mathbf{o}(t)) \cdot & \\ \qquad \displaystyle\sum_{\boldsymbol{\rho} \in \mathcal{R}^N} \alpha_{t-1}^{\mathcal{U}}(\boldsymbol{\rho}) \cdot p_{\mathcal{U}}(\boldsymbol{\rho}, \mathbf{r}) & \text{if } t > 0 \end{cases} \tag{2.5}$$

and

$$\beta_t^{\mathcal{U}}(\mathbf{r}) = \begin{cases} \displaystyle\sum_{\boldsymbol{\rho} \in \mathcal{R}^N} l_{t+1}(\boldsymbol{\rho}, C) \cdot \beta_{t+1}^{\mathcal{U}}(\boldsymbol{\rho}) \cdot & \\ \qquad p_{\mathcal{U}}(\mathbf{r}, \boldsymbol{\rho}) \cdot f_{\mathcal{U}}(\boldsymbol{\rho}, \mathbf{o}(t+1)) & \text{if } t < T \\ 1 & \text{if } t = T \end{cases} \tag{2.6}$$

where $\mathbf{r} = (r_1, \dots, r_N) \in \mathcal{R}^N$, $\boldsymbol{\rho} = (\rho_1, \dots, \rho_N) \in \mathcal{R}^N$, $\mathbf{r}' = (r'_1, \dots, r'_N) \in \mathcal{R}'^N$, and

$$\pi_{\mathcal{U}}(\mathbf{r}) = \prod_{i=1}^{N} \pi_{u_i}(r_i),$$
$$f_{\mathcal{U}}(\mathbf{r}, \mathbf{r}') = \prod_{i=1}^{N} f_{u_i}(r_i, r'_i),$$
$$p_{\mathcal{U}}(\boldsymbol{\rho}, \mathbf{r}) = \prod_{i=1}^{N} p_{u_i}(\rho_i, r_i).$$

$l_t(\cdot,\cdot)$ denotes the joint probability that the users report the set of co-locations observed at time t, when the configuration of their actual locations at t is given. That is, formally,

$$\begin{aligned} l_t(\mathbf{r}, C) &\triangleq \Pr(C_t | \mathbf{a}(t) = \mathbf{r}) \\ &= \prod_{u_i \neq u_j \in \mathcal{U}} \begin{cases} g_{u_i,u_j}(r_i, r_j) & \text{if } (u_i \leftrightarrow_t u_j) \in C_t \\ 1 - g_{u_i,u_j}(r_i, r_j) & \text{otherwise} \end{cases} \end{aligned} \tag{2.7}$$

More specifically, this is a likelihood function that captures the probability that the co-locations in C_t are reported, and that takes into account the individual co-location reporting function for every pair of users. As we assumed that co-locations are reported independently of one another, this likelihood can be expressed as a product of individual co-location reporting functions for all pairs of users.

The intuition behind Equation (2.5) is that the forward variable at time t can be expressed recursively, with respect to time, by combining, for all possible locations of the users at time $t-1$: (1) the joint probability that the users were at location $\boldsymbol{\rho}$ at time $t-1$ and reported the obfuscated locations and co-locations observed by the adversary up to time $t-1$ (this is captured by $\alpha_{t-1}^{\mathcal{U}}$), (2) the joint probability that the users move from the locations $\boldsymbol{\rho}$ to the locations $\mathbf{r}$ (this is captured by $p_{\mathcal{U}}$), (3) the joint probability that the users obfuscate their locations $\mathbf{r}$ to those observed by the adversary $\mathbf{o}(t)$ (this is captured by $f_{\mathcal{U}}$), and (4) the joint probability that the users report co-locations C_t observed by the adversary, assuming their locations $\mathbf{r}$ (this is captured by $l_t(\mathbf{r}, C)$). Because users obfuscate their locations independently from each other, the joint obfuscation probability is the product of the individual obfuscation probabilities (hence the expression of $f_{\mathcal{U}}$). The same applies to $p_{\mathcal{U}}$ and $l_t(\mathbf{r}, C)$. A similar line of reasoning applies to Equation (2.6).

The function $l_t(\cdot,\cdot)$ captures the likelihood of observing a set of co-location information (or not) given the actual users' locations. Schematically speaking (with a deterministic vision where only real co-locations are reported, for the sake of clarity), the set of possible locations for a user u_i (at time t), co-located with a user u_j, consists of the locations that can be obfuscated into the location reported by u_i at time t, that can be reached (according to u_i's mobility profile) from a possible location of u_i at time $t-1$, that can be obfuscated into the location reported by u_j at time t, **and** that can be reached (according to u_j's mobility profile) from a possible location of u_j at time $t-1$.

2.3.1 Development of Equation (2.5)

As the adversary does not have knowledge about conditional mobility profiles for the users, their mobility profiles are independent of each other – formally, $\Pr(a_u(t)=r|a_{u'}(t)=r') = \Pr(a_u(t)=r)$, for any users u and u'. Using *Bayes' rule* it follows that, for any $\mathbf{r}\in\mathcal{R}^N$

$$\Pr(\mathbf{a}(t)=\mathbf{r}) = \prod_{i=1}^{N}\Pr(a_{u_i}(t)=r_i) \tag{2.8}$$

We start the development of Equation (2.5) by proving its base case: $t=0$.

$$\begin{aligned}
\alpha_0^{\mathcal{U}}(\mathbf{r}) &= \Pr(\mathbf{a}(0)=\mathbf{r}\,|\,\mathcal{K}) && (2.9)\\
&= \Pr(a_{u_1}(0)=r_1\,|\,\mathcal{K})\times\cdots\times \\
&\qquad\qquad \Pr(a_{u_N}(0)=r_N\,|\,\mathcal{K}) && (2.10)\\
&= \pi_{u_1}(r_1)\ldots\pi_{u_N}(r_N) && (2.11)\\
&= \pi_{\mathcal{U}}(\mathbf{r}) && (2.12)
\end{aligned}$$

In step (2.9)→(2.10) of the derivation, we use the independence assumption (2.8); in step (2.10)→(2.11), we use the fact that the probability of a user u being in some region r at time $t=0$, given her mobility profile, is captured by the steady state vector, *i.e.,* $\pi_u(r)$, as there are no observations at, or before, $t=0$.

We now complete the development for any $t>0$.

$$\begin{aligned}
\alpha_t^{\mathcal{U}}(\mathbf{r}) &= \Pr(\mathbf{o}(1)\ldots\mathbf{o}(t),C_1\ldots C_t,\mathbf{a}(t)=\mathbf{r}\,|\,\mathcal{K}) && (2.13)\\
&= \Pr(C_t\,|\,\mathbf{o}(1)\ldots\mathbf{o}(t),C_1\ldots C_{t-1},\mathbf{a}(t)=\mathbf{r},\mathcal{K})\cdot \\
&\quad \Pr(\mathbf{o}(1)\ldots\mathbf{o}(t),C_1,\ldots,C_{t-1},\mathbf{a}(t)=\mathbf{r}\,|\,\mathcal{K}) && (2.14)\\
&= \Pr(C_t\,|\,\mathbf{a}(t)=\mathbf{r},\mathcal{K})\cdot \\
&\quad \Pr(\mathbf{o}(1)\ldots\mathbf{o}(t),C_1\ldots C_{t-1},\mathbf{a}(t)=\mathbf{r}\,|\,\mathcal{K}) && (2.15)\\
&= l_t(\mathbf{r},C)\cdot \\
&\quad \Pr(\mathbf{o}(1)\ldots\mathbf{o}(t),C_1\ldots C_{t-1},\mathbf{a}(t)=\mathbf{r}\,|\,\mathcal{K}) && (2.16)\\
&= l_t(\mathbf{r},C)\cdot\Pr(\mathbf{o}(t)\,|\,\mathbf{a}(t)=\mathbf{r},\mathcal{K})\cdot \\
&\quad \Pr(\mathbf{o}(1)\ldots\mathbf{o}(t-1),C_1\ldots C_{t-1},\mathbf{a}(t)=\mathbf{r}\,|\,\mathcal{K}) \\
& && (2.17)\\
&= l_t(\mathbf{r},C)\cdot f_{\mathcal{U}}(\mathbf{r},\mathbf{o}(t))\cdot
\end{aligned}$$

$$
\begin{aligned}
&\Pr\left(\mathbf{o}(1)\ldots\mathbf{o}(t-1), C_1\ldots C_{t-1}, \mathbf{a}(t)=\mathbf{r}\,|\,\mathcal{K}\right) && (2.18)\\
&= l_t(\mathbf{r},C)\cdot f_{\mathcal{U}}\left(\mathbf{r},\mathbf{o}(t)\right)\cdot \\
&\quad\sum_{\boldsymbol{\rho}\in\mathcal{R}^N}\Pr\left(\mathbf{o}(1)\ldots\mathbf{o}(t-1), C_1\ldots C_{t-1},\right.\\
&\qquad\qquad \left.\mathbf{a}(t)=\mathbf{r}, \mathbf{a}(t-1)=\boldsymbol{\rho}\,|\,\mathcal{K}\right) && (2.19)\\
&= l_t(\mathbf{r},C)\cdot f_{\mathcal{U}}\left(\mathbf{r},\mathbf{o}(t)\right)\cdot \\
&\quad\sum_{\boldsymbol{\rho}\in\mathcal{R}^N}\Pr\left(\mathbf{o}(1)\ldots\mathbf{o}(t-1), C_1\ldots C_{t-1},\right.\\
&\qquad\qquad \left.\mathbf{a}(t-1)=\boldsymbol{\rho}\,|\,\mathcal{K}\right)\cdot\\
&\qquad \Pr\left(\mathbf{a}(t)=\mathbf{r}\,|\,\mathbf{a}(t-1)=\boldsymbol{\rho},\mathcal{K}\right) && (2.20)\\
&= l_t(\mathbf{r},C)\cdot f_{\mathcal{U}}\left(\mathbf{r},\mathbf{o}(t)\right)\cdot \\
&\qquad\sum_{\boldsymbol{\rho}\in\mathcal{R}^N}\alpha^{\mathcal{U}}_{t-1}(\boldsymbol{\rho})\cdot p_{\mathcal{U}}\left(\boldsymbol{\rho},\mathbf{r}\right) && (2.21)
\end{aligned}
$$

In step (2.13)→(2.14) of the derivation, we apply the *chain rule*. In step (2.14)→ (2.15), we use *conditional independence*: given $\mathbf{a}(t) = \mathbf{r}$, the probability that the locations $\mathbf{r}$ can represent the reported C_t depends neither on the observations, nor on $\mathcal{K}$. In step (2.15)→(2.16), we use Definition (2.7). In step (2.16)→(2.17), we apply the chain rule and use conditional independence: given $\mathbf{a}(t) = \mathbf{r}$, $\mathbf{o}(t)$ does not depend on the past observations. In step (2.17)→(2.18), we use the fact that the location obfuscation process is applied independently for each user. In step (2.18)→(2.19), we apply the *law of total probability*, conditioning over all the possible actual locations $\boldsymbol{\rho}$ users could have been at, at time $t-1$. In step (2.19)→(2.20), we use the chain rule and conditional independence: given $\mathbf{a}(t-1) = \boldsymbol{\rho}$, $\mathbf{a}(t)$ does not depend on the past observations. In step (2.20)→(2.21), we use Definition (2.4). ■

2.3.2 Development of Equation (2.6)

We start the development of Equation (2.6) with the case $t = T$. As there are no observations at or after $T + 1$, using Definition (2.4), we consider

$$\beta^{\mathcal{U}}_T(\mathbf{r}) = \Pr\left(\varnothing|\mathbf{a}(T)=\mathbf{r},\mathcal{K}\right) = 1 \tag{2.22}$$

We now complete the development for any $t < T$.

$$
\begin{aligned}
\beta^{\mathcal{U}}_t(\mathbf{r}) &= \Pr\left(\mathbf{o}(t+1)\ldots,\mathbf{o}(T), C_{t+1},\ldots,C_T\,|\,\mathbf{a}(t)=\mathbf{r},\mathcal{K}\right) && (2.23)\\
&= \sum_{\boldsymbol{\rho}\in\mathcal{R}^N}\Pr\left(\mathbf{o}(t+1)\ldots,\mathbf{o}(T), C_{t+1},\ldots,C_T, \mathbf{a}(t+1)=\boldsymbol{\rho}\,|\,\mathbf{a}(t)=\mathbf{r},\mathcal{K}\right) && (2.24)\\
&= \sum_{\boldsymbol{\rho}\in\mathcal{R}^N}\Pr\left(C_{t+1}\,|\,\mathbf{a}(t+1)=\boldsymbol{\rho},\mathcal{K}\right)\cdot\\
&\qquad \Pr\left(\mathbf{o}(t+1)\ldots,\mathbf{o}(T), C_{t+2},\ldots,C_T, \mathbf{a}(t+1)=\boldsymbol{\rho}\,|\,\mathbf{a}(t)=\mathbf{r},\mathcal{K}\right) && (2.25)\\
&= \sum_{\boldsymbol{\rho}\in\mathcal{R}^N} l_{t+1}(\boldsymbol{\rho},C)\cdot\\
&\qquad \Pr\left(\mathbf{o}(t+1)\ldots,\mathbf{o}(T), C_{t+2},\ldots,C_T, \mathbf{a}(t+1)=\boldsymbol{\rho}\,|\,\mathbf{a}(t)=\mathbf{r},\mathcal{K}\right) && (2.26)
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{\boldsymbol{\rho}\in\mathcal{R}^N} l_{t+1}(\boldsymbol{\rho}, C)\cdot \Pr\left(\mathbf{o}(t+1)\,|\,\mathbf{a}(t+1)=\boldsymbol{\rho},\mathcal{K}\right)\cdot \\
&\qquad \Pr\left(\mathbf{o}(t+2)\ldots,\mathbf{o}(T),C_{t+2},\ldots,C_T,\mathbf{a}(t+1)=\boldsymbol{\rho}\,|\,\mathbf{a}(t)=\mathbf{r},\mathcal{K}\right) \quad (2.27)\\
&= \sum_{\boldsymbol{\rho}\in\mathcal{R}^N} l_{t+1}(\boldsymbol{\rho}, C)\cdot f_U(\boldsymbol{\rho},\mathbf{o}(t+1))\cdot \\
&\qquad \Pr\left(\mathbf{o}(t+2)\ldots,\mathbf{o}(T),C_{t+2},\ldots,C_T,\mathbf{a}(t+1)=\boldsymbol{\rho}\,|\,\mathbf{a}(t)=\mathbf{r},\mathcal{K}\right) \quad (2.28)\\
&= \sum_{\boldsymbol{\rho}\in\mathcal{R}^N} l_{t+1}(\boldsymbol{\rho}, C)\cdot f_U(\boldsymbol{\rho},\mathbf{o}(t+1))\cdot \Pr\left(\mathbf{a}(t+1)=\boldsymbol{\rho}\,|\,\mathbf{a}(t)=\mathbf{r},\mathcal{K}\right)\cdot \\
&\qquad \Pr\left(\mathbf{o}(t+2)\ldots,\mathbf{o}(T),C_{t+2},\ldots,C_T,\,|\,\mathbf{a}(t+1)=\boldsymbol{\rho},\mathcal{K}\right) \quad (2.29)\\
&= \sum_{\boldsymbol{\rho}\in\mathcal{R}^N} l_{t+1}(\boldsymbol{\rho}, C)\cdot f_U(\boldsymbol{\rho},\mathbf{o}(t+1))\cdot p_{\mathcal{U}}\left(\mathbf{r},\boldsymbol{\rho}\right)\cdot \beta^{\mathcal{U}}_{t+1}(\boldsymbol{\rho}) \quad (2.30)
\end{aligned}
$$

In step (2.23→2.24) we apply the *law of total probability*, conditioning over all the possible actual locations users could have been at, at time $t+1$. In step (2.24→2.25) we apply the *chain rule* and use *conditional independence*: given $\mathbf{a}(t+1)=\boldsymbol{\rho}$, the probability that these actual locations are consistent with C_{t+1} does not depend on other past observations, actual locations or $\mathcal{K}$. In step (2.25→2.27) we apply the *chain rule* and use *conditional independence*: given the actual locations of users $\mathbf{a}(t+1)=\boldsymbol{\rho}$ at time $t+1$, the probability of observing $\mathbf{o}(t+1)$ does not depend on other observations or actual locations of other time instants. In step (2.27→2.28) we use the fact that the obfuscation process is independently applied for all users. In step (2.28→2.29) we apply the *chain rule* and use *conditional independence*: given $\mathbf{a}(t+1)=\boldsymbol{\rho}$, the probability of observations at times ar or after $t+1$ does not depend on $\mathbf{a}(t)=\mathbf{r}$. Finally, in step (2.29→2.30) we use Definition (2.4). ■

Finally, the posterior probability distribution of the users' locations can be computed based on the forward and backward variables, by using the following formula, for $u_i\in\mathcal{U}$ and at time t:

$$
\begin{aligned}
h^{u_i}_t(r) &\triangleq \Pr\left(a_{u_i}(t)=r\,|\,\{\mathbf{o}(t)\}_{t=1..T},C,\mathcal{K}\right)\\
&= \frac{\displaystyle\sum_{\mathbf{r}\in\mathcal{R}^N\,|\,r_i=r} \alpha^{\mathcal{U}}_t(\mathbf{r})\cdot\beta^{\mathcal{U}}_t(\mathbf{r})}{\displaystyle\sum_{\mathbf{r}\in\mathcal{R}^N} \alpha^{\mathcal{U}}_t(\mathbf{r})\cdot\beta^{\mathcal{U}}_t(\mathbf{r})} \quad (2.31)
\end{aligned}
$$

In short, the probability that the users are at given locations at time t is computed based on all the observations before and at time t (α_t) and the observations after time t (β_t). The denominator is simply a normalization factor.

2.3.3 Development of Equation (2.31)

Using Definition (2.4), the *conditional independence* given $\mathbf{a}(t)=\mathbf{r}$ and the *chain rule*, we first compute the following, $\forall\mathbf{r}\in\mathcal{R}^N$:

$$
\begin{aligned}
\alpha_t^{\mathcal{U}}(\mathbf{r}) \cdot \beta_t^{\mathcal{U}}(\mathbf{r}) &= \Pr\left(\mathbf{o}(1) \ldots \mathbf{o}(t), C_1 \ldots C_t, \mathbf{a}(t) = \mathbf{r} \,|\, \mathcal{K}\right) \cdot \\
&\quad \Pr\left(\mathbf{o}(t+1) \ldots \mathbf{o}(T), C_{t+1} \ldots C_T | \mathbf{a}(t) = \mathbf{r}, \mathcal{K}\right) && (2.32) \\
&= \Pr\left(\mathbf{o}(1) \ldots \mathbf{o}(t), C_1 \ldots C_t, \mathbf{a}(t) = \mathbf{r} \,|\, \mathcal{K}\right) \cdot && (2.33) \\
&\quad \Pr\left(\mathbf{o}(t+1) \ldots \mathbf{o}(T), C_{t+1} \ldots C_T | \mathbf{o}(1) \ldots \mathbf{o}(t), C_1 \ldots C_t, \mathbf{a}(t) = \mathbf{r}, \mathcal{K}\right) \\
&= \Pr\left(\mathbf{o}(1) \ldots \mathbf{o}(T), C_1 \ldots C_T, \mathbf{a}(t) = \mathbf{r} \,|\, \mathcal{K}\right) && (2.34) \\
&= \Pr\left(\{\mathbf{o}(t)\}_{t=1..T}, C, \mathbf{a}(t) = \mathbf{r} | \mathcal{K}\right) && (2.35)
\end{aligned}
$$

Using Bayes Rule, the *total law of probability* and (2.35), we get

$$
\begin{aligned}
h_t^{u_i}(r) &\triangleq \Pr\left(a_{u_i}(t) = r \,|\, \{\mathbf{o}(t)\}_{t=1..T}, C, \mathcal{K}\right) \\
&= \frac{\Pr\left(\{\mathbf{o}(t)\}_{t=1..T}, C, a_{u_i}(t) = r | \mathcal{K}\right)}{\Pr\left(\mathbf{o}(t)\}_{t=1..T}, C | \mathcal{K}\right)} && (2.36) \\
&= \frac{\sum\limits_{\mathbf{r} \in \mathcal{R}^N} \Pr\left(\{\mathbf{o}(t)\}_{t=1..T}, C, a_{u_i}(t) = r, \mathbf{a}(t) = \mathbf{r} | \mathcal{K}\right)}{\sum\limits_{\mathbf{r} \in \mathcal{R}^N} \Pr\left(\mathbf{o}(t)\}_{t=1..T}, C, \mathbf{a}(t) = \mathbf{r} | \mathcal{K}\right)} && (2.37) \\
&= \frac{\sum\limits_{\mathbf{r} \in \mathcal{R}^N \,|\, r_i = r} \Pr\left(\{\mathbf{o}(t)\}_{t=1..T}, C, a_{u_i}(t) = r, \mathbf{a}(t) = \mathbf{r} | \mathcal{K}\right)}{\sum\limits_{\mathbf{r} \in \mathcal{R}^N} \alpha_t^{\mathcal{U}}(\mathbf{r}) \cdot \beta_t^{\mathcal{U}}(\mathbf{r})} + \\
&\quad \frac{\sum\limits_{\mathbf{r} \in \mathcal{R}^N \,|\, r_i \neq r} \Pr\left(\{\mathbf{o}(t)\}_{t=1..T}, C, a_{u_i}(t) = r, \mathbf{a}(t) = \mathbf{r} | \mathcal{K}\right)}{\sum\limits_{\mathbf{r} \in \mathcal{R}^N} \alpha_t^{\mathcal{U}}(\mathbf{r}) \cdot \beta_t^{\mathcal{U}}(\mathbf{r})} && (2.38) \\
&= \frac{\sum\limits_{\mathbf{r} \in \mathcal{R}^N \,|\, r_i = r} \Pr\left(\{\mathbf{o}(t)\}_{t=1..T}, C, \mathbf{a}(t) = \mathbf{r} | \mathcal{K}\right)}{\sum\limits_{\mathbf{r} \in \mathcal{R}^N} \alpha_t^{\mathcal{U}}(\mathbf{r}) \cdot \beta_t^{\mathcal{U}}(\mathbf{r})} + 0 && (2.39) \\
&= \frac{\sum\limits_{\mathbf{r} \in \mathcal{R}^N \,|\, r_i = r} \alpha_t^{\mathcal{U}}(\mathbf{r}) \cdot \beta_t^{\mathcal{U}}(\mathbf{r})}{\sum\limits_{\mathbf{r} \in \mathcal{R}^N} \alpha_t^{\mathcal{U}}(\mathbf{r}) \cdot \beta_t^{\mathcal{U}}(\mathbf{r})} && (2.40)
\end{aligned}
$$

■

We now take a simple example. Consider regions $\mathcal{R}' = \mathcal{R} = \{a, b\}$ and no LPPM, hence $f(a,a) = f(b,b) = 1$, $f(a,b) = f(b,a) = 0$. Consider users $\mathcal{U} = \{u, v\}$ with mobility profiles $p_u = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$ and $p_v = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix}$ (thus $\pi_u = \pi_v = \begin{bmatrix} 0.5 & 0.5 \end{bmatrix}$). Assume users always report co-locations and do not lie, then $g_{u,v}(r_u, r_v) = \begin{cases} 1 & \text{if } r_u = r_v \\ 0 & \text{otherwise} \end{cases}$. Furthermore, assume one time instant ($T = 1$) and two observations at this instant: $u \leftrightarrow_1 v$ and $v \,@_1\, a$. By Equation (2.5), it follows that $\alpha_0^{\mathcal{U}}([\cdot, \cdot]) = 0.25$; $\alpha_1^{\mathcal{U}}([a, b]) = \alpha_1^{\mathcal{U}}([b, a]) = 0$ (because the likelihood of reporting being co-located when users' locations are different is 0); $\alpha_1^{\mathcal{U}}([b, b]) = 0$ (because $f(b, a) = 0$ - user v cannot report a while in b));

and $\alpha_1^{\mathcal{U}}([a,a]) = 0.25$. Similarly, by Equation (2.6), it follows that $\beta_1^{\mathcal{U}}([\cdot,\cdot]) = 1$; $\beta_0^{\mathcal{U}}([\cdot,a]) = 0.45$ and $\beta_0^{\mathcal{U}}([\cdot,b]) = 0.05$. Finally, using Equation (2.31) to localize user u at $t=1$ yields: $h_1^u(a) = 1$ and $h_1^u(b) = 0$. This result can easily be verified: if the adversary knows that u is with v and v is in a, u must also be in a.

We now evaluate the complexity of the joint localization attack. The first observation is that the size of the state space (*i.e.,* the locations of all users) is M^N. To attack a user at time t, the adversary needs to compute the values of α *up to* time t and the values of beta *down to* time t (using dynamic programming for optimal performance). At each time instant, the adversary needs to compute the values of these two variables for all possible values of their inputs $\mathbf{r} \in \mathcal{R}^N$ (there are M^N possible values for $\mathbf{r}$). The computation of each of these values requires summing over the M^N possible locations $\boldsymbol{\rho}$ at time $t-1$; for each of the possible locations, the computation of one element of the sum takes $\Theta(N^2)$ operations (the complexity of the computation of l dominates for the computation of β). Therefore, the computation of the forward and backward variables, at all time instants, for all possible values of the localizations is $\Theta(N^2TM^{2N})$ operations. Note that the complexity is the same whether the adversary attacks one or all the users at one or all time instants. In fact, the adversary can pre-compute the h_t^u for all u and all t, with a complexity that is dominated by that of the computations of the forward and backward variables. In summary, the complexity of the localization attack on one or all of the users in $\mathcal{U}$ is

$$c_{\text{opt}}(N,T,M) = \Theta(N^2TM^{2N}) \ . \tag{2.41}$$

The complexity of the optimal localization attack is prohibitively high and prevents its use for the entire set of users of a mobile social network; the optimal localization attack is tractable only for small values of N, *i.e.,* 2 or 3. In the next section, we propose low-complexity alternatives for performing low-complexity approximate localization attacks.

2.4 Approximate Localization Attacks

We propose two low-complexity alternatives for performing approximate localization attacks. Essentially, the first carefully selects a small set of users to consider when attacking a target user and performs an optimal joint localization attack on this small set of users (*i.e.,* considering only the co-locations between these users). The intuition behind this heuristic is that the locations of a user are significantly correlated with those of only a limited number of users (*e.g.,* a few co-workers during work hours, and her family and close friends the rest of the time). The second alternative makes use of all available location and co-location information (from all users) but only performs an approximate inference attack to localize users. We formulate the localization problem as a Bayesian network and apply a well-known inference algorithm, namely, loopy belief propagation.

2.4.1 Limited User-Set Heuristic

As discussed in Section 2.3, the optimal localization attack can be efficiently performed only on small sets of users. This is because the location of a target user u depends on locations of *all* other users that are connected to u in the co-location graph (where there is an edge between two users u and v if $u' \leftrightarrow_t v$ for some time t). The rationale of our first approximation is to limit the number of users, to whom the target user's location depends on, and to consider only those that have a high location correlation with u. Concretely, we choose the user(s) who have the largest number of reported co-locations with the targeted user, and we perform an optimal localization attack on the resulting set of users. We call these users the *co-targets* of the targeted user. Depending on his computational power, the adversary can choose one or two such users (*i.e.,* $N = 2$ or $N = 3$) to attack the target with. The co-targets of a user u are chosen as follows:

$$\begin{aligned}
\text{co-target}_1(u) &\triangleq \operatorname*{argmax}_{v \in \mathcal{U} \setminus \{u\}} \left|\{t \in \{1, \dots, T\} \,|\, u \leftrightarrow_t v\}\right| && (2.42)\\
\text{co-target}_2(u) &\triangleq \operatorname*{argmax}_{v \in \mathcal{U} \setminus \{u, u'\}} \Big[\left|\{t \in \{1, \dots, T\} \,|\, u \leftrightarrow_t v\}\right| + \\
&\qquad\qquad \left|\{t \in \{1, \dots, T\} \,|\, u' \leftrightarrow_t v\}\right| \Big] && (2.43)
\end{aligned}$$

where $u' = \text{co-target}_1(u)$ and $|\cdot|$ denotes the cardinality of the set. More specifically, the first co-target of a user u is the user with whom u has the most reported co-locations during the time interval considered for the localization attack. The second co-target of u is chosen so as to maximize the number of co-locations with u **plus** the number of co-locations with u's first co-target. Note that the set of considered users can be different for every targeted user; in particular $v = \text{co-target}_1(u) \not\Longrightarrow u = \text{co-target}_1(v)$. In practice, an adversary could also take into account the function g when choosing u's co-targets, to better reflect the trustworthiness of the reported co-locations. The complexity of this heuristic is $\Theta(TM^4)$ for $N = 2$ and $\Theta(TM^6)$ for $N = 3$ (obtained by replacing N by its value in the generic expression (2.41) of the complexity of the optimal attack).

2.4.2 Bayesian Network-Based Approximation

We propose using approximation algorithms on Bayesian networks, as a low-complexity alternative solution to the localization problem. A Bayesian network is a graphical model that encodes the probabilistic dependencies between different random variables of interest [33,34]. More specifically, a Bayesian network is a directed acyclic graph in which nodes represent random variables and where the edges model conditional dependence between the variables corresponding to the nodes they connect. In addition to its (graph) structure, a Bayesian network is also specified by its parameters: Each node has an associated conditional probability distribution (CPD) that specifies the probability that the corresponding variable will take a certain value, given a combination of values of the variables associated with its predecessor nodes. Modeling our problem as a Bayesian network enables us to exploit existing approximate inference algorithms, such as the belief propagation (BP) algorithm [34, 35] (which we use in the evaluation). BP is an algorithm that converges to the optimal solution by iteratively updating the posterior of a random variable, based on that of its neighbors and on its CPD, by using values of the observed variables. For Bayesian networks that do not contain undirected loops, which is *not* the case of our model, the BP algorithm converges to the optimal solution in only one iteration. Because of its iterative aspect, it balances (through the number of iterations) execution time and accuracy. A typical choice for the number of iterations is two times the number of nodes in the Bayesian network. Moreover, by running the BP-based solution, the adversary can obtain coarse-grained estimates of the users' locations after a few iterations and update them with better estimates as BP progresses. The heuristic

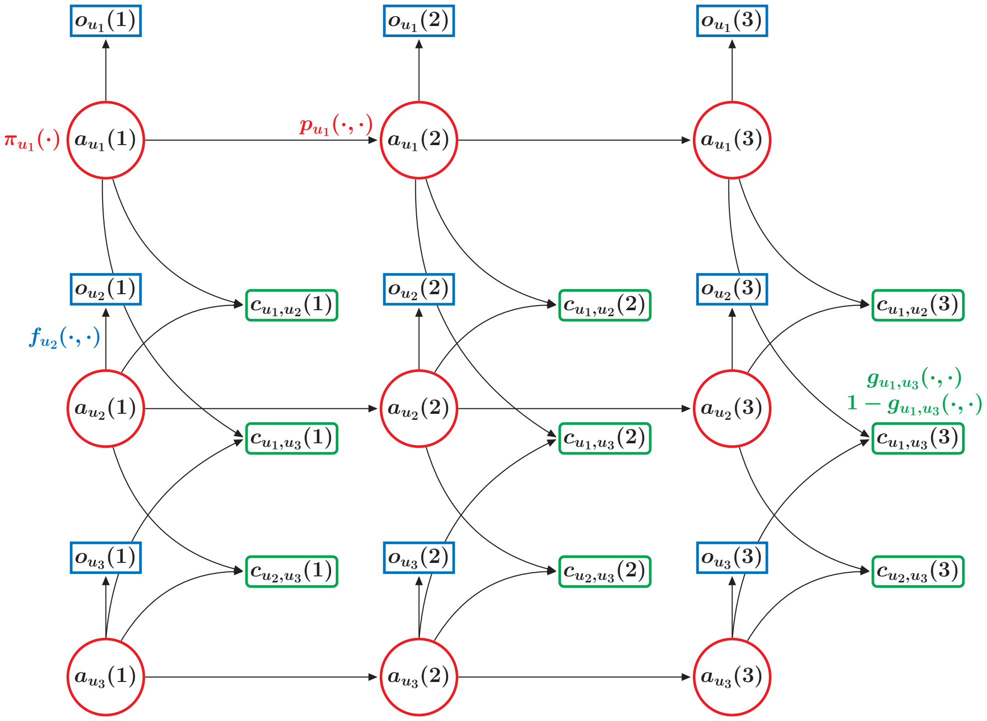


Figure 2.4: Sample Bayesian network for $N = 3$ users and $T = 3$ time instants. *Actual location nodes* are represented by red circles, *observed location nodes* by blue rectangles and *observed co-location nodes* by green rectangles with rounded corners. Probabilistic dependencies are specified by edges and conditional probability distributions (CPD), e.g., a co-location observation depends only on the actual locations of the two involved users and the probabilistic dependency is captured by g.

presented in the previous sub-section makes the most out of a subset of the available information (i.e., optimal inference on the data of the target user and her co-targets), whereas the BP-based solution only approximates the optimal solution but exploits all the available information (approximate inference on the data of all the users).

We build a Bayesian network, as illustrated in Figure 2.4 (for $N = 3$ and $T = 3$): For any user u and any time instant t, a node is associated with the variable $a_u(t)$ and another with the variable $o_u(t)$. To represent the fact that the observed location depends only on a user's actual location at that time, an edge connects the corresponding nodes and the corresponding CPD is f_u. Additionally an edge connects the node corresponding to a user u's actual location at time t to her actual location node at time $t+1$, with its CPD determined by her mobility profile p_u (following from the Markov assumption). For any pair u, v of users and any time instant, an observed co-location node is associated with variable $c_{u,v}(t)$, with its CPD specified by $g_{u,v}$ (it depends on the actual location of the two users involved). Our Bayesian network consists of $T \cdot N$ actual/observed location nodes and $T \cdot N(N-1)/2$ observed co-location nodes.[5] Location nodes have one incoming

[5] Note that when the probability of two users reporting a co-location between them is null (e.g., non-friend users in a social network), the corresponding nodes can be removed. As suggested by Dunbar's number [36], a user has a limited number of friends. Therefore, in many contexts, the number of co-location nodes grows linearly with N.

edge, and co-location nodes have two. Consequently, the complexity for one iteration of the belief propagation algorithm is $O(N^2 \cdot T \cdot M^2)$.

Specifically, for nodes $\{a_u(1)\}_{u\in\mathcal{U}}$, we define the CPD $\forall r \in \mathcal{R}$ as

$$\Pr\left(a_u(1) = r\right) = \pi_u(r) \tag{2.44}$$

For nodes $\{a_u(t)\}_{u\in\mathcal{U},t\in\{2,\dots,T\}}$, we define the CPD $\forall \rho, r \in \mathcal{R}$ as

$$\Pr\left(a_u(t) = r \mid a_u(t-1) = \rho\right) = p_u(\rho, r) \tag{2.45}$$

For nodes $\{o_u(t)\}_{u\in\mathcal{U},t\in\{1,\dots,T\}}$, we define the CPD $\forall r, r' \in \mathcal{R}$ as

$$\Pr\left(o_u(t) = r' \mid a_u(t) = r\right) = f_u(r, r') \tag{2.46}$$

Finally, for nodes $\{u \leftrightarrow_t v\}_{u,v\in\mathcal{U},t\in\{1,\dots,T\}}$, we define the CPD $\forall r, r' \in \mathcal{R}$ as

$$\Pr\left(u \leftrightarrow_t v = T \mid a_u(t) = r, a_v(t) = r'\right) = g_{u,v}(r, r') \tag{2.47}$$

$$\Pr\left(u \leftrightarrow_t v = F \mid a_u(t) = r, a_v(t) = r'\right) = 1 - g_{u,v}(r, r') \tag{2.48}$$

We compare the approximate localization attack to the optimal localization attack, and we measure its accuracy by the average Hellinger and statistical distance between their output region distributions. Specifically, if h denotes the output of the optimal localization attack $\hat{h}$ that of the approximate localization attack, then

$$\frac{1}{N\cdot T}\sum_{u\in\mathcal{U}}\sum_{t\in\{1,\dots,T\}}\frac{1}{\sqrt{2}}\sqrt{\sum_{r\in\mathcal{R}}\left(\sqrt{h_t^u(r)}-\sqrt{\hat{h}_t^u(r)}\right)^2}$$

$$\frac{1}{N\cdot T}\sum_{u\in\mathcal{U}}\sum_{t\in\{1,\dots,T\}}\frac{1}{2}\sum_{r\in\mathcal{R}}\left|h_t^u(r)-\hat{h}_t^u(r)\right|.$$

2.5 Countermeasures

So far, we have presented and analyzed a localization attack that exploits co-location information. In this section, we propose two countermeasures that mitigate the (negative) effect of co-locations on the users' location privacy. These countermeasures apply to the case where users explicitly report their co-locations, typically on a social network. For co-location information leaked by the underlying technologies, such as IP addresses and Bluetooth and Wi-Fi scans of neighboring devices, technology-dependent techniques should be used. For instance, a user can hide her IP address from the service provider by using a proxy, a VPN or a peer-to-peer anonymization network such as Tor. Note that countermeasures are not limited to those presented in this section. Altering the individual LPPM settings (the value of λ, using obfuscation or cloaking) would also reduce, to some extent, the privacy risk. Unfortunately there is not much else a user can do to protect herself, other than hide or generalize co-location information or prevent it from being inferred. In practice, this would translate to hiding IP addresses, disabling Bluetooth, or blurring faces in pictures posted on online social networks, as proposed in [37]. Simply put, the proposed countermeasures operate as follows: The first consists in making co-located users report the same (obfuscated) location and the second consists in generalizing time and/or user information in the reported co-locations.

2.5.1 Coordinated LPPMs

In order to make the inference attacks we described in previous sections less effective, we propose a simple countermeasure: user *coordination*. This means that if users report being co-located at some time instant and also want to report obfuscated individual check-ins, they should coordinate (*i.e.,* report the *same obfuscated* location). Such a mechanism requires collaboration between users, which can be challenging to achieve in practice. A possible solution, in the case of explicitly reported co-locations, is that a user who posts a co-location information embeds her obfuscated location so that all the co-located users report the same obfuscated location (if they do report their locations). Collaboration could also be achieved by means of short-range ad-hoc communication technologies, such as Wi-Fi Direct or Bluetooth, as the co-located users are physically close. We emphasize that this does not mean that co-located users have to also report individual check-ins, rather that if they want to report individual check-ins, they must agree to make them the same. We argue this would bring no detriment to users' utility of individual check-ins, as the obfuscation mechanism selects a *random* neighboring location to the actual location, which users have no control or preference over. Intuitively, reporting single co-locations in a coordinated fashion should give an adversary less information, because it maximizes the set of possible locations co-located users could be in. As described in Figure 2.1 on page 8, based on individual check-ins of co-located users, an adversary can infer that both users should be located in the intersection of possible locations of each of the co-located users. With coordination, the possible locations of users are the same, thus maximizing their intersection. Note that this countermeasure has an effect only if both users use obfuscation.

2.5.2 Generalization of Co-locations

We propose another countermeasure for reducing the effectiveness of inference attacks that make use of co-location information. In the case of single location observations, a recommended privacy-protection technique is obfuscation by generalization (i.e., report a large area that contains the user's actual location). Similarly, we propose that users generalize co-location information, in a coarse-grained fashion; specifically, this implies generalizing the time component of a co-location, and/or the co-located user(s) component. Generalizing the time component in a co-location information means reporting a time range instead of the exact time (*e.g.,* use "morning" instead of "10am"). Generalizing the user component means excluding the names of the friends a user is with and reporting only the *number* of friends (*e.g.,* instead of reporting being with her friend Alice, a user would just report being with *a friend*). More generally, in the case where a user u is co-located with k friends $u'_1, u'_2, \ldots, u'_k$ he would no longer report k pairwise co-locations with each of them ($u \leftrightarrow_t u'_1, \ldots, u \leftrightarrow_t u'_k$), but instead report one *generalized co-location* $u \leftrightarrow_t k$ friends. The user component of the co-location could also contain information on social ties, such as "with two colleagues", or "with some friends".

We analyze in more depth the case of generalizing the co-located user(s) component of a co-location. Intuitively, if this mechanism is employed by the users, it is harder for the adversary to exploit a co-location information because he has to explore all possible combinations of real users a user is with and assign a likelihood to each of them. This leads to $\binom{N_u^{\text{colleagues}}}{k}$ possible choices for exploring the generalized co-location "with k colleagues", where $N_u^{\text{colleagues}}$ is the number of user u's colleagues in the social network, and $2^{N_u^{\text{friends}}+N_u^{\text{family}}}-1$ choices for the generalized co-location "with some friends and/or family". More specifically, the joint variables $\alpha_\cdot^{\mathcal{U}}(\cdot)$ and $\beta_\cdot^{\mathcal{U}}(\cdot)$ (Equation (2.5) and Equation (2.6)) would include a summation in the computation of the likelihood of observing the obfuscated co-locations ($l_\cdot(\cdot, C)$) for *all* possible instantiations of *all* reported co-locations at time t (for α) or $t+1$ (for β) by *all* users. This would drastically increase the complexity of the optimal inference attack. Note that generalizing the user component of the co-locations would also drastically increase the complexity of the BP-based solution; the current Bayesian network (Figure 2.4 on page 21) could not be used anymore. We will investigate this as part of future work. In summary, this countermeasure protects the users' privacy by making the inference prohibitively computationally expensive for the adversary.

Obfuscating the time component of co-locations would also lead to a drastic increase in complexity because the adversary would have to consider all combinations of exact time instances when users are co-located (which makes the computation of the joint α, β variables nontrivial). Naturally, obfuscating both components of co-location information would result in the greatest complexity increase. We leave the design and in-depth analysis of inference algorithms when a combination of the proposed counter-measures is employed by users to future work. We intend to analytically evaluate the inference complexity and empirically evaluate the users' *privacy gain* and potential *utility loss* in different scenarios of employed countermeasures.

2.6 Differential-Privacy Perspective

In this section, we complement our inferential approach to privacy quantification, presented in the previous sections, with a brief analysis of the effect of co-locations on users' location privacy from a differential-privacy perspective. In the geo-indistinguishability framework [38,39] (i.e., the application of differential privacy to geo-location), each observation has a privacy cost that depends on the level of noise added by the mechanism used (typically drawn from a planar Laplace distribution). For instance, in order to guarantee ε-differential privacy, noise must introduced with an amplitude such that the expected distance between the actual location and the reported location is proportional to $1/\varepsilon$. Consider the case of a single time instant. If two co-located users each report one obfuscated version of their actual locations, the adversary has access to two observations of the same variable, i.e., the users' common location. Following the composability property of differential privacy, this means that, to guarantee ε-differential privacy for the users' location, each individualy reported obfuscated location should satisfy $(\varepsilon/2)$–differential privacy (unless the two users agree on reporting the same obfuscated location, as discussed in Section 2.5 dedicated to countermeasures). This means that the expected distance between the users' actual locations and the obfuscated locations they report is doubled, thus causing a substantial utility loss. This reasoning can be generalized to an arbitrary number of co-located users: Intuitively, at every time instant, the level of noise a user must introduce (and thus the utility loss she faces), in order to retain the same privacy level in the presence of co-locations, is proportional to the number of co-located users.

A study by Kifer et al. [40], effectively shows that if data correlations are ignored, the released data will have a privacy guarantee lower than expected. A more complex analysis of the effect of co-locations, from a differential-privacy perspective, could be carried out

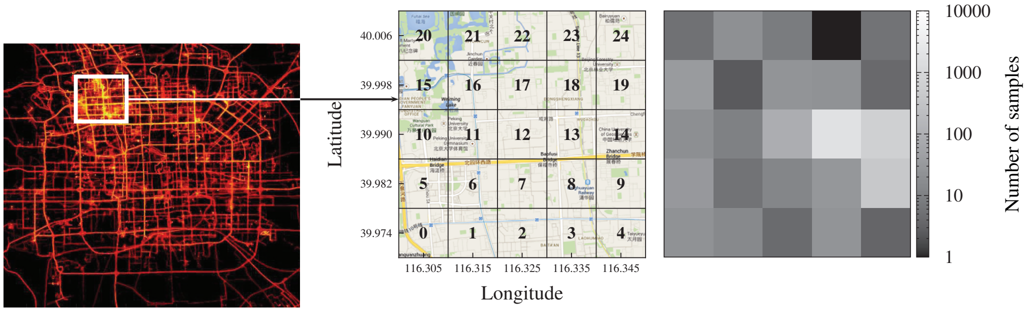


Figure 2.5: Illustration of the dataset used in the evaluation. Most traces are located in the region of Beijing (left); we focus on a small active area that corresponds to the campus of Tsinghua University and we partition it by using a 5×5 square grid (middle). The heat-map (right) shows the number of samples in each region (logscale), for the users of interest

by using the Pufferfish framework [41, 42] (or more recently [43–45]) that enables taking into account the correlation between entries in a differential-privacy analysis.[6]

2.7 Experimental Evaluation

Using a dataset of mobility traces, we evaluate the effect of co-locations on users' privacy, with respect to the various localization attacks presented in the previous sections.

2.7.1 Dataset, Methodology, and Experimental Setup

The dataset was collected by Microsoft Research Asia, in the framework of the GeoLife project [48]. It comprises the GPS traces (*i.e.,* sequences of time-stamped latitude-longitude couples, sampled at a rate of one point every 1-5 seconds) of 182 users, collected over a period of over three years. The GPS traces are scattered all over the world; but most of them are located in the region of Beijing, China. We processed the data as follows, in order to fit in our formalism.

Space Discretization. We select the area of $\sim$4.4 km × 4.4 km, within Beijing, that contains the largest number of GPS samples, and we filter out GPS samples that are outside of this area. This geographic area corresponds to the campus of Tsinghua University (longitude ranging from 116.3 to 116.35 and latitude ranging from 39.97 to 40.01, see Figure 2.5). We partition the selected area into 25 regions by using a $5{\times}5$ square grid. The GPS coordinates of each sample are translated into the region (*i.e.,* the grid cell) they fall into.

Time Discretization. We divide the continuous time interval into one-hour time sub-intervals that correspond to time instants in our formalism. For each time sub-interval t and for each user u, we set the user's actual location in that time interval (*i.e.,* $a_u(t)$) to the region corresponding to the sample that is the closest to the midpoint of the

[6] Since the publication of our article [22], several other studies (*e.g.,* [46, 47]) have proposed methods to effectively take data correlations into account when proposing protection techniques based on differential privacy.

considered time sub-interval. If a user's trace does not contain any samples in a given time sub-interval, the user's actual location is set to a dummy region $r_\perp$, leaving us with partial user traces.

Co-location Generation. As the dataset does not contain explicit co-location information reported by the users, we use synthetic co-locations that we generate as follows: At each time instant, we generate a co-location between two users according to the probabilistic co-location reporting function $g_{\cdot,\cdot}(\cdot,\cdot)$, based on their discretized actual locations (if they are different from $r_\perp$). We consider a special case of the co-location reporting function (Equation (2.1)) as follows:

$$g_{\cdot,\cdot}(r_u, r_v) = \begin{cases} \nu & \text{if } r_u = r_v \\ \mu & \text{if } r_u \neq r_v \end{cases} \tag{2.49}$$

As stated in the model, the adversary is assumed to know the values of μ and ν. Intuitively, μ represents the probability a *fake* co-location is reported, and ν represents the probability a *true* co-location is reported. This model assumes that for any user, reporting a co-location does not depend on the actual location where she and her friend are and that the user chooses to report their co-location with a fixed probability. In order to simplify the evaluation, we assume that the co-location reporting function is the same among any pair of users, as in the case of a Bluetooth scenario. We could relax this assumption and make ν and μ functions of the particular pair of users; for example, if a social graph of relationships between users were available, we could consider $\nu, \mu > 0$ only for pairs of users for which a social relationship exists, and 0 for all other user pairs, as users typically report co-locations on social networks only with their friends. Regarding the values of ν and μ, several cases can also be considered: $\nu = 1$ and $\mu = 0$ would correspond, for example, to an ideal Bluetooth scenario, in which devices automatically discover each other and report co-locations with all neighboring devices; $\nu < 1$ and $\mu = 0$, could also correspond to a Bluetooth scenario, where co-locations are reported with only some of the neighboring devices. In our evaluation, we will consider both cases.

For each user, we compute the number of *real* co-locations[7] she has with every other user in the dataset, across the full user traces. We keep only the users for which there exists another user with whom they have at least 200 co-locations. For these users, we consider their *common* time interval (*i.e.,* the longest time interval during which all these users have at least one sample); we obtained an interval of ~6000 hours. Within this interval, we sample 10 short traces of 300 continuous hours such that (1) all users have at least 10% of valid samples (*i.e.,* , different from $r_\perp$) and (2) all users have at least 20 co-locations with their co-target$_1$ (as defined in Eq. (2.43)). This leaves us with a total of 5 users.

User Mobility Profile Construction. We build the mobility profiles $\{p_u\}_{u\in\mathcal{U}}$ of the users based on their entire discretized traces by counting the transitions from any region to any region (in $\mathcal{R}$) in one time instant.

Obfuscation. We consider that users report a single (or none), potentially obfuscated, location at each time instant.[8] This means that the set $\mathcal{R}'$ in which the obfuscated loca-

[7] Note that, by real co-locations, we mean that the users are at the same location (i.e., their actual locations at a given time instant are the same), regardless of the fact that the co-location is reported or not.

[8] We assume this because of the limited size of the considered grid.

tion $o_u(\cdot)$ takes values is $\mathcal{R} \cup \{r_\perp\}$. We consider, for each user u, that two location-privacy protection mechanisms are used together: First, the location is hidden (*i.e.,* obfuscated to $r_\perp$) with a probability λ_u and then, if the location has not been hidden, it is replaced by a region (chosen uniformly at random) at a distance of at most d_u from the user's actual discretized location (*i.e.,* a region). If the actual location of a user is not known (*i.e.,* set to $r_\perp$), the LPPM returns $r_\perp$ with probability 1. In our evaluation, we vary λ_u from 0 to 1 and we set d_u to the size of one grid cell; this means that, if it is not hidden, a user's location is obfuscated either to its actual value (with probability 0.2) or to one of the four adjacent regions (*e.g.,* 2, 6, 8 and 12 for Region 7 in Figure 2.5 on page 25), each with probability 0.2.

Privacy Evaluation. We evaluate the location privacy of the users based on the metric defined in (2.3). For each user and for each short trace, we generate 20 random obfuscated traces (remember that obfuscation is a random process), and we perform a localization attack on each of them. We compute the average location privacy of each user across the different obfuscated traces and across the different time instants. Time instants for which the location of a user is not known (*i.e.,* set to $r_\perp$) are not taken into account in the computation of their average over time.

Limitations. Unfortunately, we could not obtain real datasets from online social networks containing both (coarse-grained) location and co-location data. Due to the synthetic nature of the reported location and co-location information in our data source, our experimental setup does not perfectly reflect a real usage case. Deciding whether to report locations and co-locations is a complicated process that involves many factors, such as the users' contexts, their privacy preferences, and the shared information itself. Therefore, the results presented in this section should be taken with a pinch of salt as they cannot directly be interpreted as the magnitude of the threat in real life. Yet, we believe that they are significant enough for understanding the effect of co-locations on location privacy, the sources of privacy loss, and the relative performance of the proposed heuristics. Also, the number of users considered in most of our evaluations (*i.e.,* 5), as well as the active area considered, are relatively small. Hence, the results might not be representative of the entire population. In order to overcome the aforementioned shortcomings, we intend to collect a large-scale dataset from an existing social network. We also intend to run experiments on large grids (*i.e.,* larger than 5×5).

2.7.2 Experimental Results

We experimentally evaluate the algorithms, presented in Section 2.4, in different scenarios, with different settings. For the solution based on belief propagation, we relied on the implementation provided in the Bayes Net Toolbox for Matlab (https://code.google.com/p/bnt/); for the optimal inference algorithm, we used our own Java implementation. The purpose of our evaluation is to assess the raw performance of our algorithms, but also to compare their results. We also analyze the effect of the different parameters of the model (including the individual LPPM settings of the users and the *differences* between the individual LPPM settings of the users) and of the set of co-locations considered in the localization attack.

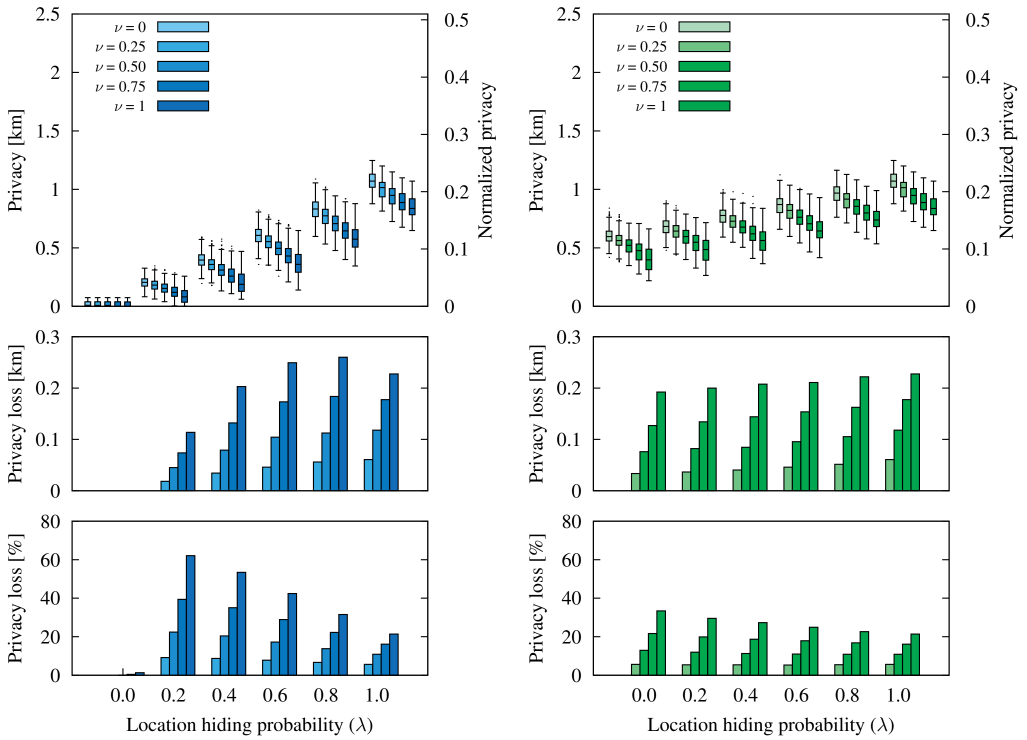
Effects of True Co-locations and LPPM Settings

We begin our evaluation by analyzing the effect of (1) the amount of reported true co-locations and (2) the LPPM settings (*i.e.,* with or without obfuscation and the location hiding probability λ, assumed to be the same across users) in the case of two users, *i.e.,* the target user and her first co-target are considered jointly in an optimal localization attack, namely the limited user set approximation with $N = 2$. For this evaluation, we consider the case where no fake co-locations are reported. The results are depicted in Figure 2.6 on page 29. Figure 2.6a on page 29 shows the case where no obfuscation is used (*i.e.,* the users disclose their *actual* locations with probability $1 - \lambda$ and hide them otherwise), and Figure 2.6b on page 29 shows the case where obfuscation is used (*i.e.,* the users disclose their *obfuscated* locations, specifically a region chosen uniformly at random among the actual location and the four immediate neighboring regions, with probability $1-\lambda$ and hide them otherwise). The top graphs show a box-plot representation (*i.e.,* first quartile, median, third quartile and outliers) of the users' location privacy expressed in terms of the expected error of the adversary, in kilometers (left axis) and in proportion of the size of the considered geographic area (right axis). For each couple of values (λ, ν), we draw one box-plot to aggregate the data-points obtained for all users and for all the 20 randomly generated obfuscated versions of each of the considered actual traces. Note that without obfuscation, the case $\lambda = 0$ leads to zero privacy, as users *always* disclose their *actual* locations. It can be observed that the proportion of reported true co-locations consistently decreases the location privacy of the users. To quantify this decrease, we plot (middle and bottom graphs) the privacy loss caused by the use of co-location information, with respect to the case where no true co-locations are reported, *i.e.,* $\nu = 0$. We show both the median absolute privacy loss, in kilometers (middle graph), and the median relative privacy loss, in percentage of the privacy in the case $\nu = 0$ (bottom graph). Note that the median privacy loss is equal to the median of the differences (w.r.t. the case $\nu = 0$) and **not** to the difference of the median privacy.
Consider for example, the case $\lambda = 0.4$ and $\nu = 0.5$: In the case without obfuscation, the median privacy loss is approximately 80m, which corresponds to a decrease of approximately 21%. The median absolute privacy loss can go up to 260m ($\lambda = 0.8$, $\nu = 1$) and the median relative privacy loss up to 62% ($\lambda = 0.2$ and $\nu = 1$). We observe the same trend, with a more modest loss, in the case where obfuscation is used. We emphasize that when there is an obfuscated location observation, the adversary has only five choices of cells to locate the user: the cell of her actual location and four neighboring cells of the actual location. Hence, an upper bound for privacy in this case is given by the inter-cell distance (0.87km). It can be observed in Figure 2.6b on page 29 that when all observations are available ($\lambda = 0$), this upper bound is indeed respected.

Effects of True Co-locations and Spatial Cloaking

For a complementary experiment, we also studied the effect of co-location information when users employ spatial cloaking instead of obfuscation.

Similarly to our experimental setup presented in Figure 2.6b on page 29, we evaluate user privacy for a different LPPM, namely, location hiding (with probability λ) or spatial cloaking (with probability $1-\lambda$). When using cloaking, a user does not report the region corresponding to her actual location, but instead a meta-region consisting of four regions, one of which is the actual location. In Figure 2.7 on page 30, we present our results.



(a) Without obfuscation

(b) With obfuscation

Figure 2.6: Privacy (top), absolute privacy loss (middle) and relative privacy loss (bottom) for the limited user set attack with $N = 2$ users , when users do not report fake co-locations ($\mu = 0$). The privacy *loss* is expressed w.r.t. the case where no co-locations are available ($\nu = 0$, $\mu = 0$); the histograms show median values. Co-location information decreases privacy. The relative privacy loss is higher for small values of the hiding probability and without obfuscation.

We conclude that the proportion of reported true co-locations consistently decreases the location privacy of the users (as was the case for the other LPPM based on location hiding and location obfuscation), but in this case the privacy loss is more evident. This could be explained by the fact that in the case of cloaking, when observing a meta-region of size four regions, the adversary has to explore four possible regions as candidates for the user's actual location; whereas, in the case of obfuscation, five possible candidates for the actual location have to be explored (one of the four neighboring regions of the observed (obfuscated) region and the observed region itself).

In the next sections, we focus on the case where users obfuscate their locations, report true co-locations with probability $\nu = 0.5$ and do not report fake co-locations ($\mu = 0$).

Effects of the Differences of Individual LPPM Settings

We now analyze the effect of the differences, in the users' LPPM settings, on the location privacy (loss) due to co-locations. To do so, we focus on the case of two users, a target and her co-target, both who obfuscate their locations but with different hiding probabilities λ_{target} and $\lambda_{\text{co-target}}$. We perform a joint optimal localization attack. The results are

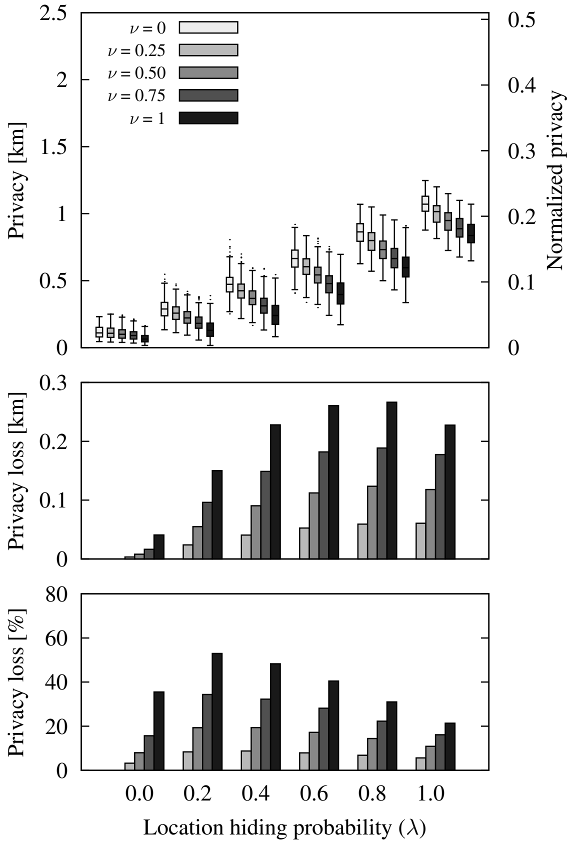


Figure 2.7: Privacy (top), absolute privacy loss (middle) and relative privacy loss (bottom) for the limited user set attack with $N = 2$ users, when users do not report fake co-locations ($\mu = 0$) and use spatial cloaking or location hiding as protection mechanisms. The privacy *loss* is expressed w.r.t. the case where no co-locations are available ($\nu = 0$, $\mu = 0$); the histograms show median values.

depicted in Figure 2.8 on page 32, under the form of heat-maps that represent the target user's location privacy (a) as well as her absolute (b) and relative (c) privacy loss (with respect to the case $\nu = 0$) as functions of the respective LPPM settings $\lambda_{\text{co-target}}$ (x-axis) and λ_{target} (y-axis).

A first observation is that co-locations always decrease the privacy of the target (*i.e.,* all values in Figure 2.8b on page 32 are positive) and that the more information the co-target discloses, the worse the privacy of the target is (*i.e.,* the cells of the heat-map depicted in Figure 2.8a on page 32 become lighter, when going from right to left on a given row).

The diagonals of the heat-maps correspond to the case $\lambda_{\text{co-target}} = \lambda_{\text{target}}$, which is depicted in more detail in Figure 2.6 on page 29. The region of the heat-map above the diagonal corresponds to the case where the target is more *conservative*, in terms of her privacy attitude, than her co-target (*i.e.,* $\lambda_{\text{co-target}} < \lambda_{\text{target}}$). It can be observed that the information disclosed by the target herself compromises her privacy more than the information disclosed by her co-target, *e.g.,* the cell (0.6,0) is lighter (which means that the target's privacy is lower) than the cell (0,0.6).

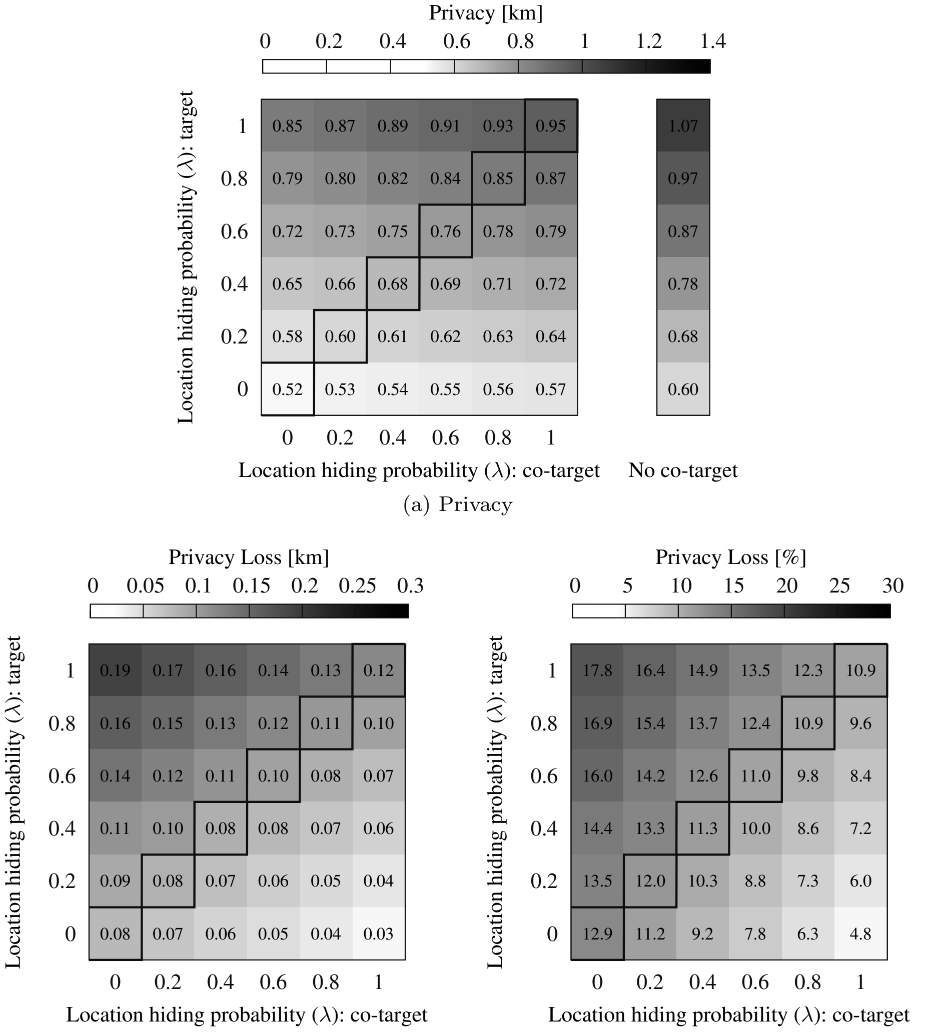
By comparing the columns "$\lambda_{\text{co-target}} = 1$" and "no co-target" (two right-most columns in Figure 2.8a on page 32), we can observe the privacy loss that stems from the use, through the co-location information, of the co-target's mobility profile alone (as the co-target never discloses her location). This is substantial. The intuition behind this result is that co-located users are likely to be at a place that is often visited by *both* of them, which narrows down the choice of locations the adversary needs to explore when localizing both users.

Finally, in the extreme case where the target never discloses location information and her co-target always does so (top-left cell of the heat-maps in Figures 2.8b and 2.8c on page 32), the privacy loss for the target is 190m, which corresponds to a decrease of 18%. This case (and in general the cases where the target never discloses location information, *i.e.,* the top row of the heat-maps) highlights the fact that, as reported co-locations involve two users, users lose some control over their privacy: Without revealing any information about herself, a user can still have her privacy decreased by other users, due to co-location information.
For the rest of the evaluation, we focus on the case where all users have the same LPPM settings (*i.e.,* same values of λ).

Comparison of the Proposed Low-Complexity Alternatives

Here, we compare, through experimentation, the proposed inference algorithms for the localization attack, by taking into account different scenarios, as depicted in Figure 2.9 on page 33. We assume all users use the same LPPM settings, *i.e.,* same value for λ and disclose only their obfuscated locations. In Scenario (a), we consider, in turn, all target users in our set and perform an individual localization attack on each of them, using only their own reported locations and no co-locations. This corresponds to the baseline case $\nu = 0$, which was presented in detail in Figure 2.6b on page 29. We then consider the case of an adversary that exploits co-locations. We assume users report only a limited proportion of their true co-locations, with probability $\nu = 0.5$, and no fake co-locations ($\mu = 0$). Scenario (b) corresponds to the case of an adversary that, in order to attack a target user, performs an optimal joint inference attack on the target and her co-target, as described in Section 2.3. This scenario corresponds to the case $\nu = 0.5$ in Figure 2.6b on page 29. Scenarios (c) and (d) correspond to the case of an adversary that performs an optimal joint attack on the target and her **two co-targets**. We distinguish two cases: (c) in which the co-locations between the co-targets are ignored, and (d) in which all co-locations between any of the three users are considered. We make this distinction solely to quantify the privacy loss that stems from the use of co-locations that do not directly involve the target. In practice, an adversary would always consider Scenario (d) because it takes into account more information at no extra cost. Finally we consider Scenario (e) that corresponds to an adversary that uses *all* reported co-locations but solves an *approximate* joint inference problem, as described in Section 2.4.2. We set the maximum number of iterations of the BP algorithm to 20.

Figure 2.10 on page 34 shows the results of our comparison. The top graph shows a box-plot representation of users' privacy, for each of scenarios (a)-(e). To quantify the different effects on the users' privacy of the set of considered co-locations and of the inference algorithm used, we show (bottom) the absolute and relative privacy loss, with respect to Scenario (a), for each of the scenarios (b)-(e). It can be observed by comparing



Location hiding probability (λ): target	0	0.2	0.4	0.6	0.8	1	No co-target
1	0.85	0.87	0.89	0.91	0.93	0.95	1.07
0.8	0.79	0.80	0.82	0.84	0.85	0.87	0.97
0.6	0.72	0.73	0.75	0.76	0.78	0.79	0.87
0.4	0.65	0.66	0.68	0.69	0.71	0.72	0.78
0.2	0.58	0.60	0.61	0.62	0.63	0.64	0.68
0	0.52	0.53	0.54	0.55	0.56	0.57	0.60

(a) Privacy

Location hiding probability (λ): target	0	0.2	0.4	0.6	0.8	1
1	0.19	0.17	0.16	0.14	0.13	0.12
0.8	0.16	0.15	0.13	0.12	0.11	0.10
0.6	0.14	0.12	0.11	0.10	0.08	0.07
0.4	0.11	0.10	0.08	0.08	0.07	0.06
0.2	0.09	0.08	0.07	0.06	0.05	0.04
0	0.08	0.07	0.06	0.05	0.04	0.03

(b) Absolute privacy loss (w.r.t. to $\nu = 0\%$)

Location hiding probability (λ): target	0	0.2	0.4	0.6	0.8	1
1	17.8	16.4	14.9	13.5	12.3	10.9
0.8	16.9	15.4	13.7	12.4	10.9	9.6
0.6	16.0	14.2	12.6	11.0	9.8	8.4
0.4	14.4	13.3	11.3	10.0	8.6	7.2
0.2	13.5	12.0	10.3	8.8	7.3	6.0
0	12.9	11.2	9.2	7.8	6.3	4.8

(c) Relative privacy loss (w.r.t. to $\nu = 0\%$)

Figure 2.8: Median values of the target's location privacy (loss), for the limited user set attack with $N = 2$ users, when the target and her co-target have different values of λ (with obfuscation, $\nu = 0.5$, $\mu = 0$). The diagonals correspond to the values of Figure 2.6b on page 29.

scenarios (a)-(d) that, unsurprisingly, the users' privacy decreases with the amount of considered co-locations. The comparison between scenarios (c) and (d) shows that co-locations between the target's co-targets improve the performance of the localization attack, but not as much as co-locations that directly involve the target user (Scenario (b) and Scenario (c)). Finally, we observe that the approximation based on belief propagation (Scenario (e)), which takes into account all co-locations and the location information of all the users, outperforms the first heuristic ($N \leq 3$), at a low computational cost. In this scenario, the median absolute privacy loss can go up to 182m ($\lambda = 0.8$) and the median relative privacy loss up to 27% ($\lambda = 0$), when $\nu = 0.5$ and $\mu = 0$. We can thus conclude that, when using belief propagation instead of joint optimal inference, the loss in inference accuracy is far less than the gain that stems from using all of the available

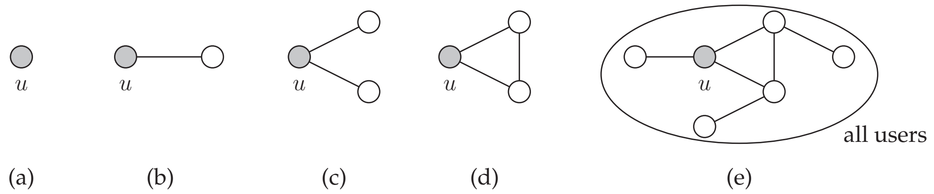


Figure 2.9: Co-locations considered in the evaluation: (a) no co-locations (also referred to as "*No co-target*"), (b) only co-locations between the target and co-target$_1$ (heuristic, $N = 2$), (c) only co-locations between the target and co-target$_1$ and between the target and co-target$_2$ (heuristic, $N = 3$), (d) all co-locations between the target, co-target$_1$ and co-target$_2$ (heuristic, $N = 3$), (e) all co-locations (belief propagation in our proposed Bayesian network formalization depicted in Figure 2.4 on page 21).

co-location information and the location information of all the users.

In order to assess the performance of the belief propagation algorithm, we also compared it with the optimal inference algorithm, for all scenarios (a)-(d). For each of these scenarios, we computed the Hellinger distance between the BP algorithm and the optimal inference. We obtained the following distances: 3.79E-4 for Scenario (a), 5.18E-3 for Scenario (b), 1.79E-2 for Scenario (c) and 3.31E-2 for Scenario (d). Similarly, we computed the statistical distances and obtained the following: 1.86E-4 for Scenario (a), 3.79E-3 for Scenario (b), 1.84E-2 for Scenario (c) and 3.10E-2 for Scenario (d). These very small values for both the Hellinger and statistical distance, for all scenarios, show that the BP algorithm converges in about 20 iterations, while also proving that our formulation of the localization problem as a Bayesian network (depicted in Figure 2.4 on page 21) is correct. In fact, we observe that the approximation provided by the BP algorithm is already quite close to the optimal after a very small number of iterations (i.e., 2-3) which suggests that the attack can be carried out efficiently by the adversary.

To further analyze and compare the performance of the different inference algorithms, we measured their execution times in a typical setting ($\lambda = 0.2$, $\nu = 0.5$ and $\mu = 0$, for a single user) on an 8-core Intel(R) Xeon(R) CPU E3-1270 V2 @ 3.50GHz with 16GB of RAM. We obtained the following results: Scenario (a): 1.82±0.0471s; Scenario (b): 3.87±0.0498s; Scenario (c): 2,711±91s; Scenario (d): 2,555±73.6s; Scenario (e): 2.66±0.0151s. These results demonstrate the practicality of the BP-based attack in comparison to the optimal localization attack, which is already very expensive for Scenarios (c) and (d), where co-locations between three users are considered.

Effects of Users Reporting Fake Co-locations

Here, we analyze the effect of reporting fake co-locations. We focus on the case of two users, a target and her first co-target, who both obfuscate their locations and use the same location hiding probability. We present the case where $\lambda = 0.2$, but we observe the same trend for all values of λ. We vary ν – the probability of reporting true co-locations, as well as μ – the probability of reporting fake co-locations. We perform a joint optimal localization attack. The results are depicted in Figure 2.11 on page 35. The top graph shows a box-plot representation of users' privacy, and the middle and bottom graphs show the median absolute and relative privacy loss, with respect to Scenario (a) (where no co-location information is considered). We observe that when all the true co-location

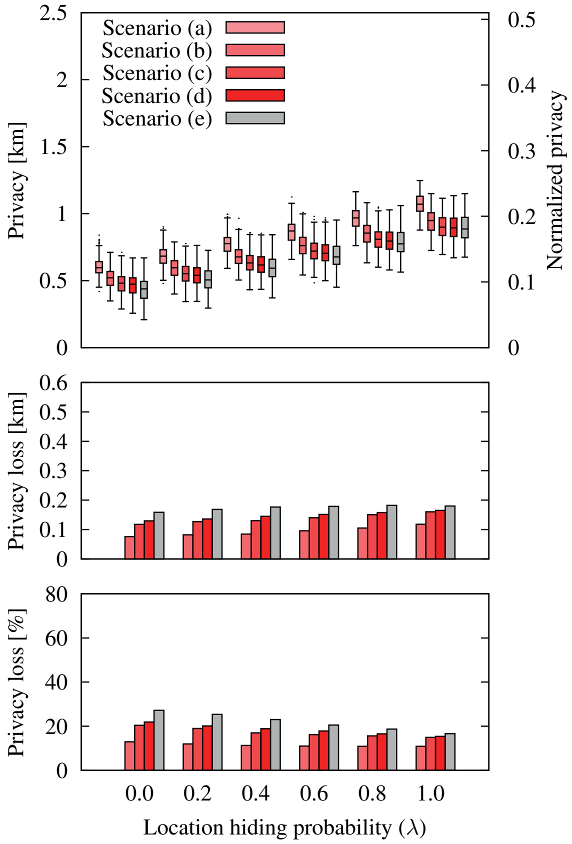


Figure 2.10: Comparison of the different localization attacks for the scenarios (a)-(e) depicted in Figure 2.9 on page 33, with obfuscation. The privacy loss (middle and bottom) is evaluated w.r.t. Scenario (a). In scenarios (b)-(e), we consider users report true co-locations with probability $\nu = 0.5$ and that they do not report fake co-locations ($\mu = 0$).

information between the target and her co-target is reported ($\nu = 1$), the users' privacy increases as there are more fake co-locations reported (as μ increases). However, when none of the true co-locations are reported ($\nu = 0$), we observe that the users' privacy *decreases* with the increase of available fake co-location information. In other words, an adversary can exploit the absence of a reported fake co-location at some time instant to infer that the users must, in fact, be co-located (for large values of μ).[9] This is an interesting observation that shows an adversary can learn not only from available co-location information but also from the *absence* of co-location information. Finally, in the case where only some of the true co-location information is reported ($\nu = 0.5$), we observe the largest users' privacy for values of μ which lead to a high uncertainty for the adversary (these are middle values of μ). We emphasize that users' privacy in the case where $\nu = \mu = 0$ (users *never* report co-location regardless of whether they are co-located or not) is the same as that where $\nu = \mu = 1$ (users *always* report co-location regardless of whether they are co-located or not). Finally, an important observation is that regardless of the amount of available co-location information (true or fake), users' privacy is never

[9] This is similar to the case of the entropy of a binary variable that is flipped with a given probability.

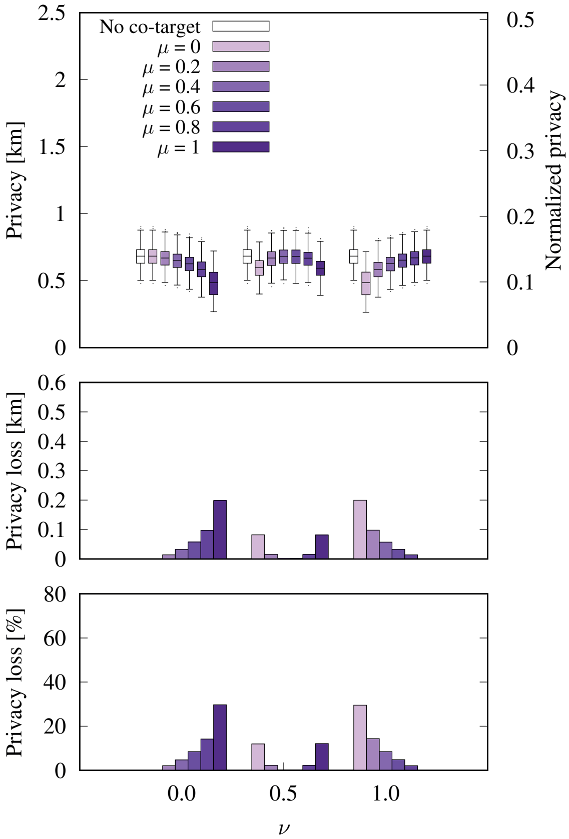


Figure 2.11: Privacy (top), absolute privacy loss (middle) and relative privacy loss (bottom) for the limited user set attack with $N = 2$ users, with obfuscation, $\lambda = 0.2$ and $\nu \in \{0, 0.5, 1\}$. The privacy *loss* is expressed w.r.t. the case where no co-locations are available; the histograms show medians. We observed similar results for other values of λ (not shown).

larger than that in the case where no co-locations are considered. This means that an adversary cannot be significantly confused by misleading co-location information, hence reporting such fake co-locations would not be an effective privacy protection practice.

Effects of User Coordination

We present the effect of using coordination for Scenario (e), where all available co-location is used. We infer the user location by using the BP algorithm for Scenario (e) and optimal inference for Scenario (a). We focus on the case where all users use obfuscation and have the same location hiding probability, λ. We assume users report true co-location information with probability $\nu = 0.5$ and no fake co-location information. We consider both the case where all users use coordination and the case where no users coordinate. We compare these with Scenario (a), where no co-location information is observed. Figure 2.12 on page 36 shows the results of our experiment: a box-plot representation of user privacy in the top graph, and the median privacy loss with respect to Scenario (a) in the bottom graphs. We observe that when users coordinate, their privacy can still decrease compared to the case where no co-locations are used, but there is a *privacy gain*

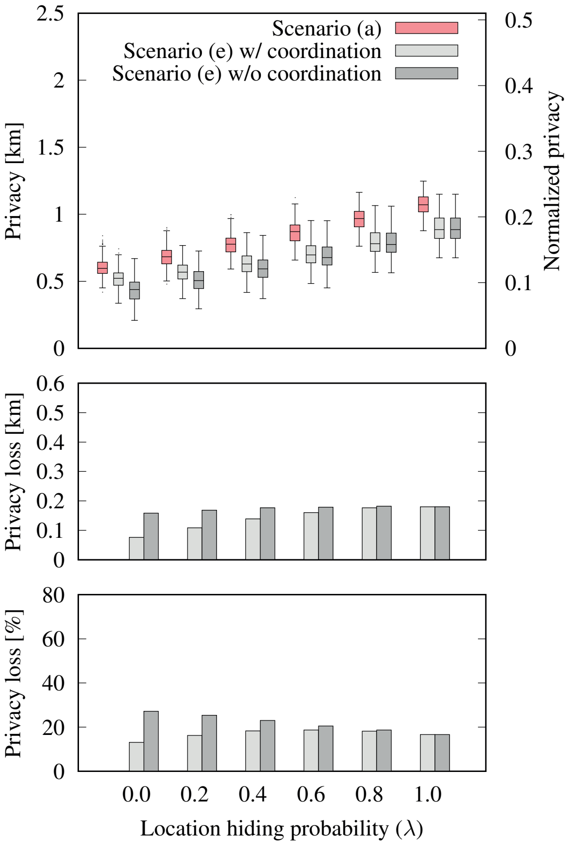


Figure 2.12: Localization attack with and without coordination for scenario (e) depicted in Figure 2.9 on page 33, with obfuscation, $\nu = 0.5$ and $\mu = 0$. The privacy loss (middle and bottom) is evaluated w.r.t. scenario (a).

with respect to the case where co-locations are reported in an uncoordinated fashion. This privacy gain is higher, as λ decreases. For instance, when $\lambda = 1$, users always hide their individual location and there is nothing to coordinate, hence coordination has no effect on users' location privacy. However, as users report more individual check-ins (λ decreases), the privacy gain stemming from coordination increases, with a peak for $\lambda = 0$ (where users' privacy loss is reduced by half when coordinating). We can conclude that, by coordinating their individual check-ins with their friends at times where users also report being co-located, users can limit the privacy loss caused by the co-location information.

Co-location Information on a Larger Scale

We evaluate the Bayesian network-based approximation on a set of 38 users. We compare it with the optimal individual localization attack (where no co-location information is used) and observe the same trend that co-location information further reduces location privacy.

In Section 2.7 and Section 2.5, we considered a small dataset of users, due to the high complexity of the optimal solution. We denote this small dataset by $\mathcal{U}_s$. Here, we

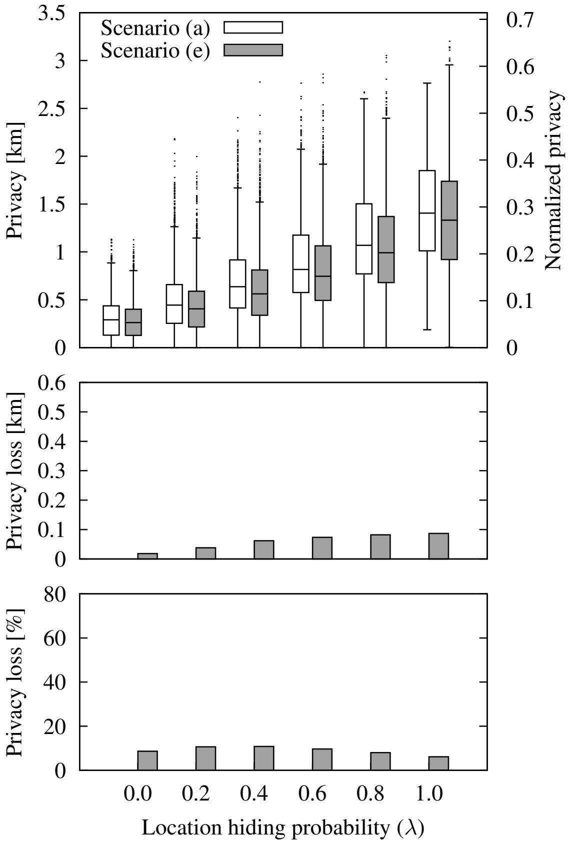


Figure 2.13: Comparison of the localization attacks for target users in $\mathcal{U}_l$ on Scenarios (a) and (e), as depicted in Figure 2.9 on page 33, with obfuscation. The privacy loss (middle and bottom) is evaluated w.r.t. Scenario (a). In Scenario (e), we consider users report true co-locations with probability $\nu = 0.5$ and that they do not report fake co-locations ($\mu = 0$).

evaluate our belief propagation solution on a larger dataset, in order to quantify location privacy loss when co-locations from a larger set of users are available. To this end, we select a subset $\mathcal{U}_l$ of users in the GeoLife dataset, such that each selected user must have at least one *real* co-location[10] with any other user in $\mathcal{U}_l$ (across their full traces). This results in 38 users being selected. Note that $\mathcal{U}_s \subset \mathcal{U}_l$. We emphasize that due to the low availability of real co-locations across the GeoLife users, this represents a weaker constraint of minimum desired co-locations, compared to that which we use when sampling the users in our small dataset $\mathcal{U}_s$. The low availability of co-locations, coupled with the sparsity of the location information available, also is also the basis for sampling 10 short *individual* collections of actual traces in the following way: For each u, a target user in $\mathcal{U}_l$, we generate actual traces for all the users in $\mathcal{U}_l$ such that (1) u has at least 10% of valid samples (*i.e.,* different from $r_\perp$) and u has at least 1 co-location with her co-target$_1$.

[10] Note that by real co-locations, we mean that the users are at the same location (i.e., their actual locations at a given time instant are the same), regardless of the fact that the co-location is reported or not.

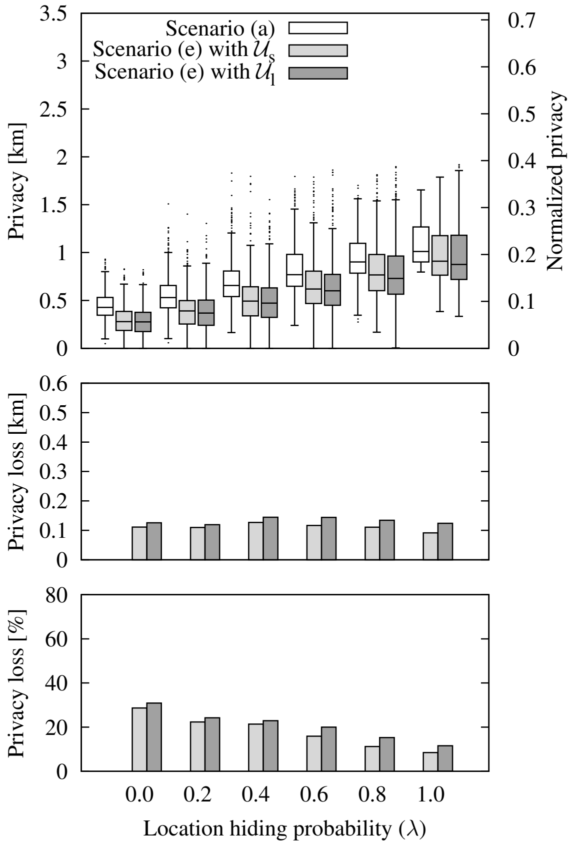


Figure 2.14: Comparison of the localization attacks for target users in $\mathcal{U}_s$ on Scenario (a), Scenario (e) considering co-locations only with and among users in $\mathcal{U}_s$ and Scenario (e) considering co-locations with and among all users in $\mathcal{U}_l$. The privacy loss (middle and bottom) is evaluated w.r.t. Scenario (a). We consider users report true co-locations with probability $\nu = 0.5$, do not report fake co-locations ($\mu = 0$) and use obfuscation.

We perform an individual localization attack by optimal inference for Scenario (a), considering, in turn, each user in the set $\mathcal{U}_l$ as the target user (using only their own reported locations and no co-locations). We then consider Scenario (e), the case of an adversary that exploits co-locations between any of the users in $\mathcal{U}_l$. We assume users report only a limited proportion of their true co-locations, with probability $\nu = 0.5$, and no fake co-locations ($\mu = 0$). We perform an approximate joint inference algorithm, by using the belief propagation algorithm with at most 20 iterations. We then compare the privacy in Scenario (e) to that in Scenario (a), in the case where all users use the same LPPM settings, *i.e.,* same value for λ and disclose only their obfuscated locations. Figure 2.13 shows the results of our comparison. It can be observed that the users' privacy decreases with the amount of considered co-locations. The privacy loss can seem somewhat modest, in comparison to the one observed in our previous experiments using $\mathcal{U}_s$. This can be explained by the fact that users in $\mathcal{U}_s$ have more real co-locations than those in $\mathcal{U}_l$ (a user has a median number of real co-locations in their actual traces of 5.5 and 2, respectively). We further compare the privacy of only the target users from $\mathcal{U}_s$ (but still using all the co-locations in the larger dataset $\mathcal{U}_l$) with that when using co-locations among users from $\mathcal{U}_s$. Figure 2.14 on page 38 shows the results of this comparison. It can be observed that the availability of co-locations with a larger number of users can further reduce privacy (privacy loss is as much as 31% when $\lambda = 0$).

2.8 Related Work

Location is identity. Even if the set of locations shared by a user is anonymized, and her true identity is hidden from the location-based service provider, the observed trajectories can be re-identified [6,49–51]. This attack is made by linking available information about users' mobility in the past with their observed traces. To protect against such attacks, many location obfuscation mechanisms have been proposed in the literature; they suggest that users hide their locations at certain locations, or that they reduce the accuracy or granularity of their reported locations [52–54]. These techniques increase users' privacy by making it more difficult for an adversary to de-anonymize users and to localize or track them over time. The location privacy of users in such settings can be computed using the expected error of an adversary in estimating their locations [28]. In such an inference framework, an adversary has some background knowledge on users' mobility models; this is used to reconstruct the full trajectories of the users, by using the anonymized and obfuscated observed traces.

The adversary's information, however, is not limited to mobility models. With most users being members of social networks, an adversary can de-anonymize location traces by matching the graph of co-traveler users with their social network graph [55]. Co-travelers are those who have been in each others' physical proximity for a considerable number of times. Researchers have extensively studied the problem of inferring social ties between users, based on their physical proximity [2, 56]. Recent revelations about NSA surveillance programs also show that this type of information is of great use for tracking and identifying individuals [57]. The dual problem, i.e., inferring location from social ties, has also been studied by the research community [58–60]. In [61], the authors exploit proximity information detected via Bluetooth, which is similar to co-location, to build an opportunistic ad-hoc localization algorithm by using intersection techniques similar to what we use in our attack. (see Figure 2.1 on page 8) In [62], the authors modeled the influence of social relationships on human movement and used this to predict future locations of the users. Location privacy risks have also been studied in the context of proximity detection (*e.g.,* finding nearby friends in OSNs) [63–66].

The correlation between different users' information also opens the door to a new type of privacy threat. Even if a user does not reveal much information about herself, her privacy can be compromised by others. In [67], the authors study how information revealed, from pictures, by a user's friends in social networks can be used to infer private information about her location. Private information about, for example, a user's profile and her age can also be inferred from shared information on online social networks [9,10]. A users' home address can also be inferred from those of her Facebook friends [68]. Mobile users, connecting to location-based services from a same IP address, can also compromise the privacy of those who want to keep their location private [69]. The loss in privacy, due to other users, has also been shown in other contexts such as genomics [12, 70]. Finally, interdependent privacy risks have been studied by using game-theoretic models for predicting the optimal behavior of rational users, in the context of OSNs [71, 72] and genomics [73]. Other game-theoretic interdependence models for security and privacy have been surveyed in [10].

Extracting co-location information about users (i.e., who is with whom) is becoming increasingly easier. More specifically, with the proliferation of mobile social networks, where users can check-in with others at various locations, the threat of available co-location information on users' location privacy is clear (as pointed out in [27]). Despite the abovementioned works on quantifying the location privacy and the privacy of users in social networks, as well as the extensive research on privacy loss due to others, there has not been a study on evaluating location privacy where co-location information is considered. We bridge the gap between studies on location privacy and social networks, and we propose the first analytical framework for quantifying the effects of co-location information on location privacy, where users can also make use of obfuscation mechanisms.

2.9 Conclusion

In this chapter, we have studied the effect on users' location privacy when co-location information is available, in addition to individual (obfuscated) location information. To the best of our knowledge, this is the first effort to quantify the effects of co-location information that stems from social relationships between users on location privacy; as such, it constitutes a first step towards bridging the gap between studies on location privacy and social networks. Indeed, most studies on geo-location and social networks look at how social ties can be inferred from co-locations between individuals and how social ties can be used to de-anonymize mobility traces. We have shown that, by considering the users' locations jointly, an adversary can exploit co-location information to better localize users, hence decrease their individual privacy. Although the optimal joint localization attack has a prohibitively high computational complexity, the polynomial-time approximate inference algorithms that we propose provide good localization performance. An important observation from our work is that a user's location privacy is no longer entirely in her control, as the co-locations and the individual location information disclosed by other users significantly affect her own location privacy.

The message of this work is that protection mechanisms must not ignore the social aspects of location information. Because it is not desirable to report dummy lists of co-located users (as this information is displayed on the users' profiles on social networks), a location-privacy preserving mechanism needs, instead, to generalize information about co-located users or to generalize the time of a social gathering, as well as the locations of users at other locations, in order to reduce the effectiveness of the attacks we suggested in this chapter. For a first attempt to mitigate the privacy risks stemming from co-location information, we have proposed a simple countermeasure that relies on cooperation between users and have demonstrated its effectiveness. We intend to address the design of social-aware location-privacy protection mechanisms (running on the users' mobile devices) to help the users assess and protect their location privacy when co-location information is available. An important aspect of generalization techniques is the tension between utility and privacy: For a user, reporting to be with "some friends" might not be sufficiently informative, and the generalized co-location information would fail to serve the user's purpose. Usability is also a crucial aspect for the adoption of technical protection mechanisms. We plan to investigate both the utility and usability aspects of such protection mechanisms through targeted user surveys.

In our future work, we also plan to investigate the case where co-locations are not explicitly reported by the users, instead the adversary has access to the social ties between the users (e.g., friends, family, colleagues). Such ties can be associated with probabilistic co-location patterns; for instance, the fact that the locations of work-colleagues are often correlated during office hours.

Chapter 3

To Share or not to Share: Insights into Users' Behavior of Sharing (Co-)Locations

> *Interdependence is a fundamental law of nature. Even tiny insects survive by cooperating with each other. Our own survival is so dependent on the help of others that a need for love lies at the very core of our existence. This is why we need to cultivate a genuine sense of responsibility and a sincere concern for the welfare of others.*
>
> Dalai Lama

3.1 Introduction

The location-sharing feature offered by major online social networks has gained momentum, as users increasingly access their favorite OSNs from their smartphones – as much as 95.1% of active users worldwide accessed their Facebook account via smartphone in January 2018 [74]. Thus 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. [37] 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 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. [75]) 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.

But, these features also raise privacy concerns. Although it has been known for years that location information leads to severe privacy issues and this has been extensively studied in the literature (e.g., [2,4,76]; see also FindYou [77], 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 [22], as we presented in Chapter 2. A critical aspect of co-locations is that they relate to all the involved users (such information is "co-owned" by the involved users [78,79]) 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 evaluate 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 [79,80].

Awareness about the interdependent nature of privacy is increasing, yet, due to its complexity, this is not explicitly addressed by current laws. Opinion 5/2009 on online social networking produced by the Working Party on Data Protection, 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 others. Yet, even in the General Data Protection Regulation (GDPR) (Regulation EU 2016/679) which became enforceable on 25 May 2018, 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 few regulations that apply to sharing on OSNs (except for the extreme case of sharing sexually explicit content, namely revenge pornography) and this serious problem deserves further study. We dedicate the next chapter, Chapter 4, for exploring solutions for this problem.

In this chapter, 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 anyone to analyze the behavior of users regarding location and co-location sharing on OSNs. To this end, we build our framework by using two well-established modeling and analytical tools: game theory [81–83] and conjoint analysis [84]. Game theory enables us to model and formalize the users' sharing rationale and behavior. Such models include a number of parameters that, typically in the expression of the users' utility, 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 [71,72,85] 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 propose the first game-theoretic framework, namely the Sharing Game, to formalize the important problem of location sharing with interdependent privacy risks (introduced by co-location). 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 and on a real dataset, Geolife [48]. 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 chapteris organized as follows. In Section 3.2, we give background information on the main techniques used. In Section 3.3, 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 3.4, 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 3.5, we evaluate our framework in a number of scenarios. We present an extended model in Section 3.6. In Section 3.7, we discuss directions for improvement and extension of our work. In Section 3.8, we survey the related work. In Section 3.9, we conclude the chapter and we discuss future work. Finally, in Section 3.9, we provide the full survey transcript.

3.2 Background

In this section, we briefy introduce the relevant concepts in game theory and the conjoint analysis technique.

Game Theory 101

Game theory is the study of the strategic interaction between multiple rational decision-makers who aim to maximize their own utility [81–83]. 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 computing 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. It is especially relevant for our work as it enables us to model and analyze users' preferences and interactions, and to understand 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. A NE is not necessarily a social optimum, but finding a socially-optimal NE is highly desirable from a mechanism design perspective. This process continues to the second to last actions, and so on until it reaches the first move of the game, i.e., the root of the game tree.

Conjoint Analysis 101

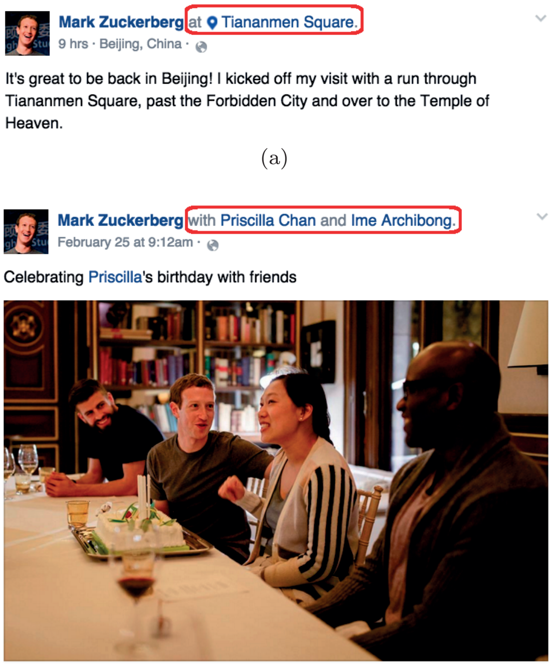
Conjoint analysis [84] is an experimental approach used to detect the hidden rules users rely on to make decisions (involving trade-offs) between services. A service is viewed as a combination of attributes, each of which has different levels (values). Users are asked to rank multiple versions of the 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* [86]. 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; these are represented as percentages for all the attributes.

3.3 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 a 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 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 3.1 on page 47 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 (once the service provider has seen the information, it cannot be "unseen"). 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 the two users – has access to some of the 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 (co-)location information, and the decision to allow a friend to post co-location information, is made strategically by both the users involved.

We are aware of the fact that users might act irrationally, especially when it comes to privacy-related decisions [87] and that privacy concerns were shown to vary decidedly with context as well as personal traits [19]. However, users' rationality in privacy-related decision making is an active research topic. For instance, recent results by Redmiles et al. show that users can actually make rational decisions in the context of adopting optional security behavior [88] and that users are more likely to behave rationally in the face of high risk [89]. We believe that privacy-protection demand will increase, notably because a



Mark Zuckerberg at Tiananmen Square.
9 hrs · Beijing, China ·

It's great to be back in Beijing! I kicked off my visit with a run through Tiananmen Square, past the Forbidden City and over to the Temple of Heaven.

(a)

Mark Zuckerberg with Priscilla Chan and Ime Archibong.
February 25 at 9:12am ·

Celebrating Priscilla's birthday with friends

(b)

Figure 3.1: 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.

growing number of people suffer the consequences of their carelessness and that of others. Furthermore, smartphones are increasingly involved in the sharing decisions users make, as demonstrated by the growing sophistication of the apps' permission systems. A tool run for this purpose *can* be "rational" and strictly follow the parametrization provided by its user to aid her in decision making. It is therefore of interest to investigate what happens under the assumption of *rationality*.

3.3.1 User Model

We model the interactions between a user and one of her friends (also called players) as a game, called the *Sharing Game*, over a time window of interest ($\{1, \ldots, T\}$). Note that the adversary, with respect to whom the users' privacy is evaluated (typically the service provider), is *not* a player of the game.

We denote by $\mathbf{a}(t) = (a_i(t), a_j(t))$ the users' (denoted by i and j) actual locations at time t. A potential strategy of a user i at time t is denoted by $s_i(t)$ and $\mathbf{s}(t) \triangleq (s_i(t), s_j(t))$ denotes a strategy profile. $s_i(t)$ is chosen from the combinations of possibilities 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

$a_i(t)$	Player i's actual location at time t
$a_j(t)$	Player j's actual location at time t
$\mathbf{a}(t) = (a_i(t), a_j(t))$	The players' actual locations at time t
$s_i(t) = (sl_i(t), sc_i(t))$	A possible strategy of player 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
$\mathbf{s}^*(t) = \left(s_i^*(t), s_j^*(t)\right)$	Equilibrium strategy profiles (decisions) at time t
α_i	Weight with which player i values privacy over benefits
$B_i\left(t, \mathbf{a}(\mathbf{t}), \mathbf{s}(t)\right)$	Player i's benefits at time t for strategy profile $\mathbf{s}(t)$
b^i_{sl}	Player i's benefit of sharing her actual location at t
b^i_{vl}	Player i's benefit of viewing her friend's location at t
b^i_{sc}	Player i's benefit of sharing co-location with a friend at t
b^i_{vc}	Player i's benefit of viewing co-loc. shared by a friend at t
f^i_{sv}	Player i's preference factor: sharing vs. viewing
f^i_{lc}	Player i's preference factor: location vs. co-location
f^i_{pb}	Player i's preference factor: privacy vs. benefits
f^i_a	The altruistic factor of player i for the other player
$\mathbf{o}(t)$	Information observed by the adversary in the time window up to t
$i \leftrightarrow_t j$	Co-location between i and j at time t, observed by the adversary
$a_i(t)\,@_t$	Player i's actual location at time t, observed by the adversary
$a_j(t)\,@_t$	Player j's actual location at time t, observed by the adversary
$\mathcal{B}_i$	The adversary's background knowledge about player i (*e.g.,* her mobility profile)
$P\left(i, t, \mathbf{a}(t), \mathbf{o}(t-1), \mathbf{s}(t), \mathcal{B}_i, \mathcal{B}_j\right)$	Player i's privacy at time t for some strategy profile $\mathbf{s}(t)$
$\theta_i(t)$	Player i's type at time t (includes actual location, benefits vs. privacy preferences, $\cdots$)
$\hat{u}_i\left(t, \mathbf{a}(t), \mathbf{o}(t-1), \mathbf{s}(t), \mathcal{B}_i, \mathcal{B}_j, \boldsymbol{\theta}(t)\right)$	Player i's *individual utility* at time t for strategy profile $\mathbf{s}(t)$ and player types $\boldsymbol{\theta}(t)$
$\bar{u}_i\left(t, \mathbf{a}(t), \mathbf{o}(t-1), \mathbf{s}(t), \mathcal{B}_i, \mathcal{B}_j\right)$	Player i's *expected individual utility* at t for strategy profile $\mathbf{s}(t)$
$u_i\left(t, \mathbf{a}(t), \mathbf{o}(t-1), \mathbf{s}(t), \mathcal{B}_i, \mathcal{B}_j\right)$	Player i's *perceived utility* at time t for strategy profile $\mathbf{s}(t)$ (includes altruism)
$U_i\left(t, \mathbf{a}(t), \mathbf{o}(t-1), \mathbf{s}(t), \mathcal{B}_i, \mathcal{B}_j\right)$	Player i's *utility/cumulative utility* (includes future considerations) at time t for strategy profile $\mathbf{s}(t)$
δ	Discount factor for future considerations in the cumulated utility
$SW(t, s_i(t), s_j(t))$	Social welfare at time t for strategy profile $(s_i(t), s_j(t))$

Table 3.1: Table of notations.

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, selecting between two possible strategies: $\bar{L}\bar{C}$ – sharing nothing and $L\bar{C}$ – sharing location information. At each time in the windows of interest, both users choose their equilibria strategies–denoted by $\mathbf{s}^*(t) \triangleq (s_i^*(t), s_j^*(t))$. Information that the users share becomes available to an adversary: their actual locations at times at which they choose L and/or the fact that they are co-located at times they choose C. For a time t, we denote by $\mathbf{o}(t-1)$ the information that the adversary observes up to time $t-1$. We consider that the adversary only observes information at $k \in \{0, \ldots, t-1\}$ time instants up to $t-1$ included. The information in this set depends on both players' equilibria decisions up to time $t-1$. Note that for $k = 0$ or $t = 1$ the set is empty. The information that the adversary observes at time t depends on $\mathbf{s}(t)$.

User i's social benefits that correspond to a strategy profile $\mathbf{s}(t)$ at time t are denoted by $B_i\left(t, \mathbf{a}(t), \mathbf{s}(t)\right)$. Note that the benefit function takes into account (i) the time t, to reflect the fact that check-ins at different times can have different meanings, (ii) both users' locations at time t, to reflect the fact that some locations can be more interesting to share or view than others (*e.g.,* a hotel versus a park), (iii) who the other user is, to reflect the fact that some co-locations can be more interesting to share than others, and (iv) both of the users' strategies, to reflect the benefit of sharing and that of viewing information shared by her friend. Her privacy at t, denoted by $P\left(i, t, \mathbf{a}(t), \mathbf{o}(t-1), \mathbf{s}(t), \mathcal{B}_i, \mathcal{B}_j\right)$, is a function of (i) both users' actual location at t, (ii) the information observed by the adversary at the last k time instants up to $t-1$ – this depends on both users' strategies at those time instants, (iii) their strategy profile at t, and (iv) background user information (denoted by $\mathcal{B}_i, \mathcal{B}_j$), *e.g.,* their mobility profiles. We emphasize that the privacy function takes into account previous time instants. In other words, a decision to disclose information at time t has privacy implications at later time instants. Due to the dependency introduced by co-locations, the privacy function also takes into account decisions made by the other user, as well as the related background information.

Naturally, the information that a user has about the adversary's background knowledge of herself and of the other user, and her information about the other user's past or current locations, social benefits and privacy preferences, could be limited. These factors would influence her computation of her own privacy and of the other's privacy and social benefits. Some of these factors could be estimated (*e.g.,* by completing surveys to compute their preference factors) and voluntarily shared among players (for instance through the service provider, in a private way).

In the decision-making process, *players can be assisted by a tool for evaluating the privacy implications*, namely the value $P(\cdot)$, of each of the players' possible decisions regarding sharing. For instance, a Facebook client could compute this and suggest players' optimal decisions, using the information regarding users' locations and preferences in the computation of the game's equilibria. Another option is that such information about the friend is unknown and the players (*i.e.,* their local tool) must estimate it or build probabilistic models of it. For the sake of keeping our model easy to understand, we consider the first option here; we also consider only immediate privacy implications in the users' estimation of the privacy (users were shown to often become "privacy myopic" and opt for *immediate gratification* in the context of their privacy decisions [90]); last, we assume players to be selfish. We present an extended model, relaxing all of these assumptions, in Section 3.6.

At any time instant 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,

$$B_i\left(t, \mathbf{a}(t), \mathbf{s}(t)\right) \quad = \quad \frac{b_{sl}^i(\cdot) sl_i(t) + b_{sc}^i(\cdot) sc_i(t) + b_{vl}^i(\cdot) sl_j(t) + b_{vc}^i(\cdot) sc_j(t)}{b_{sl}^i(\cdot) + b_{sc}^i(\cdot) + b_{vl}^i(\cdot) + b_{vc}^i(\cdot)} \tag{3.1}$$

where $b_{sl}^i(\cdot)$ and $b_{sc}^i(\cdot)$ denote user i's benefit of sharing location and co-location, and $b_{vl}^i(\cdot)$ and $b_{vc}^i(\cdot)$ her benefit of viewing location and co-location. Note that these benefits also take into account the parameters of $B_i(\cdot)$ and it is possible that they are correlated, *e.g.,* if user i has a large value of $b_{sl}^i(.)$, she might also have a large value of $b_{sc}^i(.)$. The utility of player i for some strategy profile $\mathbf{s}(t)$ at time t captures both her social benefits and her privacy. Specifically,

$$\begin{aligned} U_i\left(t, \mathbf{a}(t), \mathbf{o}(t-1), \mathbf{s}(t), \mathcal{B}_i, \mathcal{B}_j\right) \quad &= \\ (1-\alpha_i) \cdot B_i\left(t, \mathbf{a}(t), \mathbf{s}(t)\right) + \alpha_i \cdot P\left(i, t, \mathbf{a}(t), \mathbf{o}(t-1), \mathbf{s}(t), \mathcal{B}_i, \mathcal{B}_j\right) & \end{aligned} \tag{3.2}$$

where $\alpha_i \in [0,1]$ denotes the weight with which user i values her privacy over her social benefits. This formulation follows a privacy calculus approach [91,92] under a pragmatic

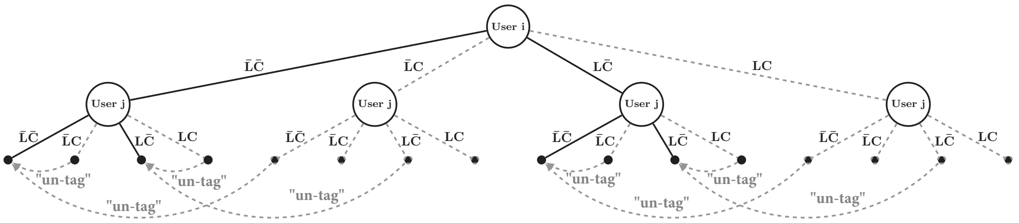


Figure 3.2: Possible strategies for one time instant of the Sharing Game in the friends adversarial models (depicted in Figure 3.3 on page 51). User i is the first player (she chooses a strategy first) and user j the second player (she reacts to i's choice). Only the black solid strategies are valid 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 to share their co-location are valid.

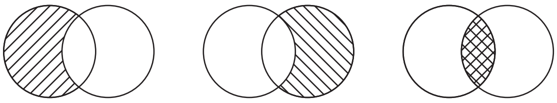
user model, as classified by Westin [93]. The choice of a linear model follows previous work (*e.g.,* Acquisti [90]).

The game is played successively, at 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 sequential: The second player (or her application implementing the decision model) knows the choice of the first player and decides (or suggests to 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 (reacting to each other's sharing actions), hence our choice of the model.

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 initiated by either of the players are visible to the service provider and cannot be removed (even if the second player removes them, the service provider still has access to 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 fixed. Once they strategically choose the best decisions at time t, they will not revisit them at later time instants. Table 3.1 on page 48 summarizes the notations used in our formalism.

3.3.2 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



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

Figure 3.3: 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. The intersection represents the common friends of i and j.

to them for the typical default privacy settings for OSN posts. Tables 3.2, 3.3, 3.4, and 3.5 summarize the information that is available to each of the adversaries, for all strategy profiles.

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 her. 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 their IP address.[1]

$sl_i(t)$	$sc_i(t)$	$sl_j(t)$	$sc_j(t)$	Adversary's info. on i = Adversary's info. on j
0	0	0	0	$\varnothing$
0	0	0	1	$\{i \leftrightarrow_t j\}$
0	0	1	0	$\{a_j(t) \, @_t\}$
0	0	1	1	$\{i \leftrightarrow_t j, a_j(t) \, @_t\}$
0	1	0	0	$\{i \leftrightarrow_t j\}$
0	1	0	1	$\{i \leftrightarrow_t j\}$
0	1	1	0	$\{i \leftrightarrow_t j, a_j(t) \, @_t\}$
0	1	1	1	$\{i \leftrightarrow_t j, a_j(t) \, @_t\}$
1	0	0	0	$\{a_i(t) \, @_t\}$
1	0	0	1	$\{i \leftrightarrow_t j, a_i(t) \, @_t\}$
1	0	1	0	$\{a_i(t) \, @_t, a_j(t) \, @_t\}$
1	0	1	1	$\{i \leftrightarrow_t j, a_i(t) \, @_t, a_j(t) \, @_t\}$
1	1	0	0	$\{i \leftrightarrow_t j, a_i(t) \, @_t\}$
1	1	0	1	$\{i \leftrightarrow_t j, a_i(t) \, @_t\}$
1	1	1	0	$\{i \leftrightarrow_t j, a_i(t) \, @_t, a_j(t) \, @_t\}$
1	1	1	1	$\{i \leftrightarrow_t j, a_i(t) \, @_t, a_j(t) \, @_t\}$

Table 3.2: Information available to the adversary in the SP model, at time t

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, the co-location

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

$sl_i(t)$	$sc_i(t)$	$sl_j(t)$	$sc_j(t)$	Adversary's info. on i	Adversary's info. on j
0	0	0	0	$\varnothing$	$\varnothing$
0	0	1	0	$\varnothing$	$\{a_j(t)\,@_t\}$
0	1	0	1	$\{i \leftrightarrow_t j\}$	$\{i \leftrightarrow_t j\}$
0	1	1	1	$\{i \leftrightarrow_t j\}$	$\{i \leftrightarrow_t j, a_j(t)\,@_t\}$
1	0	0	0	$\{a_i(t)\,@_t\}$	$\varnothing$
1	0	1	0	$\{a_i(t)\,@_t\}$	$\{a_j(t)\,@_t\}$
1	1	0	1	$\{i \leftrightarrow_t j, a_i(t)\,@_t\}$	$\{i \leftrightarrow_t j\}$
1	1	1	1	$\{i \leftrightarrow_t j, a_i(t)\,@_t\}$	$\{i \leftrightarrow_t j, a_j(t)\,@_t\}$

Table 3.3: Information available to the adversary in the MF model, at time t

$sl_i(t)$	$sc_i(t)$	$sl_j(t)$	$sc_j(t)$	Adversary's info. on i	Adversary's info. on j
0	0	0	0	$\varnothing$	$\varnothing$
0	0	1	0	$\{a_j(t)\,@_t\}$	$\varnothing$
0	1	0	1	$\{i \leftrightarrow_t j\}$	$\{i \leftrightarrow_t j\}$
0	1	1	1	$\{i \leftrightarrow_t j, a_j(t)\,@_t\}$	$\{i \leftrightarrow_t j\}$
1	0	0	0	$\varnothing$	$\{a_i(t)\,@_t\}$
1	0	1	0	$\{a_j(t)\,@_t\}$	$\{a_i(t)\,@_t\}$
1	1	0	1	$\{i \leftrightarrow_t j\}$	$\{i \leftrightarrow_t j, a_i(t)\,@_t\}$
1	1	1	1	$\{i \leftrightarrow_t j, a_j(t)\,@_t\}$	$\{i \leftrightarrow_t j, a_i(t)\,@_t\}$

Table 3.4: Information available to the adversary in the FF model, at time t

$sl_i(t)$	$sc_i(t)$	$sl_j(t)$	$sc_j(t)$	Adversary's info. on i = Adversary's info. on j
0	0	0	0	$\varnothing$
0	0	1	0	$\{a_j(t)\,@_t\}$
0	1	0	1	$\{i \leftrightarrow_t j\}$
0	1	1	1	$\{i \leftrightarrow_t j, a_j(t)\,@_t\}$
1	0	0	0	$\{a_i(t)\,@_t\}$
1	0	1	0	$\{a_i(t)\,@_t, a_j(t)\,@_t\}$
1	1	0	1	$\{i \leftrightarrow_t j, a_i(t)\,@_t\}$
1	1	1	1	$\{i \leftrightarrow_t j, a_i(t)\,@_t, a_j(t)\,@_t\}$

Table 3.5: Information available to the adversary in the CF model, at time t

information potentially shared by the first player can be removed by the second one (*e.g.,* by un-tagging). Figure 3.2 on page 50 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 3.3: (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 α.) for a public adversary – though this is not the default visibility for posts, users can post information with public visibility.

We emphasize that the action of un-tagging is not a strategy in our game. Yet, it is modelled in different ways: In the FF, CF and MF models, un-tagging is equivalent to the strategies that do not share co-location ($\bar{C}$) – these adversaries can no longer see the co-location; in the SP model, un-tagging has no effect – the SP has access to all the information shared by either player.

3.3.3 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. Observations made by the adversary at prior time instants, stemming from the players' equilibria decisions, 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_i, essentially

$$\forall s_i,\ s_j^*(s_i) = \arg\max_s\ U_j\left(t, \mathbf{a}(t), \mathbf{o}(t-1), (s_i, s), \mathcal{B}_i, \mathcal{B}_j\right) \tag{3.3}$$

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

$$s_i^* = \arg\max_s\ U_i\left(t, \mathbf{a}(t), \mathbf{o}(t-1), \left(s, s_j^*(s)\right), \mathcal{B}_i, \mathcal{B}_j\right) \tag{3.4}$$

The equilibrium decisions at time t are given by

$$\mathbf{s}^*(t) = \left(s_i^*, s_j^*(s_i^*)\right) \tag{3.5}$$

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

$$\begin{aligned} SW(t, \mathbf{s}^*(t)) &= U_i\left(t, \mathbf{a}(t), \mathbf{o}(t-1), \mathbf{s}^*(t), \mathcal{B}_i, \mathcal{B}_j\right) + \\ &\quad U_j\left(t, \mathbf{a}(t), \mathbf{o}(t-1), \mathbf{s}^*(t), \mathcal{B}_i, \mathcal{B}_j\right) \end{aligned} \tag{3.6}$$

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

3.3.4 Equilibria Properties

We are interested in different properties for the players' equilibria decisions: social optimality and utility maximization.

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:

$$\forall \mathbf{s}(t) \neq \mathbf{s}^*(t): \qquad SW\left(t, \mathbf{s}^*(t)\right) \geq SW\left(t, \mathbf{s}(t)\right) \tag{3.7}$$

Individual Utility Maximization at Equilibrium

A player's i utility is maximized for the equilibrium decisions at time t if the following property holds:

$$\forall \mathbf{s}(t) \neq \mathbf{s}^*(t) : U_i\left(t, \mathbf{a}(t), \mathbf{o}(t-1), \mathbf{s}^*(t), \mathcal{B}_i, \mathcal{B}_j\right) \geq U_i\left(t, \mathbf{a}(t), \mathbf{o}(t-1), \mathbf{s}(t), \mathcal{B}_i, \mathcal{B}_j\right) \quad (3.8)$$

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. Note that social optimality is defined only at time instants where both players play the game.

3.4 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 for some 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.

3.4.1 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 use the information about the participants' Facebook account only for screening purposes and we did not store any such information (we made this point clear in the advertisement page 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 [84] and making use of a dedicated tool, namely the XLSTAT statistical software [94].

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 the posts to which we refer are posts they once shared and, hence, 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 [22]. 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.

It took approximately ten minutes to complete the survey; the participants were paid 2 USD. We ruled out the 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 relatively 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.

Estimation of the Model's Parameters. We estimate the parameters in our model (α, $b_{sl}(t)$, $b_{sc}(t)$, $b_{vl}(t)$ and $b_{vc}(t)$) from the survey data. As we wanted to keep the number of questions for our participants low, we quantified only three preference factors f_{pb}, f_{lc} and f_{sv}; to estimate the model's parameters from these, we make a few 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, that (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, and that (3) the users' benefits of sharing/viewing are the same over time. We derive the values of the model parameters as follows:

$$\begin{aligned} \alpha &= f_{pb} \\ b_{sl} &= \frac{f_{sv}}{1-f_{sv}} \cdot \frac{f_{lc}}{1-f_{lc}} b_{vc} \\ b_{sc} &= \frac{f_{sv}}{1-f_{sv}} b_{vc} \\ b_{vl} &= \frac{f_{lc}}{1-f_{lc}} b_{vc} \end{aligned} \tag{3.9}$$

where b_{vc} is considered a free variable (we set it to 1).

3.4.2 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. Table 3.6 on page 57 presents relevant statistics (*e.g.,* mean and standard deviation) and Figure 3.4 on page 57 illustrates the CDFs of the derived preference factors. Note that these results should be taken with a grain of salt as previous works (e.g., [95, 96]) 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, $.57 \pm .15$, $.56 \pm .15$ and $.60 \pm .39$ 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 [97, 98]. 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 ($f_{pb} > 0.5$).

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. 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

	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 3.6: 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}.

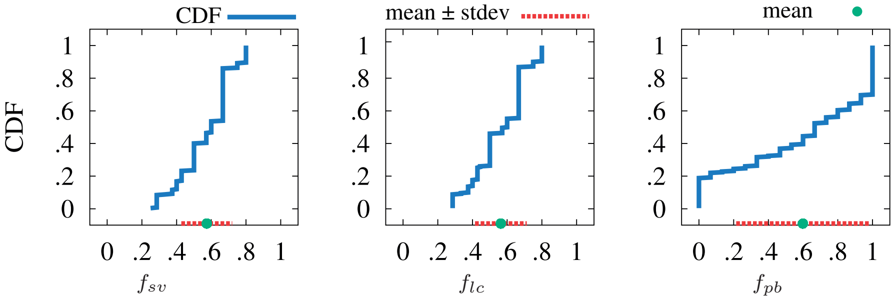


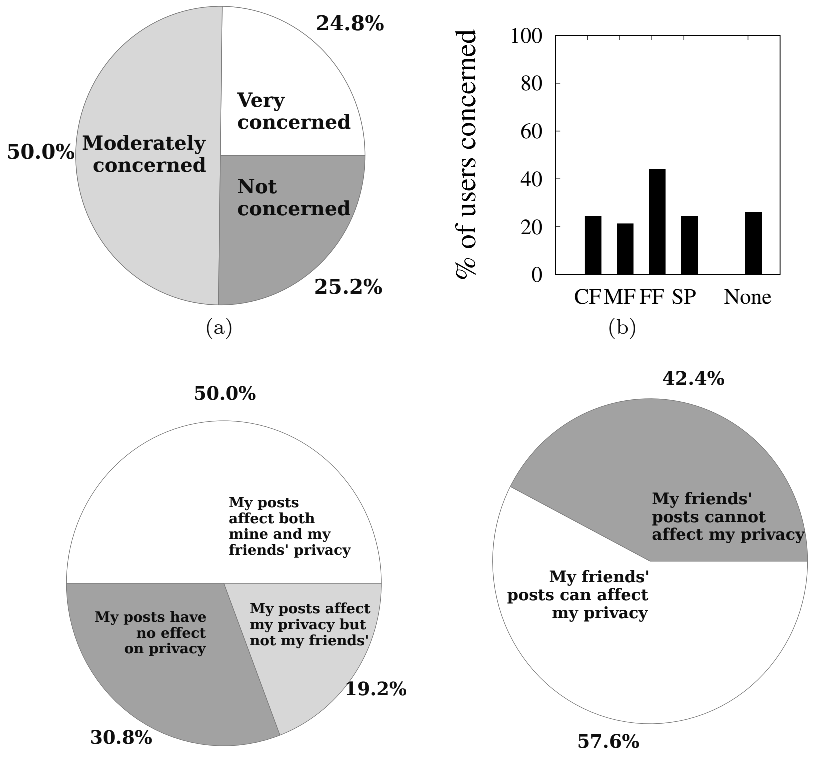
Figure 3.4: CDFs of the preference factors of our survey participants.

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%); 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 3.5 on page 58). 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, for instance, by 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.

The full transcript of our survey, as well as an anonymized and sanitized version of the answers (part II and some of part III, password *FbS250*) are available at https://infoscience.epfl.ch/record/218755?&ln=en.



(a) (b) (c) (d)

Figure 3.5: (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). Users' awareness about (c) privacy risks stemming from their own posts; (d) own privacy risks stemming from friends' posts.

3.5 Evaluation

We evaluate our framework by simulating and analyzing the users' decisions in different experimental setups. Note that the space of the parameters to explore is very large, therefore we isolate some of the parameters and we present a selected set of experiments which shed light on key insights.

3.5.1 Experimental Setup

In this section, we describe the experimental setup of the different building blocks of our framework, including the location privacy function and the evaluation scenarios.

Quantification of the Users' Privacy

We quantify users' privacy ($P(\cdot)$) as their location privacy, by relying on the inference framework proposed in Chapter 2 (note that our model is flexible enough to enable the use of other frameworks for inferring location privacy, for instance, the one proposed by Xu et al. [99].); 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

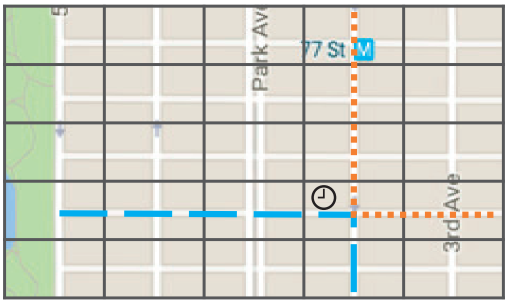


Figure 3.6: Canonical meeting scenario considered in the evaluation: Two users, Alice (dotted) and Bob (dashed), coming from distinct directions, meet for some time, and later separate in distinct directions.

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 [33] 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 3.3.2, 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 time instances where a user's actual location is not known (sparse data), her privacy cannot be evaluated. At each other time instant t, the adversary considers *all* past location and co-location posts from the users when inferring their locations.

Scenarios

In order to evaluate our framework and to gain insight about the effects of the different parameters, we first consider the **canonical meeting scenario**, illustrated in Figure 3.6: 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. Both users estimate the mobility profiles that an adversary would use in the inference process (*i.e.*, $\mathcal{B}_i, \mathcal{B}_j$) by a very basic one: 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);[2] note that, other than from a velocity point of view (a user cannot move further than one cell in one time instant), this captures no real mobility information (all locations are equally probable). 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. This scenario is, thus, also representative for the case of users who naively estimate the adversarial background information.

[2] We assume, for simplicity, a Manhattan-like model where users can move vertically or horizontally.

Additionally, we consider a **real dataset scenario**, using the Geolife dataset [48] (collected in 2008) and the same co-location generation, user mobility profiles construction and space and time discretization as presented in Chapter 2 (25 geographic regions covering the campus of Tsinghua University in Beijing and one-hour time sub-intervals splits of the continuous time interval). In this scenario, two users, Alice and Bob, follow their individual actual location traces (we consider sub-samples of $T = 300$ time instants from their full traces). At each time instant, both Alice and Bob can either report or hide their actual location (if this is known). Additionally, if co-located, they can also choose to report their co-location. In the privacy computation, mobility profiles constructed from real users' *full* location traces are used. Note that these are different and no longer uniform (among locations) and that they illustrate user-specific patterns of movement.

3.5.2 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.

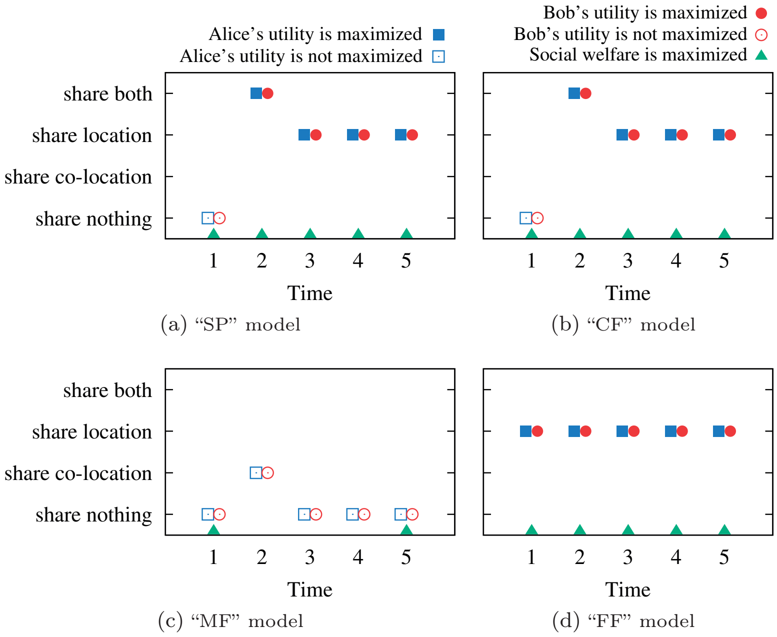
Canonical Meeting Scenario

We first present several simulations on the canonical meeting scenario.

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, based on the canonical meeting scenario, we consider a *homogeneous* setup, 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 3.4 on page 57. Figure 3.7 on page 61 illustrates the different game outcomes, for the four adversarial models we presented in Section 3.3.2. 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 3.7a and 3.7b on page 61), 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 for 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 their incurred privacy loss, weighted by $1 - \alpha$ and α, 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



(a) "SP" model

(b) "CF" model

(c) "MF" model

(d) "FF" model

Figure 3.7: Players' decisions at equilibrium, $(s^*_{Alice}(t), s^*_{Bob}(t))$, 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.

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 benefits) and this effect propagates at successive time instants.

In the *MF* model (Figure 3.7c), 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 [22]), 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 3.7d on page 61), 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 $(L\bar{C}, L\bar{C})$ and it maximizes both the social welfare and the players' utilities.

The Effect of Privacy vs. Benefits Preferences.

We present a *heterogeneous* setup, based on the canonical meeting 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 $[0, 1]$. Figures 3.8 on page 63 and 3.9 on page 64 illustrate our results (see caption for details). 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 (recall that "share co-location" and "share both" decisions can only occur when the players are co-located, *i.e.,* 20% of the times) 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).

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 $(\bar{L}\bar{C}, \bar{L}\bar{C})$, $(\bar{L}C, \bar{L}C)$ 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* also but 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 permitted. An example can be observed in Figure 3.8 on page 63, 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 $(\bar{L}\bar{C}, L\bar{C})$: 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 α 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 α_{Bob} decreases

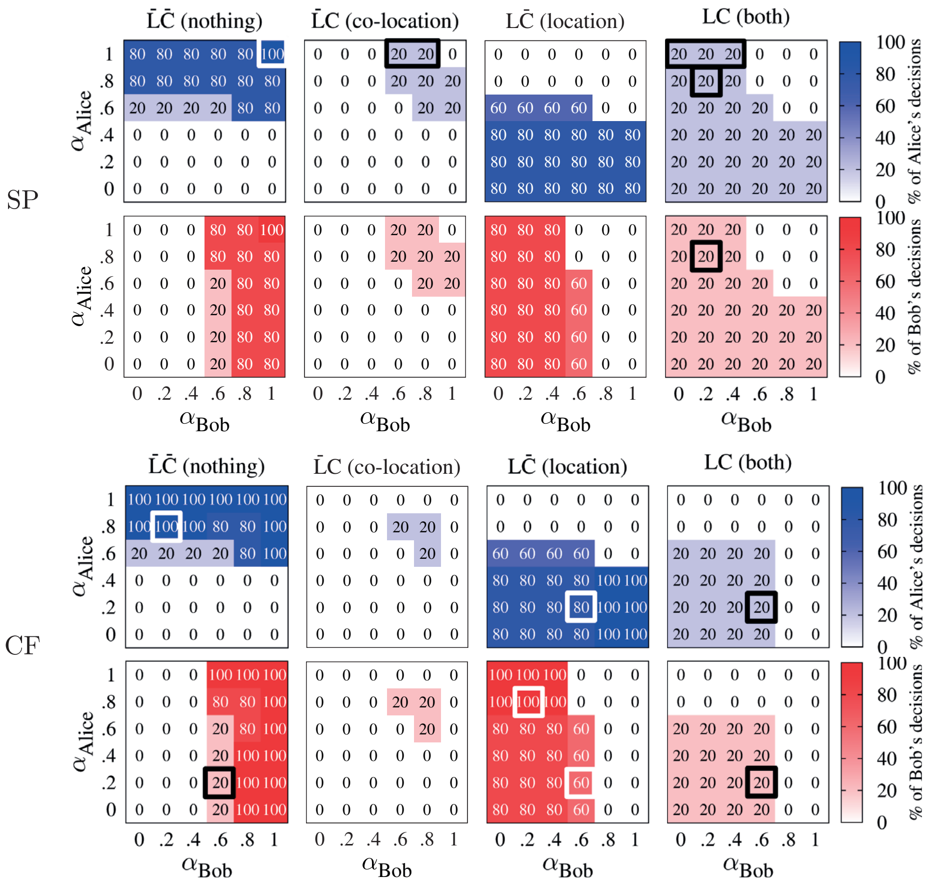


Figure 3.8: Players' decisions at equilibrium, aggregated over time for $f_{sv} = 0.57, f_{lc} = 0.56$, and different adversarial models: Service Provider (SP)–first row, Our friends in common (CF)–second row. For each adversarial model and each possible combination of values for α_{Alice} and α_{Bob}, eight heatmaps (top blue four for Alice, bottom red 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 α_{Alice}-α_{Bob}, the values of the four cells for a player sum to 100). We highlight with rectangles the cases that we discuss in Section 3.5.2.

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 α_{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 $t = 1$ - a detail that is not directly readable form Figure 3.8 on page 63, as it presents statistics aggregated over time instants), Bob decides to not share anything (20% of the times).

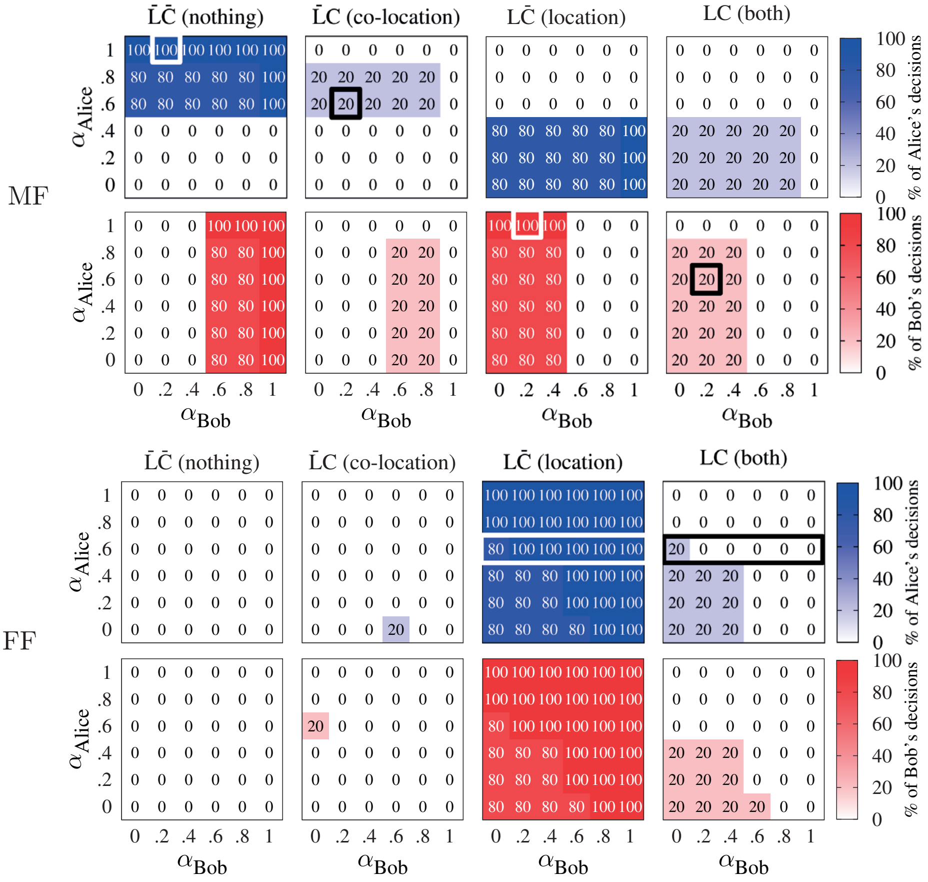
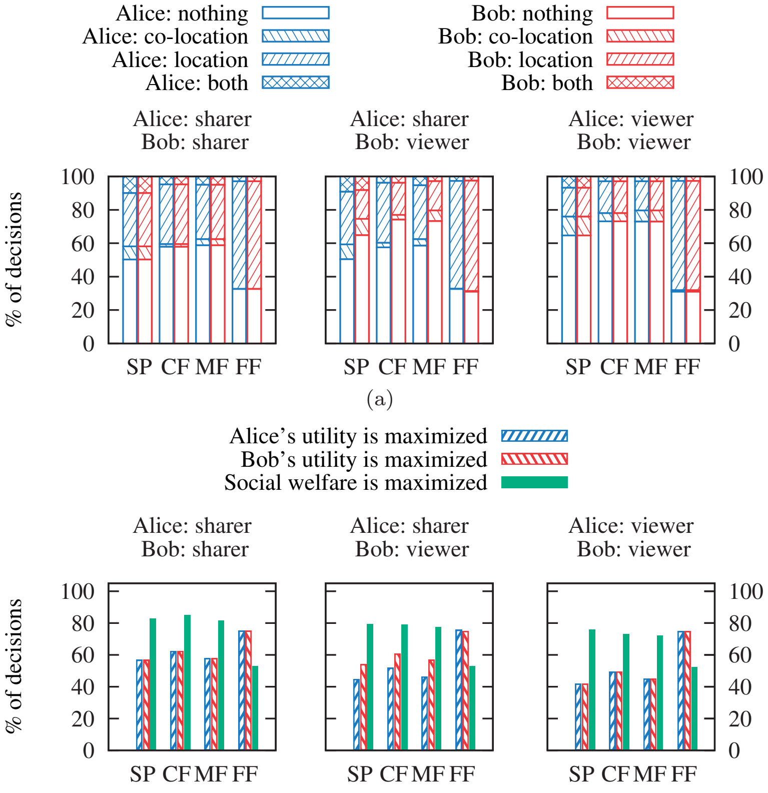


Figure 3.9: Players' decisions at equilibrium, aggregated over time for $f_{sv} = 0.57, f_{lc} = 0.56$, and different adversarial models: My other friends (MF)–first row, My friend's other friends (FF)–second row models. For each adversarial model and each possible combination of values for α_{Alice} and α_{Bob}, eight heatmaps (top blue four for Alice, bottom red 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 α_{Alice}-α_{Bob}, the values of the four cells for a player sum to 100). We highlight with rectangles the cases that we discuss in Section 3.5.2.

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).

The Effects of Multiple Users' Preferences.

We present a more realistic setup, based on the canonical meeting scenario, 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,



(a)

(b)

Figure 3.10: 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.

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 3.10 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 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: (1) the *privacy-oriented* (158 profiles), for which $f_{pb} > 0.5$; and (2) the *benefits-oriented* (92 profiles), for which $f_{pb} < 0.5$. We evaluate the outcome of the Sharing Game in three cases: (1) when Alice is *privacy-oriented* and Bob is *benefits-oriented*, (2) when both are *privacy-oriented* and (3) when both are *benefits-oriented*, for each possible pairs of preference profiles. We note that the interplay between the various parameters of the preference profiles (*e.g.,* a *benefits-oriented* profile encourages sharing because $f_{pb} < 0.5$, but it could also discourage sharing if $f_{sv} < 0.5$) results in a large variety in the distribution of players' equilibria decisions. Yet, a few trends are still distinguishable.

Figure 3.11 on page 67 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. To infer these numbers from Figure 3.11 on page 67, we sum the values for "co-location" and "both". As discussed in Section 3.5.2, in any adversarial model, both players share the same amount of co-location. 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

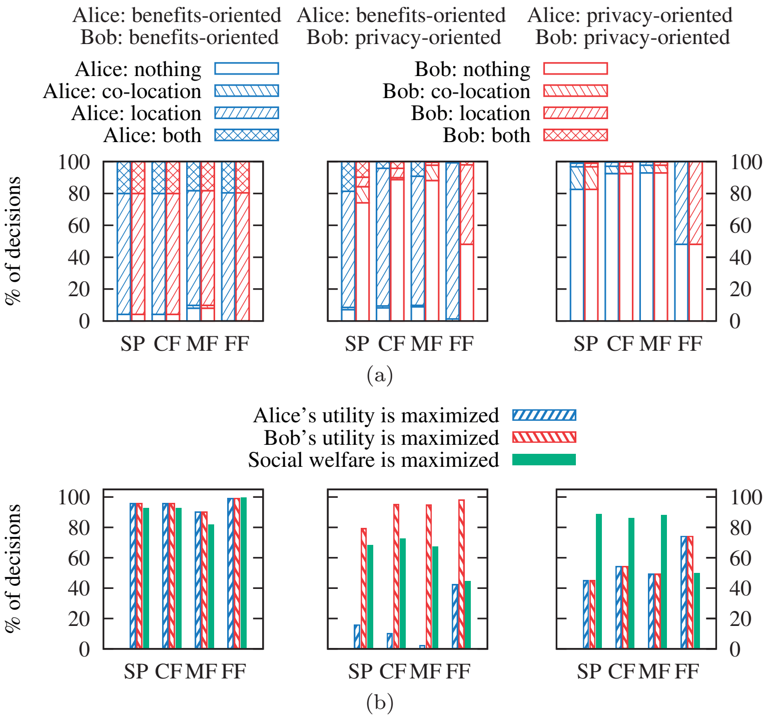
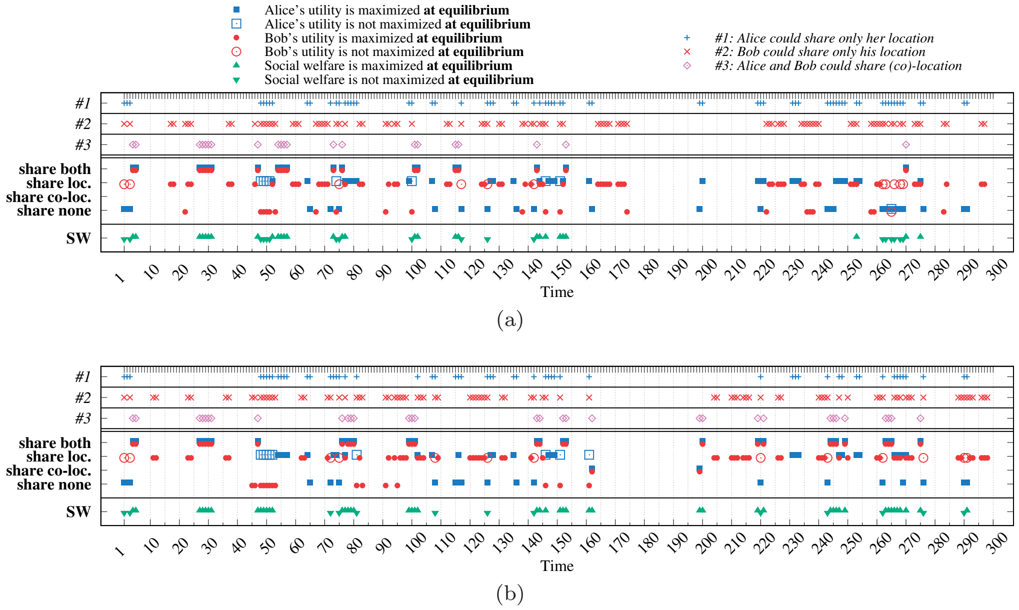


Figure 3.11: Equilibria decisions (a) and their properties (b), when Alice and Bob have different preference profiles, corresponding to real survey data: 92 *benefits-oriented* profiles ($f_{pb} < 0.5$) and 158 *privacy-oriented* profiles ($f_{pb} > 0.5$). We present three scenarios: both Alice and Bob 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*. 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 (left blue for Alice and right red 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 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.

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), and 0.4% of all time instants (only 2% of the time instants when the players are co-located) in the FF case. 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-



(a)

(b)

Figure 3.12: Possible strategies and equilibria decisions (CF adversarial model) when Alice and Bob use two random traces from the Geolife dataset, for $f_{sv} = 0.57$, $f_{lc} = 0.56$, $f_{pb} = 0.60$. We illustrate the game outcome between Alice and Bob in two cases: %7 (a) and 10% (b) co-locations, respectively – the trace used by Alice is exactly the same in both cases. The x axis shows the entire time window of interest. In the top section, on the y axis, the *possible strategies* at each time instant of the game are illustrated: Alice can only share location (blue 'plus'), Bob can only share location (red 'cross'), Alice and Bob can both share location and co-location (violet 'rhombus'); For time instants where the game cannot be played (missing location data) there is no symbol. The middle and bottom sections of the y axis illustrate the *equilibria decisions*: 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 at which both Alice and Bob can play is marked by an upward-pointing green triangle, if the equilibrium decisions maximize social welfare or by a downward-pointing green rectangle if they do not.

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.

Real Dataset Scenario

In this section, we describe results on our **real dataset scenario**. We consider homogenous preference parameters in both the users' utility and set these using the average values of f_{sv}, f_{lc} and f_{pb} obtained in our survey, as presented in Table 3.6 on page 57. For the users' traces, we consider pairs of real traces (60 pairs of traces consisting of 300 time instants) sampled in Chapter 2 and the corresponding complex (and different) mobility profiles. On these traces, users can report co-locations (*i.e.,* meet), on average, 14.6% of the time instants (first quartile, median, third quartiles for the number of co-locations are 10.67%, 12.83%, and 16.67% of the time instants, respectively) and there are location samples at 37.55% of the time instants, on average. We simulate the game between Alice and Bob, for all the 60 pairs of traces, and twice for each of the pairs (ensuring that each of the two traces in a pair is attributed to the first player).

An Individual Snapshot.

We first present two randomly selected simulations of the game interactions between Alice and Bob, illustrated in Figure 3.12 on page 68: We chose a trace for Alice and two different traces for Bob; Alice's trace contains co-locations with that of Bob in 7% and 10% of the 300 time instants, respectively. A first observation is that a **vicious circle effect is noticeable**. After their first shared co-location ($t = 4$), users' behavior changes and they share more than they had done before that co-location: Alice shifts from not sharing anything at $t = 1, 2, 3$ to sharing both location and co-location at $t = 5$ and Bob from sharing only his location ($t = 1, 3$) to sharing both ($t = 5$). We repeatedly observe that, after a shared co-location, users continue to share. Specifically, in the first case (Figure 3.12a on page 68), after a co-location was shared by either of the players, a player's subsequent decision involves sharing location (either only location, or both location and co-location) 95% and 90% of the times, for Alice and Bob, respectively. We consider only the time instants where location is available. This frequency is quite high, indeed higher than the frequency of deciding to share their locations (when available) over the entire traces: 68% and 81% of the time instants, respectively for Alice and for Bob. In the second case (Figure 3.12b on page 68), the same effect is observed 91% and 88% of the times, respectively. Again, this is higher than the frequency of deciding to share their locations over the whole traces: 76% and 86% of the time instants, respectively. **The oversharing effect can, occasionally, be overcome at later time instants.** This can happen when users did not share for some time (either because they did not meet or because their location data is sparse) or when their location is particular (*i.e.,* a rarely visited location, which would yield a higher error for the adversary, thus a higher privacy). Consequently, both Alice and Bob occasionally choose to not share anything (*e.g.,* in Figure 3.12a on page 68, Bob shares nothing at $t = 22, \cdots$; Alice shares nothing at $t = 65, 72, \cdots$; At $t = 74$, Bob shares nothing even though Alice shares her location and, *immediately after*, at $t = 75$, *the situation flips* and Alice shares nothing, while Bob shares his location; in Figure 3.12b on page 68, at $t = 162$ both users only share co-location. Overall, of the times a player's location is available and he decides to share it, 33% and 21% in the first case (Figure 3.12a on page 68), and 43% and 27% in the second case (Figure 3.12b on page 68), respectively for Alice and Bob, happen right after a co-location was shared. To conclude, **even though Alice's location data, mobility profile, and preferences are constant, her behavior can change depending on the friend with whom she interacts, *even if he has the same preferences as she has***. Hence, **the specificities of the data** – the actual locations in the traces, the density of the location data, the meeting frequency (*i.e.,* density of co-locations) and the patterns of the meetings, as well as the quality of the users' mobility profiles – **strongly affect the users' sharing decisions**, even when they agree on their privacy preferences.

The Impact of Co-locations.

We now present the aggregated results of our simulations between all the pairs of traces, illustrated in Figure 3.13 on page 70. We split and aggregate the results in two sets: those with traces that contain fewer co-locations than the median value for

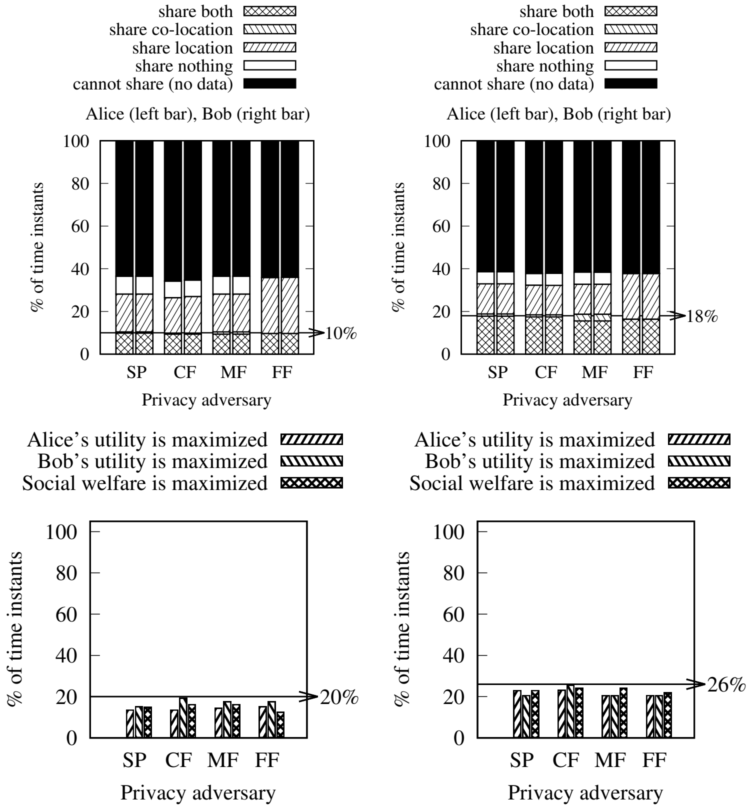


Figure 3.13: Equilibria decisions for Alice and Bob using traces from the Geolife dataset, for $f_{sv} = 0.57, f_{lc} = 0.56$, $f_{pb} = 0.60$. We present two scenarios, depending on the number of co-locations they share: # co-locations lower than the median (left) and # co-locations higher than the median (right). Different adversarial models (SP, CF, MF, FF) are illustrated on the x axis. In the top plots, for each adversarial model, two bars (left for Alice and right for Bob) indicate–on the y axis–the proportion of times (aggregated over the total number of time instants–300–and the number of runs considered in that scenario) a player made one of the four 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); The fact that a player cannot share anything (*i.e.,* missing location data) is illustrated by the solid pattern. The arrows represent, in each scenario, the average number of co-locations. The two bottom plots show, on the y axis, for each adversarial model, the proportion of times (aggregated over the total number of time instants–300–and the number of runs considered in that scenario) social welfare and individual utilities are maximized. Note that at times where only one of the two players has location data, his utility is always maximized, therefore we only illustrate the cases where both players have location data–the arrows represent, in each scenario, the average number of times that this happens.

the entire dataset (12.83%) and those that contain more. We notice an **incentive to over-share co-locations and locations**: In the second case, where Alice and Bob have more co-locations, they *both share more locations and more co-locations, in all adversarial models*. For example, in the CF model, if in the first scenario Alice shares co-location and location 9.8% and 25.8% of the time instants, respectively, in the second one she shares co-location and location 18.8% (representing a relative increase of 87.8%) and 31.3% (representing a relative increase of 21.7%) of the times, respectively. Finally, on this dataset, the average users (in terms of their privacy preferences) would often choose to share some information; This can be attributed to the relatively high number of co-locations in the dataset, as well as to the quality of the mobility profiles. Therefore, the results presented in this section should be taken with a grain of salt, as they will likely differ on other datasets or when considering heterogenous privacy preferences.

3.5.3 Conclusions on the Carried out Experiments

In the *canonical meeting scenario*, which abstracts the specificities of the data, we exposed the fact that **the users' different preferences factors lead to complexity in their interactions**. We noted in the *real dataset scenario* that **the users' equilibria decisions are also highly data dependent**: The users' actual traces, as well as the quality of their mobility profiles, can greatly influence users' sharing decisions, **even when they agree in terms of their preference factors**. It is not hard to imagine how much more complex the interactions between the users can become when different preference factors of the users are taken into account along with real (co-)location data. Understanding them is not a trivial problem and it is hard to draw generally applicable and quantifiable conclusions that take all these variations into account. We exposed, however, a few trends that we believe to be interesting. First, we identified the model parameters and the data specificities that have the strongest effect on the users' decisions: the frequency of co-locations, the quality of the users' mobility profiles, the considered adversary, the value of α (f_{pb}, the preference for privacy versus social benefits of *one* user); whereas other model parameters have a more moderate effect. Second, some interesting patterns of behavior emerge, for instance, 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 shares 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 several later times as well. This effect is propagated and stronger if Alice still wants to share her own location at other time instants, further damaging Bob's privacy. We also observed that the effect of a shared co-location can (sometimes) eventually fade away and that Bob, too, can influence Alice into sharing; these effects can quickly alternate, making the players' decisions vary over time. In other words, *the privacy effects propagate not only in space (influences from a friend) but also in time.* Third, we noticed that, in the FF model, a natural tendency is to share few co-locations but the users still share a significant amount of location information. Finally, we showed (*e.g.,* in the SP and CF models) that a (common) decision to share **co-location** creates the **incentive to *over-share locations at later times after the co-location***.

3.6 Extended Model

We now propose an extended model, relaxing the assumptions of selfishness, privacy-myopia and complete information being available to the players. We define the type of a player i (denoted by $\theta_i(t)$) as the information that is private to her. This can include but is not limited to her actual location ($a_i(t)$), the background information that the adversary has on her ($\mathcal{B}_i$), her specific parameters that influence the computation of her social benefits and her privacy. Thus, $\theta_i(t) \triangleq (a_i(t), \mathcal{B}_i, f_{pb}, f_{lc}, f_{sv}, \dots)$. We assume that a player i has information only about the possible domain and probability distributions of the other player's type and that she incorporates these into her utility function in the form of an expected value. Given $\boldsymbol{\theta}(t) = (\theta_i(t), \theta_j(t))$, the *individual utility* of player i for some strategy profile $\mathbf{s}(t)$ at time t captures both her social benefits and her privacy

$$\begin{aligned}\hat{u}_i\left(t, \mathbf{a}(t), \mathbf{o}(t-1), \mathbf{s}(t), \mathcal{B}_i, \mathcal{B}_j, \boldsymbol{\theta}(t)\right) \quad &= \\ (1-\alpha_i) \cdot B_i\left(t, \mathbf{a}(t), \mathbf{s}(t), \boldsymbol{\theta}(t)\right) + \alpha_i \cdot P&\left(i, t, \mathbf{a}(t), \mathbf{o}(t-1), \mathbf{s}(t), \mathcal{B}_i, \mathcal{B}_j, \boldsymbol{\theta}(t)\right) \quad (3.10)\end{aligned}$$

Thus, the expected individual utility of a player can be computed as

$$\bar{u}_i\left(t, \mathbf{a}(t), \mathbf{o}(t-1), \mathbf{s}(t), \mathcal{B}_i, \mathcal{B}_j\right) = \mathbb{E}_{\boldsymbol{\theta}(t)}[\hat{u}_i\left(t, \mathbf{a}(t), \mathbf{o}(t-1), \mathbf{s}(t), \mathcal{B}_i, \mathcal{B}_j, \boldsymbol{\theta}(t)\right)] \quad (3.11)$$

As previous studies have shown (*e.g.,* [100, 101]), in the context of OSNs where users are friends, they might exercise altruism by taking into account their friends' individual utility and hence choose their strategy based on their *perceived utility*, which we define for some strategy profile $\mathbf{s}(t)$ as follows

$$\begin{aligned}u_i\left(t, \mathbf{a}(t), \mathbf{o}(t-1), \mathbf{s}(t), \mathcal{B}_i, \mathcal{B}_j\right) \quad &= \qquad (3.12) \\ \bar{u}_i\left(t, \mathbf{a}(t), \mathbf{o}(t-1), \mathbf{s}(t), \mathcal{B}_i, \mathcal{B}_j\right) + f_a^i \cdot \bar{u}_j&\left(t, \mathbf{a}(t), \mathbf{o}(t-1), \mathbf{s}(t), \mathcal{B}_i, \mathcal{B}_j\right)\end{aligned}$$

where $f_a^i \in [0,1]$ denotes the altruistic factor of user i for the other user. These factors can be experimentally measured using techniques based on conjoint analysis (*e.g.,* [72, 102, 103]).

Furthermore, as current decisions have privacy implications at future time instants, we consider the *cumulative utility* of user i at time t as the discounted sum of all perceived utilities from time t (the present) until time T (the future) as follows

$$\begin{aligned}U_i\left(t, \mathbf{a}(t), \mathbf{o}(t-1), \mathbf{s}(t), \mathcal{B}_i, \mathcal{B}_j\right) \quad &= \qquad (3.13) \\ \mathbb{E}_{\mathbf{a}[t+1,\cdots,T]}&\left[\sum_{t'=t}^{T} \delta^{t'-t} \cdot u_i\left(t', \mathbf{a}(t'), \mathbf{o}(t'-1), \mathbf{s}(t'), \mathcal{B}_i, \mathcal{B}_j\right)\right]\end{aligned}$$

where δ is a discount factor, taking values in the interval $[0,1]$ and $\mathbf{a}[t+1,\cdots,T]$ denotes the vector of actual locations at times $t+1,\cdots,T$ for both i and j. Note that, in computing her cumulated utility at time t, a user does not know any of the actual locations in the future ($t' > t$), hence the expectation value over $\mathbf{a}[t+1,\cdots,T]$. Given these locations, $\mathbf{s}(t')$ can be deterministically predicted.

At every time instant, we model the interactions as an incomplete information, extensive-form game (where the players observe each other's strategies). Solving such an extended game is not straightforward. Intuitively, players could choose the strategy that maximizes their expected utility with respect to the unknown information (*i.e.,* the other player's type). However, the way in which players choose to estimate and use the unknown information (in the backward induction algorithm) strongly affects the complexity of the game. We plan to study the different possibilities and give a numerical solution in future work.

3.7 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. For the first attempt to tackle the problem of understanding users' interaction in such a complex context, we focused on a number of specific scenarios and assumptions, that open several interesting directions for future work. We assumed that 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 included, in our evaluation, location privacy with respect to one adversary. 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* can be included in the utility function. 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. Our evaluation focused on a few specific cases and maintained the number of free parameters low. We quantified only a limited number of preference factors (whose values could be specific to Facebook users) 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 and their respective weights in the utility function; the benefit users gain for sharing fake co-locations). Furthermore, previous works (e.g., [95, 96]) have shown that user (reported) privacy attitudes do not always correspond to actual behaviors, and we have demonstrated that users decisions are highly data dependent. Hence our results should also be taken with a grain of salt: We believe that the trends that we have exposed are generic, but their magnitude will likely differ on other datasets or when considering heterogenous user privacy preferences. Finally, in future work, we plan to study the N-player T-time Sharing Game, by using multi-agent influence diagrams (MAIDs), which were introduced by Koller et al. [104] for efficiently solving complex games. Ultimately, our extensible model with quantifiable parameters can serve as the first building block to assist user decision-making in an informed manner. Specifically, we envision that a software tool (e.g., a Facebook client) would use our framework to take these interactions into account and to assist users in improving their awareness and decision-making process.

3.8 Related Work

Our work is related to two broad research areas: information sharing on OSNs and interdependent privacy with game theory. We survey related existing works in these areas and position our work.

3.8.1 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. Laufer et al. [91, 92] coined the term *privacy calculus*; it consists in a psychological framework formalizing users' decision making process through a cost-benefit analysis when sharing information. However, 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 or the service providers who can access it [105–108]), the shared information itself, and the benefits and privacy risks [109] 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 [100] and experimentally measured using techniques based on conjoint analysis in [102, 103]. Conjoint analysis studies were also used to quantify the value that users attribute to their friends' information in the context of app adoption (*e.g.,* in [101]). In practice, deciding whether to share information often comes down to finding a sweet spot between privacy and benefits [110]. The decision process can be automated by (1) maximizing privacy under benefits (service quality) constraints [111] (or conversely), (2) taking a game-theoretic approach for modeling the interplay between the users and the adversaries [112], or (3) by mimicking the users' sharing decisions using machine-learning techniques, after a training phase [113].

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.

3.8.2 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 [71]. 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 [71, 72], sensitive attributes inferred from those of a users' friends on OSNs [9, 10, 68], demographic information inferred from a user's interests [8], detecting private OSN profiles through attributes correlation within social circles [11], genomic data inferred from that of relatives [73, 114], location leaked from geo-tagged pictures that friends upload online [67], relationships inferred from pictures [115], and co-locations detected from the users' IP address at hotspots [69] or reported on OSNs [22].

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) [79, 80, 116–123]. These works focus mostly on the case of photo sharing on online social platforms and take an experimental and empirical approach to the problem–*i.e.,* they rely on interviews and surveys. Misra et al. [124] propose a personal agent that recommends personalized access control decisions; Fogues et al. [125] propose a machine-learning based policy recommender to predict the optimal sharing policy in multiuser scenarios. 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 [71, 72], the dynamics of individuals' privacy preferences regarding shared content [126], and privacy decision-making [127], such as sharing genomic data [73]. The study of interdependent privacy risks from an economic perspective follows the long line of research on interdependent security games surveyed in [128].

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 in the game-theoretic approaches surveyed above, in our framework we take into account the time dimension, future considerations, incomplete information, and an altruistic factor. In addition, we rely on a rigorous approach, based on user surveys, to determine realistic values of the different parameters of our model.

3.9 Conclusion

It is well-known that the behavior of others affects 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 (*e.g.,* we exemplified on a mobility dataset how a vicious-circle effect emerges) and we demonstrated that sharing co-location information can additionally encourage users to over-share their locations. This is an interesting finding from a design perspective for the OSN service providers but a dangerous one for the end users: Advertising features that permit the sharing of co-location information could also encourage users to share their locations more often. Furthermore, we showed that user's decisions are strongly influenced by the adversary that they consider and dependent on the mobility data. We emphasized the need to develop appropriate warning mechanisms for the users, which we intend to develop in the future; these would help users better understand and anticipate the consequences of their (co-)location sharing decisions.

Survey transcript

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)

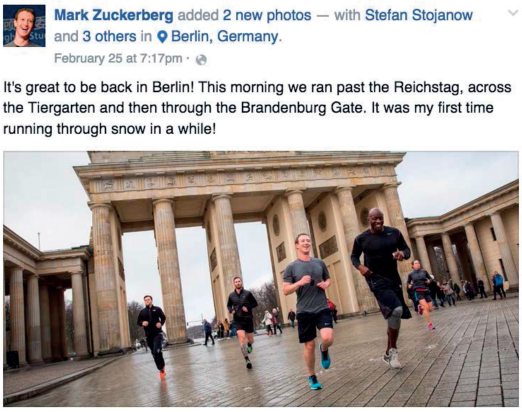
   [. . .]

   ○ 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. The picture depicts an example of a check-in post.



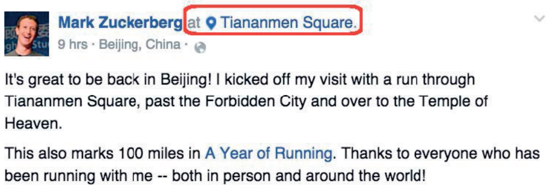
Mark Zuckerberg added 2 new photos — with Stefan Stojanow and 3 others in Berlin, Germany.
February 25 at 7:17pm ·

It's great to be back in Berlin! This morning we ran past the Reichstag, across the Tiergarten and then through the Brandenburg Gate. It was my first time running through snow in a while!

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, as illustrated below.

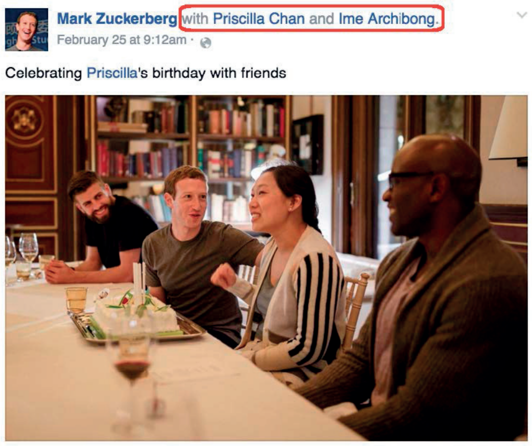


Mark Zuckerberg at Tiananmen Square.
9 hrs · Beijing, China ·

It's great to be back in Beijing! I kicked off my visit with a run through Tiananmen Square, past the Forbidden City and over to the Temple of Heaven.

This also marks 100 miles in A Year of Running. Thanks to everyone who has been running with me -- both in person and around the world!

A co-location post is a post in which you tag the friends you are with - either through a status message or a picture - as illustrated below.



Mark Zuckerberg with Priscilla Chan and Ime Archibong.
February 25 at 9:12am ·

Celebrating Priscilla's birthday with friends

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	

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.

Station
Park Central Square
50m
Big Whiskey's American Bar & Gril
The Coffee Ethic
S Patton Ave
Park Central Office Building
W McDaniel St
200
300
Ernie Biggs
n Ave

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	

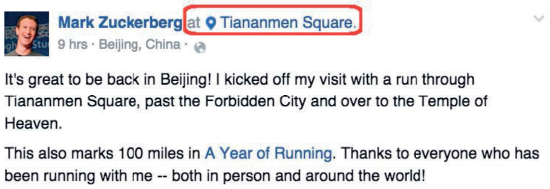
* These numbers were extracted from the experimental results presented in [22].

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, as illustrated below.



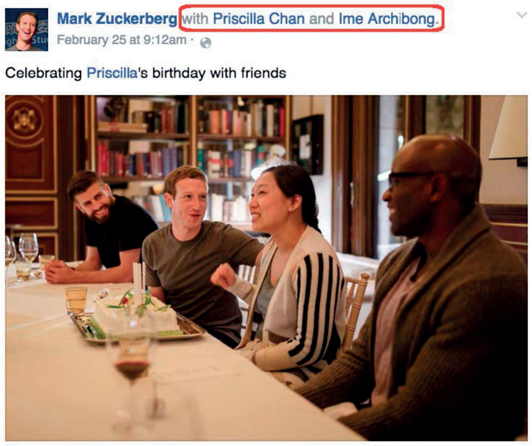
Mark Zuckerberg at Tiananmen Square.
9 hrs · Beijing, China ·

It's great to be back in Beijing! I kicked off my visit with a run through Tiananmen Square, past the Forbidden City and over to the Temple of Heaven.

This also marks 100 miles in A Year of Running. Thanks to everyone who has been running with me -- both in person and around the world!

○ 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? An example is illustrated below.



Mark Zuckerberg with Priscilla Chan and Ime Archibong.
February 25 at 9:12am ·

Celebrating Priscilla's birthday with friends

○ 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?

Chapter 4

Privacy-preserving Sharing with Consent

> *What can you do against the lunatic who is more intelligent than yourself, who gives your arguments a fair hearing and then simply persists in his lunacy?*
>
> George Orwell

4.1 Introduction

Individuals share increasing amounts of personal data online. Powered by the emergence of specialized platforms, such as OSNs, the variety of the personal data shared online has also substantially increased over the last decade, including content as diverse as contact data (address books), multimedia data (photo, audio, videos), location data and genomic data.

Recent studies highlighted the fact that such data often involves (and has privacy implications for) data subjects other than the individual who shares them online [129]. This concept, referred to as *multiple-subject personal data* (MSPD; the term was coined by Gnesi et al. [26]) or as *co-owned/multi-party data* (by Such et al. [80]), applies to numerous types of data, one of the most widespread examples being group photos and videos. A sadly popular example [15, 16, 130], with dramatic consequences, is revenge pornography (*i.e.,* the disclosure of photos or videos portraying sexually explicit activity, typically after the end of the relationship between the partners), which can also occur on platforms such as Facebook [14, 131]. Our preliminary results show that the number of potential victims is very high (40.2% of our survey participants declare that they have explicit pictures of other people).

Beyond MSPD data, recent studies showed that seemingly strictly personal data reveals information about other individuals [9, 10, 12], an effect which we also study in Chapter 2. This concept is referred to as *interdependent personal data* (IPD; the term *interdependent privacy* was coined by Biczók and Chia [71]). The root cause of interdependent privacy is the fact that the personal data of somehow related people (e.g., friends, colleagues, relatives) are correlated. A typical example, introduced by Humbert et al. [12], is genomic data: The genomes of individuals, shared on specific platforms (*e.g.,* 23andme), reveal information about the genomes of their relatives.

Most of the time, the sharing and the disclosure of multiple-subject or interdependent personal data occurs without the consent of the involved individuals, possibly creating so-called multi-party privacy conflicts [79, 80, 123], which are known to be difficult to resolve. Although the notion of consent is known to be fundamental and at the core of most of data-protection and privacy laws, as well as terms of use of online sharing platforms, very few technical solutions exist, to the best of our knowledge, for detecting and sharing such data, in a consensual and privacy-preserving way. Several protocols have been studied [117–122], but there are no associated tools to aid users to implement these and, more importantly, they are all based on the assumption that users are aware when data regarding them is shared, which is not always the case. Existing technical solutions are limited in terms of the considered adversary (*i.e.,* they typically disregard the case where the data is disclosed to the service provider), of the detection of the data-subjects and of the privacy guarantees. For instance, Facebook enables its users to review the tags that identify them in photos before they are made visible to other users, and possibly remove them; yet, even though such a tag could eventually be removed by a user, Facebook does have access to the corresponding information, *i.e.,* the fact that the tagged user most probably appears in the photo.[1]

In this chapter, we tackle the problem of designing and building a system for sharing, in a consensual and privacy-preserving way, multiple-subject or interdependent personal data (MSPD/IPD). Specifically, in accordance with Nissenbaum's definition of privacy as contextual integrity [132], we seek to give individuals control on the dissemination of data that involves them. This problem is difficult for several reasons. Identifying the data subjects of some data, or more generally the individuals whose privacy can be affected by the disclosure of the data, is far from trivial and highly data-dependent. In addition, the fact that this identification, as well as the collection of the consent and the preprocessing/sharing of the data (in compliance with the obtained consent) must be done in a privacy-preserving way, with respect to the involved service providers and individuals, makes the design of such a system even more difficult. Our survey results (Section 4.7) suggest that users are both concerned about this threat and potentially interested in our proposed solution (53.6% of the participants said they would use such a system).

We propose a generic solution able to handle various types of such data, and we identify the different building blocks of a system for sharing data online, as well as the design choices to adapt to the specifics of the different data types. We focus on the case of photos, and we design and implement a working solution named ConsenShare. ConsenShare relies on two different entities: an identity management service (IMS) and a content management service (CMS). The first is in charge of identifying[2] and contacting the individuals involved in the data about to be shared on the platform that is operated

[1] Note that the term sharing includes posting on social platforms but also sending data through (instant) messaging platforms, e.g., WhatsApp.

[2] While not privacy-mindful, an application for identifying people from a picture taken in public, FindFace [133], is becoming popular in Russia.

by the second. The second is in charge of collecting the data and the consent, and of preprocessing and sharing the data. At the core of ConsenShare lies a distributed protocol based on standard cryptographic primitives and image processing operations, which ensures that the information learned by the IMS, the CMS and the involved individuals is minimal, especially in the case where some of the involved individuals do not give their consent. An example of a typical setting for ConsenShare would be, for the case of photos, Facebook acting as the IMS and Flickr as the CMS. ConsenShare is, to the best of our knowledge, the first such system; it addresses an important and timely problem. In fact, using such a system before sharing MSPD/IPD data online might become mandatory by law in a few years. Service providers and law makers are already making efforts in this direction, in particular for revenge pornography [134–139]; these are not perfect–from a privacy perspective–for the users. Such a solution would aid with law suits avoidance, as a CMS might be held liable for allowing the sharing of MSPD/IPD data (as was the case with fake news on OSNs). Furthermore, as our solution would represent a user-desired feature in an CMS, adoption might also lead to increasing the user base (and the revenue).

We perform a security and privacy analysis of ConsenShare. By using an unbiassed random sample of 17k+ photos from Yahoo's YFCC100m dataset (Flickr [140]), we also evaluate its performance in terms of CPU and bandwidth consumption, in the (worst case) scenario where all the individuals who appear on a photo are asked for consent. Our experimental results show that the CPU time is negligible for the users and for the CMS. As for the bandwidth overhead (w.r.t. to the baseline case where users directly upload their photos to the CMS), this is approximately equal to the photo size for the user who uploads the photo (as the photo must be sent to the IMS, in addition to the CMS) and 34.78% for the CMS; for the IMS, the bandwidth usage is roughly equal to the size of the uploaded photos. We complement our evaluation with an online survey on multiple-subject and interdependent personal data, targeted at Facebook users and conducted via the Amazon Mechanical Turk platform (N=321). The survey results indicate that a system like ConsenShare could be desirable. For instance, 69.5% of the participants are concerned by the sharing of multimedia data that involves them, 27.4% are potential victims of revenge pornography (*i.e.,* they have shared intimate photos or videos), and 53.6% would certainly use a system like ConsenShare. We also study the potential adoption of such a system by analyzing the incentives (*e.g.,* business opportunities and models) of the different stakeholders, namely the end-users, the IMS and the CMS. In summary, our contributions are the following: (1) We frame the timely and critical problem of consensual and privacy-preserving sharing of MSPD/IPD data (2) We design, implement and evaluate the first system to address this problem for photos; we also propose a generic system for other types of data and identify the different challenges inherent to its design, as well as incentives for adoption for all the parties involved. Our results are quite encouraging: The experimental results demonstrate the feasibility of our approach and the survey results demonstrate potential interest from the users.

The remainder of the chapter is organized as follows. We describe the system model and list our design goals in Section 4.2. We give a high-level description of a generic solution, namely ConsenShare, in Section 4.3. We propose and give a detailed description of a solution specific to photos in Section 4.4. We provide a security and privacy analysis of ConsenShare in Section 4.5. We report on our data-driven experimental evaluation of ConsenShare in Section 4.6. We discuss the adoption of ConsenShare, based on (among other things) the results of our user survey, as well as its limitations and its extension to data other than photos in Sections 4.7 and 4.8 respectively. We survey the related work, with an emphasis on legal, social and technical aspects of the problem, in Section 4.9. We conclude the chapter in Section 4.10. Finally, in Section 4.10, we provide the full survey transcript.

4.2 System Model & Design Goals

We describe next our system model, the adversaries and the threat model we consider, our assumptions and design goals.

System Model. In our model, we consider the following major entities: Users and a Content Management Service (CMS) – *e.g.,* Flickr for photos, YouTube for videos, or OpenSNP for genomic data. Users[3] can upload content to the CMS (with a certain target audience for visibility consisting of a set of users and/or the general public); any part of the content that concerns (or has privacy implications for) another user is sent to her for approval, along with any relevant contextual data (*e.g.,* the identity of the uploader, description, upload time, target audience, *etc.*); this content is visible to these parties (and to the CMS) *only if* the concerned user grants their consent.

Threat Model. In our model, we assume that the adversaries are the users (individuals), the online services (*e.g.,* the CMS–other services can be included in the protocol as we shall see, *e.g.,* the IMS) and third parties (*e.g.,* external observers). Individuals can be active adversaries. For instance, a malicious user could try to bypass the system to fully publish sensitive content (e.g., compromising photos of other users) without obtaining consent from the affected users (possibly by colluding with other malicious users or by creating fake profiles). A malicious user might also try to monitor and tamper with the communications among the different parties in our system to infer private information about other users, e.g., their real names. The CMS and the IMS are assumed to be honest-but-curious, *i.e.,* they will follow the protocol, but they could try to infer sensitive information from the data observed. For instance, the CMS might try to learn the sensitive content specific to particular users before they give consent or infer the social networks or strength of social ties of some users based on the consent requests that are sent out and their responses (*e.g.,* if Bob often accepts that Alice share content regarding him, they are likely to be close).

System Assumptions. We assume that secure two-way communication channels have been established between all parties in our system, typically over SSL; we assume that the CMS and the IMS are independent parties and that they do not collude (we discuss the case of collusion in Section 4.8). We further assume that data from the network layers (*e.g.,* IP address) cannot be used to leak users' identities: This is a reasonable assumption as many mobile users only access the Internet through a NAT gateway offered by their Internet provider, but could be relaxed if, for instance, users make use of VPN service or anonymous networks (e.g., Tor) to access the Internet. We do not consider fingerprinting attacks in our model. Finally, we assume that software that is run locally is trusted (trusted execution environment can be used).

[3] Note that we refer to "regular" users; we do not consider professionals such as journalists, who follow specific accountability rules regarding the publication of content, are liable for it and have a reputation to uphold.

Design Goals. Our main goal is to design a mechanism that, in a *private* way, (1) informs users every time a piece of content regarding them is submitted and (2) enables them to grant their consent *before* such content is available to any other party (except from the uploader, of course). To this end, the design goals of our system are as follows.

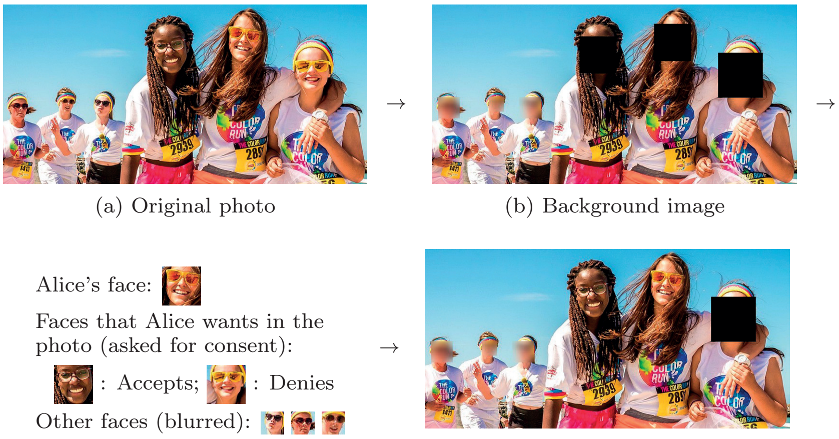
- **Effectiveness**: the registration process should be secure, registered users should be detected in uploaded content and the sensitive content involving them should not be revealed to any component of our system until *after* they consent.
- **Privacy as anonymity** for the users: data submission and consent operations should not leak information about the identity of the users involved (other than the uploader).
- **Unlinkability**: An adversary should not be able to aggregate or link consent operations regarding different individuals associated with a particular content.
- **Detection of any malicious user behavior** (*e.g.,* attempts to bypass the protocol); the system should not allow the sharing of sensitive parts of the content, in such cases.
- **Usability and transparency** to the users. This includes the fact that consent operations should provide the users with enough contextual information for them to make an informed decision.

4.3 Highlevel Solution

In this section we describe our proposed framework, its core components and the main technical challenges.

Framework Overview. We envision a system to which a user registers with identity information. The system's role is (i) to detect, for any content that is uploaded, what are the users affected by this content (*e.g.,* for genomic data these are the close relatives; for photos and videos these are the people who appear; *etc.*), (ii) to contact them and ask them for consent, providing them the option to express a decision either manually, through policies, or automatic (machine learning based) and (iii) to publish the content with the proper restrictions and obfuscations (depending on the users' consent decisions). Such a system consists of several components, which we describe next.

- **Content Management Service (CMS)**. These are User Generated Content sites, such as Flickr, YouTube or OpenSNP.
- **Identity Management Service (IMS)**. This handles users' identities and offers services to identify the users associated with a given content. The IMS is in charge of users' relationships (social and family). In practice, the role of the IMS could be played by the public administration or by popular OSNs (*e.g.,* Facebook) or, it could be distributed across several entities.
- **User applications (CMS and IMS)**. This is the component that users interact with to publish content form their devices to the CMS, to review consent requests either manually, through the use of policies, or machine learning automation (learning a user's decisions from a few initial manual decisions).



(a) Original photo

(b) Background image

(c) Face images

(d) Final photo

Figure 4.1: Example application of ConsenShare on a real photo: (a) Original photo, taken by Alice; (b) Background image produced by Alice's application and sent to the IMS and CMS. Note that faces that Alice wants to appear in the photo will be asked for consent, whereas others (*e.g.,* people in the background) will be simply blurred as in Google Street View; (c) Face images, extracted from the original image, by Alice's application; (d) Final photo, hosted by the CMS, assuming one of the people accepts and the other denies (note that Alice's consent is automatically granted, as she is the uploader). In order to enable the people who appear in the photo to make an informed consent decision with the general context in mind, they receive the background image with only their own faces shown (for privacy reasons). Note that the sizes of the blurred/cropped out regions can easily be customized and increased even to include the full body (*e.g.,* [141,142]) to avoid people being recognized through features other than their face, such as clothes. Image source: Pixabay, (https://pixabay.com/p-1517163/?no_redirect).

Challenges. Key and challenging components of our framework include: claiming identity; determining the users involved in some particular content; contacting them privately and providing them with enough context to make informed consent decisions (while hiding information for which other involved users should grant consent); the variety of types of the consent decisions (removing or obfuscating all or some parts from the content, reducing the visibility audience, *etc.*); reducing the number of consent decisions (through policies or machine learning automation); and enforcing multiple and possibly conflicting individual consent decisions–all of these in a private way. Most of these components are very data-dependent. Therefore, for simplicity's sake, in what follows we will focus on photos. In Section 4.8, we discuss the extension to other types of MSPD/IPD.

4.4 Specific Solution: The Case of Photos

In this section, we present a working solution, for the case of photos. The main entities in this case remain the photo uploader (we refer to her as Alice), the CMS, the IMS and (potentially) the other users that appear in a particular photo (consenters). We refer to any consenter as Bob.

4.4.1 Overview of ConsenShare

ConsenShare enables any user, Alice, to upload photos to the CMS (see Figure 4.3 on page 89. If such a photo, P, contains faces of other people, Alice can choose to remove these (*e.g.,* by blurring them, similarly to blurring on Google Street View [143])[4] or – if she wants these to be visible in the photo – she must first remove the faces from the photo, upload them (encrypted) separately such that the corresponding people are asked for consent (Figure 4.1 illustrates on a concrete example how some of the photos look). In this latter case, only the background corresponding to photo P, namely P_B, is uploaded to the CMS (after some validation from the IMS to certify that no (known) faces appear in it). This version of the photo (P_V) is made visible to the target audience desired by Alice as soon as the upload completes. Faces for which consent must be asked are cropped out from P_B and a protocol to identify the owner of the face, contact him, provide him the photo for review and collect his consent decision is executed; this involves the IMS at different stages. We emphasize that the parts of the photo for which other users must consent are protected, as one consenter, Bob, will only be provided with the photo consisting of the background and his own face. Before Bob grants consent, only Alice (who already has access to the full photo) and Bob are able to see the part of the photo containing Bob's face, as Bob's face image is encrypted using a key created by him. In addition to this, Bob is also provided with some contextual information about the photo (such as the identity of the uploader, upload time, description and the target audience for photo visibility). If Bob denies consent, his face will remain cropped out in the published photo, P_V. If Bob grants consent for his face to appear in the photo, he provides the CMS with the needed information to decrypt his face (*i.e.,* a key). Before adding Bob's face to P_V, the CMS performs validation steps to ensure that Alice or another party has not tampered with the original face appearing in P and that consent has been granted by the correct user.

4.4.2 Technical Challenges and Choices

Our solution comes with several challenges, discussed here.

Identity Claim

In order to use ConsenShare, users (or their legal guardians for minors) must register by providing information for face recognition, which would be used to detect them in photos shared by others in the future. A typical way to do this, which became mainstream, is to provide the system with an ID and/or photos (which can be verified by humans), as exemplified, for instance, by Uber [145,146], Airbnb [147] and potentially Facebook [148]. To reinforce the proof of identity, webcams can be used, similarly to Microsoft's Windows Hello [149], other solutions for biometric authentication proposed in the literature [150–152] or Apple's biometric facial recognition (FaceID [153, 154]) for unlocking iPhones.

[4] Note that recent works, such as that by Li et al. [144], study various common techniques for obfuscating photos, both in terms of effectiveness and impact on viewing experience. Our design is flexible enough to enable the use of other photo obfuscation techniques, beyond blurring.

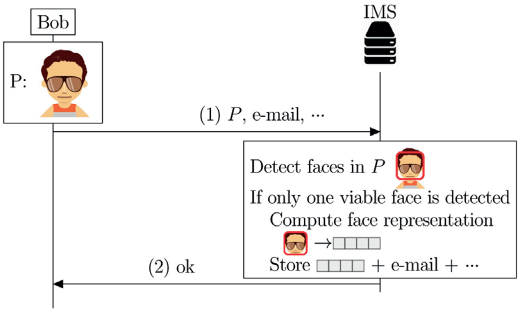


Figure 4.2: ConsenShare: Register user protocol

Privacy-preserving Face/Body Recognition

A major challenge in the design of ConsenShare is the privacy-preserving face detection and recognition. Although there is work in the literature detailing how classification could be performed on encrypted data (*e.g.,* [155–159]), it is not clear how applicable these would be to the rapidly evolving face recognition algorithms, or how efficient these would be. Furthermore, this option would raise the problem of the authenticity of the classification results. Thus, in our design, we focus on face recognition on regular images; We consider feature vector based face recognition, as described in Google's popular and efficient FaceNet framework [160]. We provide the desired privacy guarantees by: performing the face detection operations locally on the uploading client's device (thus the photo is not shared with any other parties); performing the face recognition on the IMS server based on the much less sensitive information, *i.e.,* the feature vectors; and validating these operations by the CMS using the original photo once consent is granted. Note that all local operations can be performed either, natively, in a mobile application, or a web app (Javascript).

To better understand these design decisions, we give here some background information about face recognition, which the familiar reader can skip. A face representation (or feature vector) is a multi-dimensional numerical vector that encodes the features of a face (*e.g.,* the eye distance). Its main properties are: (i) A face representation is unique to a face image, hence different face photos (even belonging to the same person) result in different, yet close in terms of distance, face representations. (ii) Typically, the Euclidean distance between face representation extracted from photos depicting the face of the same person is smaller than the distance between representations extracted from photos depicting faces of different people. Thus, distances can be translated into a measure of face (dis)similarity and the problem of face recognition for some input feature vector reduces to identifying the closest feature vector to it – in the distance space – from a set of available feature vectors of the registered users.

Note that detecting faces might not be enough, as recent work shows that identifying people is possible from features other than their face, such as their clothes [161]. Our framework can be extended to include body detection and obfuscation techniques (*e.g.,* [141, 142]), as for faces; such solutions would provide more privacy, but would also involve a utility loss due to the increased obfuscation.

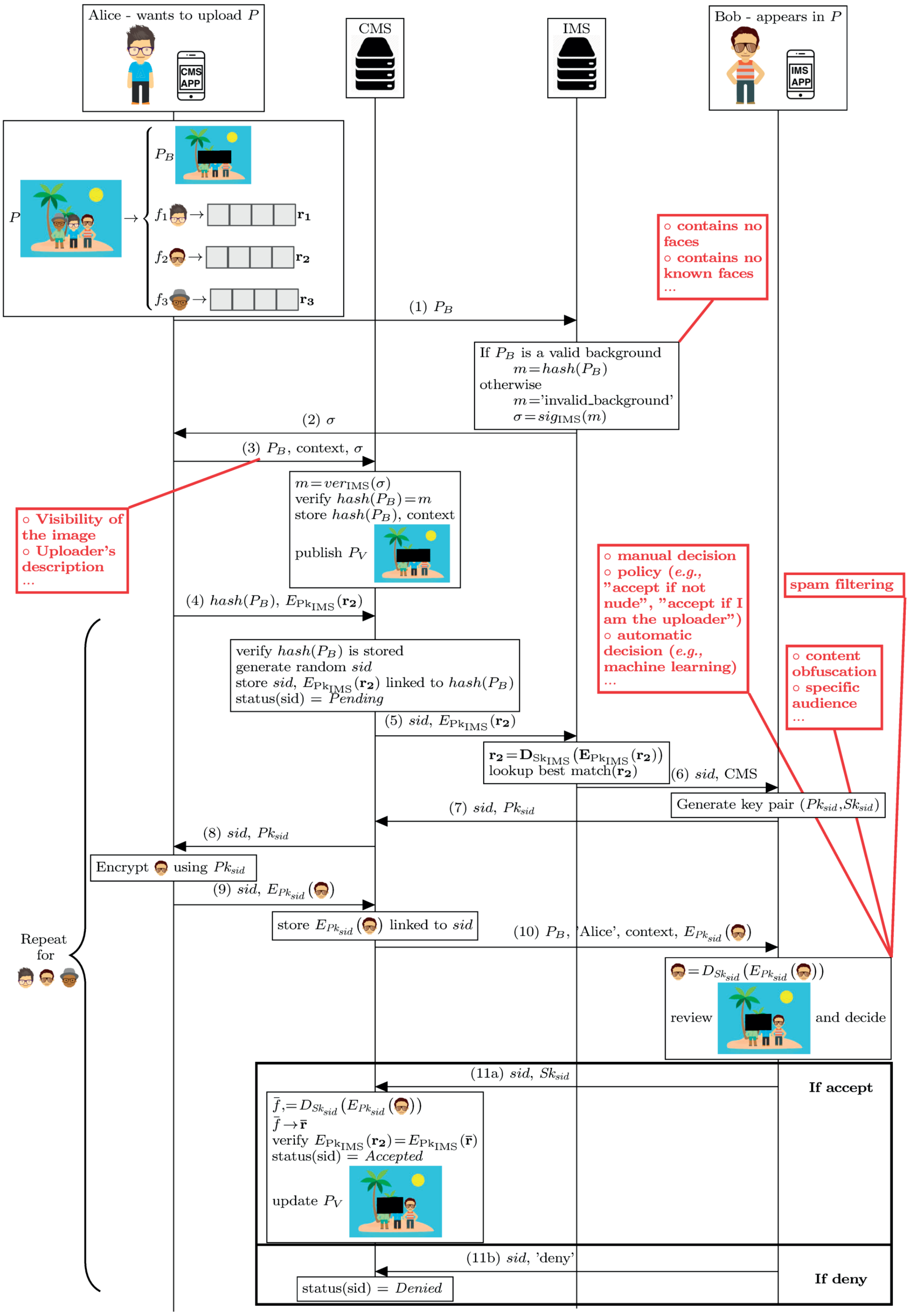

Alice - wants to upload P
CMS APP
CMS
IMS
Bob - appears in P
IMS APP
P
P_B
f_1 → $\mathbf{r_1}$
f_2 → $\mathbf{r_2}$
f_3 → $\mathbf{r_3}$
○ contains no faces
○ contains no known faces
...
(1) P_B
If P_B is a valid background
$m = hash(P_B)$
otherwise
m='invalid_background'
$\sigma = sig_{\text{IMS}}(m)$
(2) σ
(3) P_B, context, σ
○ Visibility of the image
○ Uploader's description
...
$m = ver_{\text{IMS}}(\sigma)$
verify $hash(P_B) = m$
store $hash(P_B)$, context
publish P_V
○ manual decision
○ policy (e.g., "accept if not nude", "accept if I am the uploader")
○ automatic decision (e.g., machine learning)
...
spam filtering
(4) $hash(P_B)$, $E_{\text{Pk}_{\text{IMS}}}(\mathbf{r_2})$
verify $hash(P_B)$ is stored
generate random sid
store sid, $E_{\text{Pk}_{\text{IMS}}}(\mathbf{r_2})$ linked to $hash(P_B)$
status(sid) = *Pending*
○ content obfuscation
○ specific audience
...
(5) sid, $E_{\text{Pk}_{\text{IMS}}}(\mathbf{r_2})$
$\mathbf{r_2} = \mathbf{D}_{\text{Sk}_{\text{IMS}}}(\mathbf{E}_{\text{Pk}_{\text{IMS}}}(\mathbf{r_2}))$
lookup best match($\mathbf{r_2}$)
(6) sid, CMS
Generate key pair (Pk_{sid},Sk_{sid})
(7) sid, Pk_{sid}
(8) sid, Pk_{sid}
Encrypt using Pk_{sid}
(9) sid, $E_{Pk_{sid}}$( )
store $E_{Pk_{sid}}$( ) linked to sid
(10) P_B, 'Alice', context, $E_{Pk_{sid}}$( )
Repeat for
$= D_{Sk_{sid}}(E_{Pk_{sid}}($ $))$
review
and decide
(11a) sid, Sk_{sid}
If accept
$\bar{f}_{,} = D_{Sk_{sid}}(E_{Pk_{sid}}($ $))$
$\bar{f} \to \bar{\mathbf{r}}$
verify $E_{\text{Pk}_{\text{IMS}}}(\mathbf{r_2}) = E_{\text{Pk}_{\text{IMS}}}(\bar{\mathbf{r}})$
status(sid) = *Accepted*
update P_V
(11b) sid, 'deny'
If deny
status(sid) = *Denied*

Figure 4.3: ConsenShare: Upload photo protocol

4.4.3 ConsenShare Main Operations

We describe the following operations: (i) register (Figure 4.2) and (ii) upload photo and grant/deny consent (Figure 4.3 on page 89).

Register

The operations performed are the following.

1. Bob, a new user, uses the ConsenShare IMS application on his device to take a set of photos/videos, P, with his webcam. The application sends P, along with Bob's login information to the IMS. The idea behind this is to prevent people claiming an identity different than theirs.

2. The IMS detects the faces in P. If P contains only one valid face, it computes the corresponding face representation, stores this along with the provided login information and sends an 'ok' response to Bob. Otherwise, an error message is returned to Bob and nothing is stored.

Upload Photo and Grant/Deny Consent

The operations performed are the following.

1. Alice selects the photo she wants to share: The CMS ConsenShare application splits it into background image, P_B, (which contains no faces) and the detected faces; this is done locally by the CMS provided application or webpage. Alice marks the faces that she does *not* want to ask for consent and these are blurred; the other faces are cropped out. Alice sends this P_B to the IMS. The IMS should provide a public API for this operation. Note that Alice does not need to have an IMS account.

2. The IMS performs face recognition on the received P_B. If this is indeed a valid background photo (*i.e.,* contains no faces, contains no known faces, *etc.* depending on the policy implemented by the IMS–it could also depend on the requirements imposed by the CMS), the IMS signs and sends a message containing the hash of the background photo (and information indicating whether the photo contains no face, no known faces, *etc.*). Otherwise, it signs and sends a message specifying the reasons for which the photo is not valid. The signed message, σ, is returned to Alice.

3. Alice logs in to the CMS using her credentials and forwards the received σ together with P_B and some *context* information (*e.g.,* the desired visibility and a description of P) to the CMS, which verifies that all of the following holds: (1) σ is a valid signature by the IMS, (2) the P_B it received from Alice has the same hash as the one in message returned by the IMS – to prevent Alice from uploading a different photo. If these checks pass, the CMS creates a visible photo with the specified audience in *context*, P_V, consisting of the background image P_B and stores $hash(P_B)$ and *context*.

For each detected face that must be asked for consent:

4. A face representation is computed locally by the ConsenShare CMS app on Alice's device and encrypted with the IMS's public key, $\mathrm{Pk}_{\mathrm{IMS}}$, using a *deterministic encryption scheme* (we discuss why using a deterministic encryption scheme is acceptable in this case in Section 4.8). This is then sent, along with $hash(P_B)$, to the CMS. $hash(P_B)$ is used as an identifier for the sharing of the photo.

5. The CMS generates a random session id, *sid*, marks it as pending and stores and forwards the received encrypted face representation and *sid* directly to the IMS. As several photo uploads likely happen at the same time, the CMS acts as a mix network for the IMS (shuffle and delay of messages is also possible), preventing the IMS to link users to the same photo (details in Section 4.5).

6. The IMS decrypts the message to obtain the face representation and finds, among all the registered users, the one with the closest feature vector (*i.e.,* smallest Euclidean distance) that also satisfies a maximum acceptable similarity threshold. It then sends a message to this user (Bob), containing *sid* (*e.g.,* a link embedded in a notification shown in the app running on Bob's device, the app would be provided by the IMS–eg Facebook Messenger). Note that Bob does not need to have a CMS account. In the case of a missing identity–*e.g.,* Bob is not yet a registered user–the protocol stops here and his face remain cropped out.

7. Upon notification, the app running on Bob's device generates a pair of public/private session keys, Pk_{sid}, Sk_{sid} and sends Pk_{sid} and *sid* directly to the CMS.

8. The CMS forwards Pk_{sid} to Alice, who uses it to encrypt the part of the original photo containing Bob's face, as well as its position coordinates in the photo. It sends this encrypted information to the CMS.[5]

9. The CMS stores the received information and forwards to Bob the encrypted face and the coordinates, as well as the background image P_B and the corresponding *context.*

10. Bob's app recreates an image consisting of the P_B and the portion in which his face appears (which he decrypts using Sk_{sid}), shows it to Bob along with the *context* and uploader identity, and Bob decides whether to give consent for allowing his face to be visible in this photo. Note that this can also be automated through machine learning techniques (*e.g.,* [162–164]), or enforced through policies (*e.g.,* "accept all from friends", "accept if not nude", "accept if I am the uploader", *etc.*). Before presenting the photo to Bob for consent, spam filtering techniques can be performed (*e.g.,* using senders white/black lists or performing face detection on the face image to ensure this is really a face image belonging to Bob and not an unsolicited ad).

11a) If the decision is to allow Bob's face to appear in P_V, Sk_{sid} is returned to the CMS as a response to *sid.* At this point, the CMS can decrypt the stored face and the coordinates it has received from Alice and verify the validity of the coordinates (*e.g.,* by verifying that the corresponding area of P_B is cropped out) and that the feature representation obtained from this face image is identical to the one Alice sent to the IMS, through the CMS, in step (4). This is possible because the encryption

[5] Note that, as the CMS is assumed to be passive, no Man-in-the-Middle attack is possible.

scheme is deterministic. Note that, for the same person, the feature representation differs when the face image differs, therefore if two feature representations are identical, this guarantees that the original face images they were computed from are identical. If the validation is successful, the CMS adds Bob's face to the published photo P_V and marks this sid as accepted.

11b) If the decision is to *not* allow Bob's face to appear in P_V, a "deny" message is returned to the CMS as a response to sid, and the CMS then marks sid as denied and the area corresponding to Bob's face remains cropped out.

Note that, in the case of no response from a user, his face simply remains obfuscated. This action can also be configured by the CMS (e.g., default option after a timeout).

4.5 Security and Privacy Analysis

In this section, we demonstrate how ConsenShare satisfies the goals described in Section 4.2. Note that the security of some parts of the system directly depend on that of the underlying technologies used (*e.g.,* face recognition is not perfect [165]). We discuss these in detail in Section 4.8.

4.5.1 Effectiveness

The identity claim process using webcams at registration, prevents the creation of fake accounts. Face spoofing detection techniques (*e.g.,* [166]) can also be used to prevent malicious individuals from creating accounts on behalf of other users. Once registered, if a user's face appears in uploaded photos, he would either be asked for consent (with his face being encrypted in all communications and his identity hidden from the CMS) or his face would be blurred to begin with (depending on the uploader's choice). We discuss why malicious users cannot bypass this part of the protocol in detail in Section 4.5.3.

4.5.2 Privacy as Anonymity and Unlinkability

Regarding the data that is visible or known to the different parties throughout the protocol, we emphasize that face detection is performed locally on the uploader's device, that is, faces are not transmitted to the IMS (only face vectors) and the CMS only receives encrypted versions of the faces to forward to the consenters. None of the consenters involved in a photo have access to each other's faces. The IMS only has access to the background image (which does not contain any faces) and to the face representations of the people appearing in all the photos (which do not disclose anything about the sensitive information–the actual face). As the CMS acts as a mix network (it forwards a large number of messages to the IMS), the IMS cannot distinguish the lookup operations belonging to the same photo, thus it cannot link faces to faces or faces to a particular content. To make this property even stronger, the CMS can randomly mix and delay messages sent to the IMS (using buffering and shuffling), as well as add dummy messages in step (5) (Figure 4.3 on page 89). As for the CMS, it has access to the face-free background image. All the other data (face vectors and the faces) is sent encrypted to the CMS and only decrypted after the concerned users grant consent. The CMS is thus not able to identify the users that are involved in the same photo before they have given consent.

Furthermore, the CMS is also not able to link different faces (from different photos) of the same user, as face representations of a person always differ (even slightly) in each photo, making the encrypted version entirely different. We consider the case where the CMS wants to identify users in photos submitted in the future, based on their face representations from previous granted consents (step 11a, Figure 4.3 on page 89); the CMS would have to build a dictionary of possible face representations for a target user by adding noise at each position of the face representation array. This quickly becomes very expensive, *i.e.,* the time complexity is exponential in the size of the face representation array (*e.g.,* 128 positions in OpenFace [167]). Furthermore, we can easily protect against this attack with very little bandwidth and CPU time overhead by concatenating a random salt to the feature vector and to the face sent in step (4) and (9), respectively; In step (11), the CMS can retrieve the salt along with the face and perform the validation of the face representation. As for other similar timing, linking or side-channel attacks potentially performed by the IMS or by the CMS, these can also be deterred by traffic aggregation and randomization at the CMS and by adding dummy request at the client side (*e.g.,* the uploader's application can send more messages, to other users (in Step (4)), which would be automatically disregarded by the consenters' application).

4.5.3 Malicious User Behavior

A malicious uploader cannot bypass the system by leaving faces visible in the background image, as this would immediately be detected by the IMS in step (2) (Figure 4.3 on page 89). Malicious uploaders can also not bypass the system by sending an incorrect feature vector in step (4)–in order for someone else to provide consent in lieu of the legitimate target consenter–as this would be detected by the CMS in step (11a). Similarly, malicious users cannot bypass the system by providing a consenter with a face different than that sent to the CMS, as verifications of the face position in the photo and a comparison of the feature vector for that face and that sent to each consenter are performed in step (11a). Every message sent by an uploader, throughout the protocol, is validated by the CMS before any consenter's face is made visible to the target audience. Thus, malicious user behavior (even colluding users) results in sensitive parts of the photo not being shared with the target audience. Privacy of the sensitive content is also guaranteed, up to the point consent is granted by the concerned user. This is due to the security of the encryption schemes and the design of the system: The sensitive content is encrypted and not visible to the IMS, to the CMS, eavesdroppers or to any other users of the system.

4.5.4 Usability and Transparency

All consent requests contain contextual information for the consenter. However, in our solution, there is a chance of spamming attacks, *e.g.,* sending unwanted information to specific users (in the background of the photo, for instance). Although these can be annoying, we do not consider them extremely privacy invasive and well-known anti-spamming techniques (such as those used for e-mail) can be used by the CMS to handle this problem.

4.5.5 Collusion Cases

User-CMS and user-IMS collusions do not lead to additional information leakage. We discuss the case of CMS-IMS collusion (*e.g.,* the role of the CMS and that of the IMS are both played by Facebook). Typically, the CMS has information about the photos, but does not know anything about the identities of the faces appearing in them; the IMS can link every face with a user. In the case of collusion, the CMS and IMS entities would be able to link the users' identities to a particular photo, but the sensitive parts of the photos belonging to these users (their faces) would still not be visible unless these users grant their consent (as their decryption is only possible if consent is granted and both the CMS and IMS are honest-but-curious). Linkability could be reduced by adding dummy messages by the uploader application.

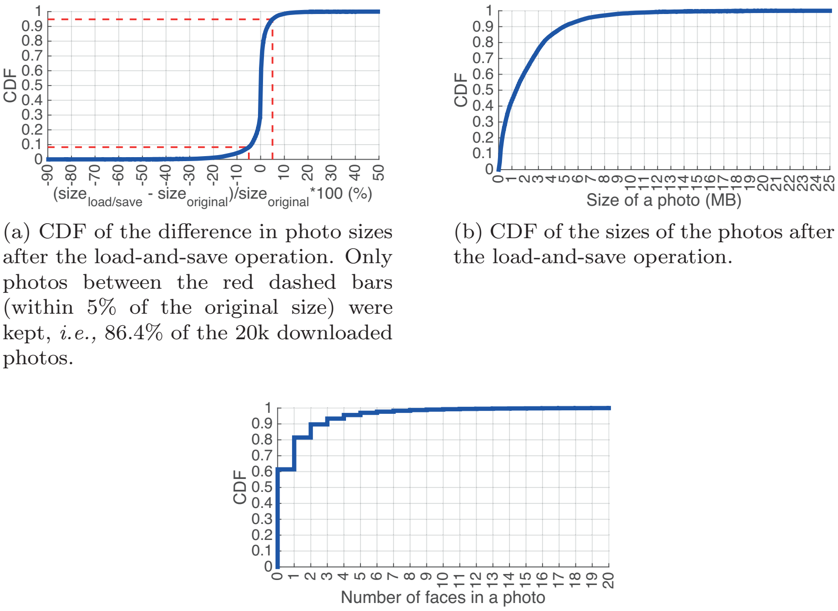
4.6 Implementation and Evaluation

To evaluate the performance of ConsenShare and demonstrate its practicality, we implemented a proof-of-concept prototype and evaluated its performance, in terms of CPU usage and bandwidth consumption, by relying on a real large photo dataset. We describe the implementation of our prototype, the dataset collection as well as our experimental setup, and we present the bandwidth and CPU consumption for the photo upload and grant consent operations.

4.6.1 Prototype Implementation

We implemented the prototype and carried out our performance evaluation by using Python 3.6. The prototype code and documentation are available at http://infoscience.epfl.ch/record/232563. Note that the prototype was not optimized; therefore, the CPU usage measurements should be considered as a loose upper bound. The prototype consists of the two server applications (the CMS and the IMS respectively), which we implemented with Flask, and a client application that supports the three main operations of the system: register (to the IMS), upload photo (to the CMS) and approve/deny consent (to the CMS). The servers use basic SQLite databases for local storage. All communication between these entities is achieved through JSON-based HTTPS requests and responses.

For face detection and feature-vector extraction, we use Dlib [168]–a C++ face detector–and OpenFace [167, 169] (v.0.2.1), which is an open source Python implementation of Google's FaceNet framework [160]. We implemented the lookup operation (*i.e.,* retrieving the best matching record for an input feature vector) to return the database record that minimizes the Euclidean distance with the input feature vector: We implemented it in a naive way, that is by comparing it to all the records in the database. Note that there exist efficient techniques for finding the most similar face, based on a low-dimensional representation (embedding), in databases of up to hundreds of million faces (*e.g.,* [170]). Basic image manipulation operations, such as loading and saving image files as well as extracting faces and replacing them with black rectangles, were performed with the Python Imaging Library (Pilow, v.4.1.1). In order to avoid discrepancies in the image file sizes (and therefore in the bandwidth measurements), we configured Pilow to retain all the parameters of the JPEG/JFIF format and encoders



(a) CDF of the difference in photo sizes after the load-and-save operation. Only photos between the red dashed bars (within 5% of the original size) were kept, *i.e.,* 86.4% of the 20k downloaded photos.

(b) CDF of the sizes of the photos after the load-and-save operation.

(c) CDF of the number of faces per picture, for pictures containing less or equal to 20 faces (99.7% of the photos.)

Figure 4.4: Dataset statistics.

(e.g., quality, color space, chrominance subsampling factor) from the original image processed, when saving (parts of) it.

For the basic cryptographic operations (hash, sign/verify, encrypt/decrypt and generate keys), we used the Python binding to the Networking and Cryptography library (PyNaCl [171], v.1.1.2). Specifically for sign/verify operations, we used the Ed25519 algorithm, with 128-bits security; for (cryptographic) hashing we used the SHA-256 algorithm with 128-bits security; for encryption, decryption and session keys generation we used the Curve25519 algorithm with 128-bits security (256-bits keys). We thus achieve a security of more than 112 bits, in compliance with the current NIST standards [172] for 2016-2030. For the simplicity of the implementation, we consider that the IMS generates the keys (Step (7)) and that no context is sent along with the photo (Steps (3) and (10)).

4.6.2 Dataset

We relied on the Yahoo Flickr Creative Commons 100 Million (YFCC100m) dataset [140] that contains the metadata of 100 million photos from the Flickr photo hosting website. More specifically, we extracted an unbiased sample of 20k photos (we drew photo IDs uniformly at random, without replacement, by using Python built-in random number generator with a seed of 0; we skipped the files that were no longer available). For each selected photo, we downloaded its full-resolution version from Flickr (if still available), and we filtered out the photos for which the size of the photo after a load and save operation differed from the original size by more than 5% (see Figure 4.4a). Our final dataset contained 17,257 photos and is available at http://infoscience.epfl.ch/record/232563.

We computed statistics related to the sizes of the photos and to the faces that appear on them (these are depicted in Table 4.1 on page 96). In our final dataset, the average number of faces in a photo is 1.0 (with a standard deviation of 3.7), and the maximum number of faces in a photo is 230. 61.2% of the photos do not contain any face, 20% of the photos contain exactly one face and 18.8% contain more than one face. Figure 4.4c illustrates the CDF of the number of faces per photo. Faces are generally small in size, compared to the actual photo: a face represents, on average, 1.1 ± 3.3 % of the photo size (with a maximum of 77.2 %), whereas the average size of all faces represents 3.0 ± 5.6 % of the photo size in MB. As for the sizes of the photos, there is substantial variation (as can be observed in Figure 4.4b), the average photo size is 2.1 ± 2.4 MB and the maximum is 25.2 MB.

	Min	Mean ± Std	Max
Number of faces in a photo	0	1.0 ± 3.7	230
Size of a photo (MB)	0.005	2.1 ± 2.4	25.2
Size of a face (MB)	0.0004	0.0 ± 0.1	10.7
$\frac{\text{Size of a face}}{\text{Size of the photo}} * 100(\%)$[6]	0.004	1.1 ± 3.3	77.2
$\frac{\text{Size of all faces}}{\text{Size of the photo}} * 100(\%)$[6]	0.007	3.0 ± 5.6	77.2

Table 4.1: Dataset statistics

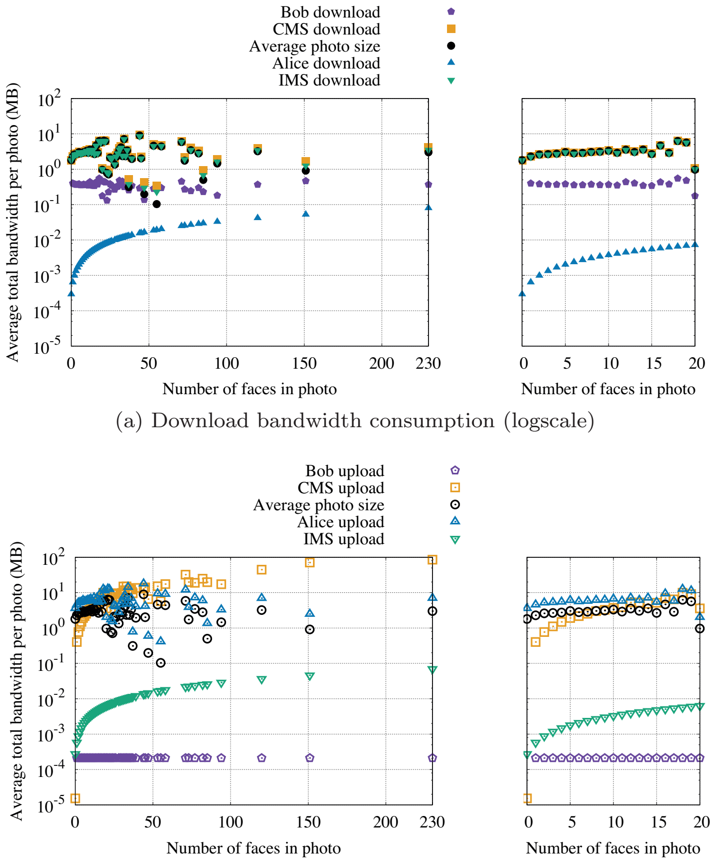
4.6.3 Experimental Set-Up

We evaluate the scenario where a user, Alice, wants to upload a photo in which potentially other people appear and we assume Alice wants all these people to appear in the photo–thus all faces are asked for consent. We do not consider the lookup operation in our evaluation (Step (6) (Figure 4.3 on page 89)). Hence, we generically refer to a consenter by the name of Bob, for photos that contain at least one face. We perform the photo upload and grant consent (for all appearing faces) operations for all the photos in our dataset, sequentially. In Step (10) (Figure 4.3 on page 89), we configured the CMS to provide Bob with a scaled version of the original background image with a maximum width of 1000 pixels, keeping the same image aspect ratio, quality and metadata as the original. We made use of one standard computer (Intel i7 CPU, 2.8 GHz, 8GB RAM) with Mac OS v.10.12.5. We did not use any optimization for Intel processors.

4.6.4 Experimental Results

We present here the bandwidth and CPU requirements of our system for the upload photo and grant consent (from all parties) operations. All the results that we present are upper bounds of the real bandwidth and CPU consumption, as face detection is done much faster, some of the users might be blurred out (and thus not asked for consent), and more advanced image transformation techniques can be used (such as JPEG transmorphing [173] to reduce bandwidth consumption).

[6] Computed only for photos containing at least one face.



(a) Download bandwidth consumption (logscale)

(b) Upload bandwidth consumption (logscale)

Figure 4.5: Average per-photo total bandwidth consumption in logscale for the uploader (Alice), the CMS, the IMS and the consenter (for the same photo, we consider the average bandwidth for one consenter, Bob). We illustrate these (y-axis) for different categories of photos, based on the number of faces they contain (x-axis). Note that for Bob, total upload and download bandwidth is 0 for photos that contain no face.

Bandwidth

We compute the average total bandwidth consumption in MB (on upload and on download) for one photo–for all the photos in our dataset, for each of the four entities: Alice, the CMS, the IMS, and one consenter, Bob. We refer to a baseline case where Alice directly uploads the photo to the CMS (providing *no privacy* for Bob). With respect to this baseline case, we compute the average *bandwidth overhead* (in MB) for one photo, which equals the total bandwidth from which the size of the original photo is subtracted (for Alice upload and CMS download) and, for the other cases, simply the total bandwidth. We also refer to the *relative bandwidth overhead*, expressed in percent, relative to the original photo size (dividing the bandwidth overhead by the original photo size). The average bandwidth requirements are presented in Table 4.2 on page 98. Notably, the average total bandwidth consumption for Alice on upload is 4.2±4.8MB (roughly twice the

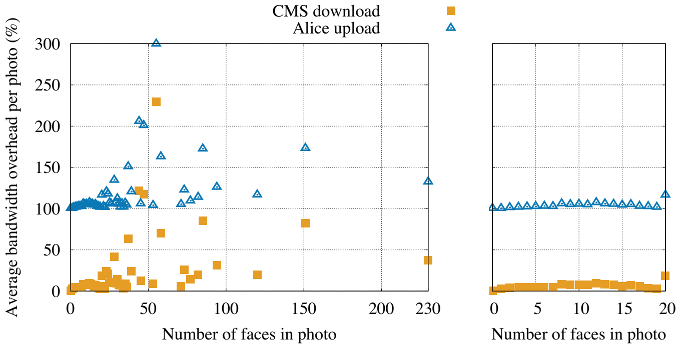


Figure 4.6: Average per-photo relative bandwidth overhead of the uploader (Alice) and the CMS (download). These are computed with respect to the baseline scenario where Alice uploads the original photo directly to the CMS and expressed as a percent of the original photo size. We illustrate these (y-axis) for different categories of photos, based on the number of faces they contain (x-axis).

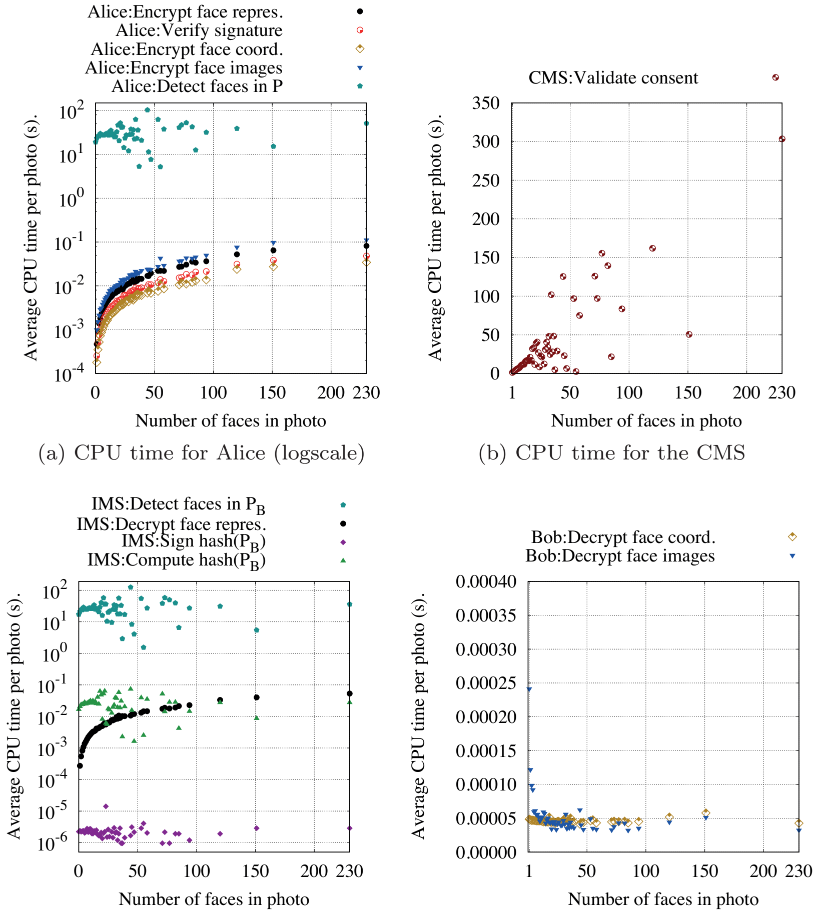
		Total bandwidth (MB)	Relative bandwidth overhead (%)
Alice (uploader)	Upload	4.2 ± 4.8	101.0 ± 4.7
	Download	0.0007 ± 0.0013	0.2 ± 0.4
CMS	Upload	0.4 ± 1.4	33.1 ± 135.0
	Download	2.1 ± 2.4	1.7 ± 5.4
IMS	Upload	0.0006 ± 0.001	0.1 ± 0.4
	Download	2.1 ± 2.4	99.8 ± 3.7
Bob (consenter)	Upload	0.0001 ± 0.0001	0.02 ± 0.07
	Download	0.2 ± 0.3	13.8 ± 26.8

Table 4.2: Average per-photo total bandwidth requirements and relative bandwidth overhead for the uploader (Alice), the CMS, the IMS and the consenter (for the same photo, we consider the average bandwidth for one consenter, Bob). We compute the bandwidth overhead relative to the baseline scenario where Alice uploads the original photo directly to the CMS.

original photo size – because she sends the background image twice[7]), the average total bandwidth consumption for the CMS and for the IMS on download is 2.1 ± 2.4MB and 2.1 ± 2.4MB, respectively (roughly the original photo size); the average total bandwidth consumption for the CMS on upload is 0.4 ± 1.4.[8] The other cases present negligible bandwidth consumption. Note that in a real system, the CMS upload cost could be substantially reduced by returning an even lower version of the background image to Bob. Figures 4.5 on page 97 and 4.6 on page 98 illustrate the total bandwidth consumptions (for all entities) and the relative bandwidth overheads (for Alice on upload and for the CMS on download), detailed for categories of photos containing a certain number of faces. Although there is some slight increase of bandwidth consumption w.r.t. to the

[7] Sending the background image to the CMS could be delegated to the IMS by providing the hash as opposed to the full P_B in step (3) (Figure 4.3 on page 89) and making the CMS request P_B from the IMS directly.

[8] While this may at first glance seem high, note that recent statistics reported that Flickr handles 1.68 million photo uploads per day, on average (this is a lower bound, as it only includes photos uploaded with public visibility) [174]. At an average photo size of 2MB, this means 6.4TB daily.



(a) CPU time for Alice (logscale)

(b) CPU time for the CMS

(c) CPU time for the IMS (logscale)

(d) CPU time for Bob

Figure 4.7: Average per-photo CPU times in seconds for different operations performed by the uploader, the CMS, the IMS and by a consenter (for the same photo, we consider the average time for one consenter, Bob). Note that the operations *Alice:Encrypt face repres.*, *Alice:Verify signature*, *Alice:Encrypt face coord.*, *Alice:Encrypt face images*, *CMS:Validate consent*, *IMS:Decrypt face repres.*, *Bob:Decrypt face coord.*, *Bob:Decrypt face images* take 0 CPU time when the photo contains no face.

number of faces in a photo (*e.g.,* for Alice on upload), this is negligible and most such increases can actually be due to an increase in photo size.

CPU time

We compute the average CPU time in seconds for one photo, for various operations, which we enumerate in Table 4.3 on page 100. Clearly the most expensive operation is the face detection performed by Alice (on the original photo) and by the IMS (on the slightly smaller sized background image for verification) with an average CPU time of 21.5 ± 21.3 s and 19.6 ± 23.6 s, respectively. We did not notice any pattern with the number of faces in the photo for these operations, but there is a noticeable increasing pattern with the photo size. Thus, a simple optimization of scaling down the photos when performing face-detection would drastically reduce this time, as shown in practice

	Operation	CPU time (s) $avg \pm stdev$
Alice (uploader)	Detect faces in P (1)	21.5 ± 21.3
	Encrypt face representations (4)	$4.7 \times 10^{-4} \pm 1.5 \times 10^{-3}$
	Verify signature (9)	$2.6 \times 10^{-4} \pm 8.7 \times 10^{-4}$
	Encrypt face coordinates (9)	$1.8 \times 10^{-4} \pm 6.2 \times 10^{-4}$
	Encrypt face images (9)	$7.5 \times 10^{-4} \pm 2.6 \times 10^{-3}$
CMS	Validate consent (11)	1.3 ± 5.4
IMS	Detect faces in P_B (2)	19.6 ± 23.6
	Compute $hash(P_B)$ (2)	0.02 ± 0.02
	Sign $hash(P_B)$ (2)	$2.3 \times 10^{-6} \pm 3.4 \times 10^{-6}$
	Decrypt face representations (5)	$2.8 \times 10^{-4} \pm 9.5 \times 10^{-4}$
Bob (consenter)	Decrypt face coordinates (10)	$1.9 \times 10^{-5} \pm 2.5 \times 10^{-5}$
	Decrypt face images (10)	$6.6 \times 10^{-5} \pm 5.0 \times 10^{-4}$

Table 4.3: Average per-photo CPU times in seconds for the uploader, the CMS, the IMS and a consenter (for the same photo, we consider the average time for one consenter, Bob).

(*e.g.,* Amos et al. [167] mention a run-time less than 0.1 s and Taigman et al. [175] mention a run-time of 1 s per photo, for images from the Labeled faces in the wild dataset [176]). The validate consent operation (Step (11) (Figure 4.3 on page 89)) – which includes face detection on all of the face photos in one photo for validation purposes and is performed by the CMS – takes, on average, 1.3 ± 5.4 s and is, as it can be seen in Figure 4.7, highly dependent on the number of faces appearing in the photo. However, even with 230 faces in a photo, this operation only takes 303 s (remember that we did not use any CPU optimizations and, in practice, face recognition operations are already performed much faster by app/services like Facebook, even on phones). The CPU times for other operations are negligible.

We conclude that these results are acceptable and demonstrate the effectiveness of a system like ConsenShare.

4.7 Incentives and Adoption

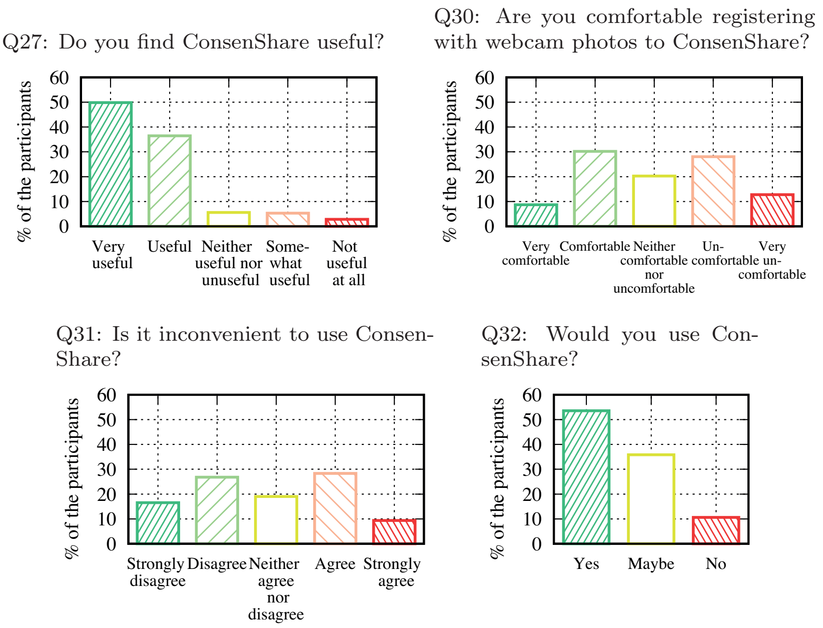
We present here the results of our survey and discuss the incentives for adoption by the different stakeholders.

4.7.1 Survey

In order to gain insight into the individuals' perceptions of Multiple-Subject/Interdependent Personal Data (MSPD/IPD) (and of the associated privacy risks) and of a ConsenShare-like system, we conducted a survey targeted at Facebook users.

Methodology

We conducted our survey in mid-2017. We recruited participants through the Amazon Mechanical Turk (AMT) 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



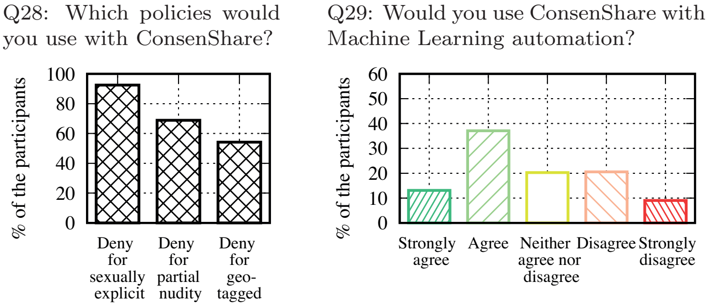
Q27: Do you find ConsenShare useful?

Q30: Are you comfortable registering with webcam photos to ConsenShare?

Q31: Is it inconvenient to use ConsenShare?

Q32: Would you use ConsenShare?

Figure 4.8: Participants' responses to ConsenShare-related survey questions.



Q28: Which policies would you use with ConsenShare?

Q29: Would you use ConsenShare with Machine Learning automation?

Figure 4.9: Participants' responses to survey questions regarding ConsenShare automation.

HITs and an active Facebook account (AMT offers the possibility to specify this admission criterion). The survey took approximately 10 minutes to complete (median completion time of 8m48s) and each participant received a financial compensation of 3 USD in exchange for their participation. The survey was approved by our institution's ethics committee/institutional review board (application #006-2017/18.05.2017).

The survey was structured as follows. After the standard demographic questions (Part I), we polled the participants about their perception of online data privacy for different data types and about their experience with discrimination causes by data available online (Part II). We polled the participants about their sharing behaviors and those of their friends regarding multimedia content on OSNs; some of the questions were specific to sexually explicit content (Part III). We polled the participants about their (un)tagging (face tags, "with tags", @ tags) behaviors and those of their friends on Facebook (Part IV); the survey questions included screenshots from the Facebook website to illustrate the aforementioned tagging features. After a brief high-level description of ConsenShare, we polled the participants about their perceptions of a ConsenShare-like system. In particular, we polled them about their willingness to use such a system (Part V), with a special emphasis on the consent decision process (manual, policy-based, machine learning-based). Finally, we asked the participants to confirm their agreement to save their responses and to use them in a scientific publication. The survey contained two duplicated questions in order to check the participants' attention; we used these to exclude the responses from inattentive participants from our dataset. The complete transcript of the survey and the anonymized and sanitized answers are available at http://infoscience.epfl.ch/record/232563.

Results

We obtained a total of 536 complete responses. We ruled out duplicates (*i.e.,* when a participant completed the survey multiple times), the responses from inattentive participants (*i.e.,* when a participant's responses to the duplicated questions were inconsistent) and the participants that chose not to allow us to save their answers. This left us with 321 complete responses. The corresponding participant sample was rather balanced and diverse in terms of the participants' demographics: 53.0% of the participants were female, the participants had various areas of employments, and their ages ranged from 20 to 75 years old, with an average of 35.3 ±10.4.

Our survey results indicate a potential user concern regarding the sharing of location data (86.3% of the participants), multimedia data (69.5% of the participants) and genomic data (60.8% of the participants). Furthermore, 10.3% of the participants claimed that they were victims of discrimination or prejudice based on online content about them. Of these, 33.3% reported that this happened more than once in the past. A staggering 66.7% reported that the cause was content shared by others, which highlights the gravity of MSPD/IPD privacy risks. As for the most common domains in which the discrimination or prejudice happened, 60.7% of the users referred to a job application, *e.g.,*

> "They checked my profile online before the interview and he started making uncomfortable jokes about the game pages I like on it and a party photo." (Male, 28)
>
> "My employer watches my online activity and asked about that later" (Male, 26)

30.3% to familial or social situations, *e.g.,*

> "My marriage proposal got cancelled" (Male, 25)

27.3% to professional situations, 15.2% to loan or mortgage and 9.1% to insurance premiums.

Regarding sharing multimedia content online, 60.1% of the participants reported that they share such content at least occasionally (a few times/month). 48.6% of the participants reported that this content contains faces of people other than themselves at least half of the times, whereas 41.4% declared that this happens sometimes, but less than half of the times. Only 10.0% said their multimedia content never features faces of others. Participants reported that their friends also share multimedia content about them at least occasionally (a few times/month) – for 45.5% of the participants.

Regarding revenge pornography, 4.1% of the participants declared that they were victims of revenge porn in the past.[9] We also polled participants about whether other people have or had access to explicit photos of themselves. 27.4% of them declared that this is the case; of these, 48.9% declared that the person who has the photos took them with a device of their own, whereas 8.0% declared that a third party took and shared the photos. Asked whether they have or had explicit photos of someone else, 40.2% participants responded positively; of these, 41.9% reported that they are the ones that took these photos, whereas 16.3% said that a third party shared these photos with them. Many participants explained, in comments, that the photos were taken in the scenario of a (former) relationship. This illustrates that the number of potential victims of revenge pornography might be quite high.

We also polled participants about their Facebook behavior. Asked how they tag or mention their friends when posting photos or videos in which they appear, a staggering 41.1% declared that they do not tag or link their friend's profile in any way (in other words, the friend can be completely unaware of the content posted and would thus not be able to remove/report the content). 7.2% of the participants said their friends ask them to remove photos that they have shared, at least occasionally and 11.2% said they noticed their friends contact Facebook about removing this content at least occasionally. 30.5% of the participants declared that they also ask their friends to remove content that they posted and 16.9% declared that they asked Facebook to remove such content.

Finally, we presented our framework to the participants. Aggregated participants' responses for this section of the survey are illustrated in Figures 4.8 on page 101 and 4.9 on page 101. Asked whether they find ConsenShare useful, 36.5% of the participants answered that this would be useful and 49.8% very useful *e.g.,*

> "I think this is a great way of giving control to individuals." (Male, 37)
>
> "I think it's best for all parties involved. " (Female, 32)
>
> "It would be nice to be asked if you wanted to be on the internet first, rather than letting any Tom, Dick, and Harry take your photo and post it[$\cdots$]" (Female, 31)

Interestingly, some participants even reported that they would be more comfortable with social platforms if such a solution was in place. Regarding the use of policies, only 3.4% of the participants declared that they would not use any policy, whereas 54.2% declared they would use a policy to deny consent for photos containing location information and 92.5% declared that they would use a policy to deny consent for photos containing explicit content. As for the use of automated decision making (via machine learning, for example), 20.2% of the participants were not sure that they would use this, whereas 50.2% declared that they would be in favor of using such a feature. Asked how comfortable they would be with a registration process similar to the one of ConsenShare (where a few photos would be required for registration), 38.9% of the participants reported being comfortable

[9] This result is in keeping with other studies, such as a report from the Data & Society Research Institute and the Center for Innovative Public Health Research [177], which found that 4% of U.S. internet users have been threatened with or experienced revenge porn.

or very comfortable and 20.3% reported being undecided about this. Of the remaining participants (who reported being uncomfortable or very uncomfortable with this), some commented that this is because they do not trust webcams in general or because this system is not provided by a known company. Some also seemed confused about how the system would work, not understanding that these photos would not be stored or that face recognition is something that Facebook (the IMS here) already does on their photos. We thus recommend taking the results for this particular question with a grain of salt.

In terms of the (in)convenience of first sharing the background photo and sanitized versions of the faces of the other people appearing in the photo, 43.3% of the participants found this convenient, *e.g.,*

> "It would not be inconvenient at all! It sounds like a great solution! You still get your photo but your friends get to decide whether or not to have their faces visible – a win win!" (Female, 33)

whereas 19.0% were undecided. Finally, asked whether they would overall use the ConsenShare system, 53.6% of the participants answered positively, *e.g.,*

> "It might seem inconvenient but I still favor this system and reasoning behind it." (Female, 32)

> "I think this could save a lot of head aches and keep people from being upset with one another over photo postings. This would allow anything that a person did not want online to be taken care of before it made it to a facebook page." (Male, 42)

> "I would definitely be interested in using this feature and really hope that you can develop it and get it out there. it would make me feel much more comfortable using social media and sharing images online." (Female, 34)

and 35.8% said that they would perhaps consider using the system.

While there are limitations to our survey and future work we envision on this–we discuss these in Section 4.8–the results indicate that a system like ConsenShare could be needed and there is a potential desire from the users to adopt it.

4.7.2 Adoption

As we saw from our survey, a system such as ConsenShare would involve some tradeoff between the user experience as an uploader (waiting for friends to give consent before the full photo is visible), as well as the data that must be provided for registration,[10] and her experience as a consenter (whose privacy would be much better protected). However, giving back control to the users represents a substantial incentive for adoption on their part, which is further enhanced by the fact that the system is transparent and lightweight. Regarding the stakeholders, although adoption would come with some costs (*e.g.,* , increased bandwidth, deploying the infrastructure[11]), a major incentive for

[10] Note that in the case where an existing system is used to play the role of the IMS (such as Facebook), the users' faces are already registered, making adoption quite straightforward for the users.

[11] For instance, the cost of adding the consent feature to a CMS such as Flickr would be minimal compared to the existing infrastructure: A simple interface with an IMS (*e.g.,* Facebook) and basic cryptography and image processing operations. In other words, the individual ConsenShare operations are not much different from what current CMS platforms are already doing.

adoption by the CMS could be following new trends (*e.g.,* the fact that such consent-based mechanisms for MSPD/IPD data might become law-enforced) and avoiding lawsuits. Furthermore, as such features are evidently desired by their users, including them would be good for reputation, providing a competitive advantage and likely increase the user base (and thus also the ads revenue), as well as the shared data. This also creates new business opportunities for the CMS providers that implement the ConsenShare in-house solution and sell it to other CMS providers. As for the IMS, the advantage would be a business-to-business arrangement with the various CMS providers (either transactional or subscription-based), an increased user base (thus revenue) and data (similar to that by the Facebook Connect feature). Furthermore, the IMS could monetize such features by providing them to the users in exchange for a premium fee. Note that existing services with large user databases including faces and relationships (typically OSNs) are perfect candidates to play this role, as they already have most of the data and technology needed. For instance, Facebook already performs face recognition in the background and could offer this service as an IMS to different CMSs.

4.8 Discussion: Limitations and Extension

We discuss here the limitations and extensions of our work. It should be noted that this work represents a first step towards proposing a privacy-preserving generic framework for sharing MSPD/IPD data. First, it is worth mentioning that there is an inherent tradeoff between the right to privacy and the right to freedom of speech. A possible middle-ground option, in the case of photos, would be to instantly publish critical content – such as photos of a mass civil action – on CMS platforms with blurred faces (similar to Google Street View and to many media outlets that already protect the identity of certain individuals, *e.g.,* minors, by blurring their faces in pictures and videos). As for the other side of the coin, the right to privacy is subject to debate for public figures/celebrities [178, 179]; such individuals could be detected using, for instance, Facebook's verified accounts feature, and their faces could be automatically posted, without the need for consent. Second, in the case of pictures, detection is not perfect. There is still a small margin of false positives (*i.e.,* detecting a face when none exists; if recognition matches such a "face" to a user, he can report this to the IMS, who can then improve its models upon checking the validity of this request) and false negatives (*i.e.,* not recognizing a face, which pose more problems from the privacy point of view, as these imply that a user appearing in a picture would not be recognized. Such a detected but unidentified face could be blurred out by default). These can be alleviated by asking the uploader for manual input in detecting users in the picture (similar to tagging on Facebook). Third, our current solution is centralized. In future work, we plan to design a decentralized P2P solution for the IMS and potentially the CMS. Fourth, we intend to extend our solution to incorporate interaction among users, providing them with the automatic tools to consider different options of sharing (*e.g.,* different obfuscation mechanisms, different target audiences, *etc.*) and iteratively achieve a consensus – thus automating the social ad-hoc mechanisms users reportedly use today [80]. Fifth, the solution is CMS-specific, which means an individual user could upload the content on a different platform and just post a link to that content, by-passing the need for consent; however, this would have less impact and could even be blocked by the CMS, *e.g.,* blocking links to dubious websites. Sixth, as the issue of balance of control and the "ideal" privacy settings are culturally-dependent and not entirely law-enforced in the case of MSPD/IPD data (and regulations can differ depending on the country), our survey sample of participants is not necessarily representative of the global population; the vast majority of Mechanical Turk workers is reported to be US-based and they might have an IT-experience higher than the average; previous works have studied the profiles of Mechanical Turk workers [180]. Furthermore, it is possible that the participants' answers do not accurately indicate their true attitudes for adoption, as users' privacy attitudes and privacy decisions are not always rationally connected [87] and reported behaviors do not necessarily match the natural behavior [181]. For a more rigorous assessment of the usability and adoption potential of ConsenShare and the users' perception on the different design alternatives, in future work, we intend to run additional surveys using a fully-functional prototype, making use of the SeBIS intention scale [182, 183] to gain more insight into the participants' expertise and following specific guidelines for designing privacy/security surveys [184]. Finally, we discuss how the main building blocks in our framework can be adapted to other data, beyond photos. Note that for any type of data, there are several options to consider in the design, such as remove vs. obfuscate the data (and the available granularity); in what follows, we discuss some of the alternatives.

The Case of Audio and Video. Considering audio and video data in our framework is a rather straightforward extension from our picture solution. Different solutions for identifying users in audio/video content have been proposed [185, 186], and various options can be used for separating the sensitive content (portions of the video in which a user appears) from the non-sensitive content: entirely cutting out the audio/video sections in which a user appears or altering the content of those sections to obfuscate that user. The privacy tradeoff of such solutions would be the temporal discontinuity.

The Case of Genomic Data. Genomic privacy is a complex subject whose discussion involves many ethical and balance of control issues and closely ties to that of the privacy of others versus personal freedom. The topics of *who* are the affected parties, *how* their consent decisions should be expressed, as well as that of genomic data credibility are still under debate both in the media (*e.g.,* [187]) and in the research community (*e.g.,* [12, 188–190]). There are several options that could be considered; we discuss how their implementation could be done in ConsenShare. Identity claim (at registration) in the case of genomic data would require formal identity proof (*e.g.,* an ID) and reporting of familial relationships. Detection of the involved individuals comes down to knowing these familial relationships (*e.g.,* through the IMS; Facebook already offers that option) and selecting the close relatives of the uploader. The degree of closeness for which individuals are considered as affected parties – and thus should grant consent – would be a configurable parameter of the system, which can be set according to the applicable regulation. In the most restrictive possible form of regulation, consent would be binary (yes/no) and a user would be allowed to share her whole genome only if all affected parties say "yes".[12]

[12] Note that in this case, the CMS/IMS do not even need to see the data to determine the involved individuals. In less restrictive regulatory options, consent can be refined to allow publication with a level of noise added to the full genome (*e.g.,* through differential privacy [191]) or after applying obfuscation techniques – typically at the SNP level – that would guarantee a certain level of privacy to the affected relatives, while allowing the user who wants to share his genome some freedom. In this case, the individual desired level of privacy would be configurable for each user and the consent decision of an affected user would be the level of noise/obfuscation that the uploader must apply to his genome; the

A unique limitation for genomic data is the fact that affected users (*i.e.,* unborn future relatives, such as children) can appear after the user has already shared his genome with proper consent from his relatives at that time. In this case, we can offer the possibility of revoking consent, which would force the uploader to remove the data (a less than perfect solution, as the data might have already been duplicated).

The Case of Co-location Data. Co-location data can be shared online by different means, for instance, by posting (and tagging) pictures or videos in which multiple people appear or directly tagging them in a post message. In the case of co-locations shared by using multimedia data, detection can be done as described above. In the case where co-located users are directly tagged by the uploader, detection is, obviously, no longer needed. The context provided to a consenter in the case of co-location data could also include an estimation of the location privacy loss stemming from that reported co-location, as quantified through the framework that we propose in Chapter 2. However, as co-locations introduce dependencies among the data of different users, once a co-location between Alice and Bob is shared, Alice's future location posts would also affect Bob's location privacy and vice versa. Hence the system should consider a window of influence of the correlation, by including an adjustable parameter for each user, specifying how much time after a shared co-location in which he is involved are other users required to ask for his consent to share location data. The temporal constraint could be coupled with a constraint specifying an allowed granularity level of obfuscation (*e.g.,* Bob wants to be asked for consent if, for the next hour after allowing Alice to share their co-location, she wants to report her exact location, but she will not require his consent if she only reports her location within an area of 1km or more of her real location).

4.9 Related Work

Consensual and privacy-preserving sharing of multi-subject and interdependent data online is a multi-faceted problem, as advocated by Good [192]: it includes legal, social and technological dimensions. In this section, we survey the related work in these dimensions, beginning with the legal aspects.

The notion of individual control over information about oneself and more specifically that of consent for information disclosure is currently the basis of most definitions of privacy, including Nissembaum's contextual integrity [132], terms of use, and data protection and privacy laws in most countries [193]. This is the case for the Consumer Privacy Bill of Rights [194], adopted by the US White House, and the General Data Protection Regulation (GDPR) (Regulation EU 2016/679), recently adopted by the EU Parliament.

Although the general case of MSPD/IPD is not explicitly addressed by current laws, because of its complexity, it is mentioned in multiple places. For instance, in case law [195], the court found that "an individual's personal autonomy makes him master of all those facts about his own identity, such as his name, health, sexuality, ethnicity, his own image [. . .] and *also of the 'zone of interaction' [...] between himself and others*". In addition, Opinion 5/2009 on OSN produced by the Working Party on Data Protec-

most restrictive of these – among all the relatives – would be applied to the uploader's genome before sharing. An individual's genomic privacy can be quantified using dedicated frameworks such as that proposed by Humbert et al. [12].

tion mentions the case of online social networks (OSN) users uploading data about other individuals, possibly not members of the OSN.

In the context of MSPD/IPD, specific data types received particular attention: photos, in light of the right to one's own image, genomic data [187], and more recently, photos and videos containing sexually-explicit content, namely revenge pornography, against which laws have been passed in Canada, France, Israel, Japan, the United Kingdom and in several states in the US (to name a few). In addition, online service providers, including Reddit [134], Facebook [135], and Twitter [137] have also reacted to this new trend and updated their terms of use accordingly. Furthermore, deepfake technology [196] has recently been used to create revenge pornography, a situation to which several service providers responded by updating their terms of use [197–199]. Yet, neither the laws nor terms of use are self-enforcing, and technical solutions are therefore needed.

Online services recently began including features to cope with content uploaded without the consent of some of the individuals whose privacy is affected by it, typically for photos. Facebook, for instance, enables its users to report such content and to remove references to their identities attached to shared content. However, such features still suffer from the following problems: (1) Individuals cannot automatically detect that content having privacy implications for them has been shared, unless an explicit reference to their identities is attached to it. (2) Even though the content is eventually removed, the damage, in terms of privacy, is done as the service provider and possibly some users have seen the content.

The need for and the design of collaborative privacy schemes for MSPD/IPD is an active topic in the literature. Gnesi et al. [26] introduce the notion of MSPD as data that contains identifiers that refer to more than one person, as is the case of pictures, phone records, co-locations or medical reports. They also discuss a technical solution for protecting MSPD based on user-defined privacy policies, however, it does not guarantee any protection against the service provider.

Aditya et al. [200] propose an image capture platform that also provides privacy protection by a combination of short-range wireless signaling to disseminate privacy policies. In the context of OSNs, Such et al. [79, 80, 123] study the so-called multi-party privacy conflicts (MPC), notably in the case of pictures. They identify the sources of such conflicts and the different non-technical strategies used by users to cope with them, including avoidance or individual/collaborative resolution. Collaborative privacy policy enforcement solutions were also proposed by Beato et al. [201] (based on secret sharing), by Squicciarini et al. [202, 203] (based on the Clarke-Tax mechanism from game theory), by Ratikan et al. [204] (based on majority voting), as well as by Hu et al. [205–208] (based on access control). Again, in contrast to our work, these solutions assume a trusted model for the service provider (OSN). Ilia et al. [209] proposes a collaborative multi-party access control model for OSNs where the service provider is considered honest-but-curious. However, it assumes that the data uploader is honest, yet privacy careless. Our work, on the contrary, assumes that the data uploader and other users could be malicious. Finally, collaborative privacy policies mechanisms have been proposed in other contexts such as sharing photos through instant messaging platforms [210] and personal data stored on others' devices (e.g., phone numbers) [211].

Researchers also proposed mechanisms to defend against untrusted providers in OSNs. De Cristofaro et al. [212] propose a privacy-enhanced alternative to microblogging OSNs such as Twitter. Their solution protects posts' contents, hashtags and follower interests from the service provider. Ion et al. [213] describe a privacy-enhancing mechanism that enables users to share data over any web-based OSN and provides confidentiality against unauthorized parties, including the service provider. Feldman et al. [214] propose a framework for OSNs that provides not only confidentiality guarantees, but also integrity protection (*e.g.,* against equivocation attacks) against an untrusted service provider. Secure JPEG techniques [215] can also be used to hide part of the photo content from the service provider. Although these works offer different levels of protection against an untrusted service provider, they do not offer mechanisms for detecting/identifying the individuals involved in the content and for implementing collaborative privacy policies for MSPD/IPD.

Identifying individuals involved in content shared online is a difficult problem. In several cases, including the case of pictures, such identification comes down to a classification problem. For instance, machine learning techniques can be used for detecting faces on encrypted images [155], [156]. Moreover, Ziad et al. [216] describe the use of homomorphic encryption for performing general image-processing operations (*e.g.,* spatial filtering, anti-aliasing) on remote (untrusted) servers in a privacy-preserving way; to use such an approach in our framework, certification of results would also be needed (*e.g.,* through blind signatures). Closer to our work, He et al. [217] describe a system for partial image sharing in OSNs that enables data owners to define private regions in an image, support for popular image transformations and set different privacy policies for each user associated with an image. In the same area, Ilia et al. [37] propose a fine-grained access control mechanism that enables users associated with an image to restrict the exposure of their own faces; this approach handles multi-party privacy policies conflicts and is compatible with existing access control mechanisms. These works, however, focus only on the problem of sharing images in OSNs. Our work, in contrast, focuses on different MSPD/IPD types, not only images, and deals with the problem of detecting involved individuals in a privacy-preserving way.

4.10 Conclusion

In this chapter, we propose ConsenShare, a generic framework for sharing MSPD/IPD data (*e.g.,* photos, videos, genomic data, *etc.*) with consent from all the involved individuals. ConsenShare is privacy-preserving by design not only with respect to other users of the system, but also with respect to the service providers. We implement and evaluate ConsenShare in the case of photos and show that it is technically possible to provide users control in the sharing of photos in which they appear, while ensuring their privacy and preserving the main features of existing sharing platforms. In doing so, our work lays the foundation for the design of privacy-preserving sharing of MSPD/IPD data.

Survey Transcript

Part I: Demographics

1. What is your gender?
   - ○ Male
   - ○ Female
   - ○ I prefer not to answer
2. What is your age? ▭
3. What is your primary area of employment?
   - ○ Homemaker
   - ○ . . .
   - ○ Retired
   - ○ Other ▭

Part II: Privacy and discrimination

4. From a *privacy* point of view, *do you care* about being associated with the following data shared online?
   - □ Textual data, such as status updates on Online Social Networks
   - □ Location data, such as check-ins on Online Social Networks
   - □ Multimedia data, such as posts containing pictures or videos on Online Social Networks
   - □ Genomic data of your relatives shared on specific platforms (e.g., 23andme)
   - □ Other ▭
5. Do *you* feel that you have been the *victim of discrimination* or prejudice (e.g., for a job application, insurance premiums, social or professional situations, loan application, etc) based on *content about you online* (e.g., pictures or videos)?
   - ○ Yes, more than once
   - ○ Yes, once
   - ○ No, never
6. [Conditioned on the positive response to Q5] Do you feel that the *discrimination or prejudice* was caused by *content* that *you shared or someone else shared*?
   - ○ Myself
   - ○ Someone else
   - ○ Both
7. [Conditioned on the positive response to Q5] In which of the following domains did *you experience discrimination or prejudice* caused by content shared online? (We would appreciate any details about the context of the stories that you are comfortable sharing).
   - □ A job application ▭

- □ Professional situations □
- □ Familial or social situations □
- □ Insurance premiums (health, car, etc.) □
- □ Loan application (including mortgage) □
- □ Other □

8. From a *privacy* point of view, *do you care* about being associated with the following data shared online?
   - □ Textual data, such as status updates on Online Social Networks
   - □ Location data, such as check-ins on Online Social Networks
   - □ Multimedia data, such as posts containing pictures or videos on Online Social Networks
   - □ Genomic data of your relatives shared on specific platforms (e.g., 23andme)
   - □ Other □

Part III: Multimedia data online

9. On average, how often do *you share* pictures or videos online?
   - ○ Very often (Almost daily)
   - ○ Somewhat often (A few times / week)
   - ○ Occasionally (A few times / month)
   - ○ Infrequently (Less than a few times / month)
   - ○ Never
10. On average, what proportion of the pictures and videos that *you share* online contains *faces of people other than yourself*?
   - ○ None
   - ○ Few of them (some, but less than half)
   - ○ About half
   - ○ Most of them (more than half, but not all)
   - ○ All
11. On average, how often do *your friends share* online pictures or videos on which *you appear*?
   - ○ Very often (Almost daily)
   - ○ Somewhat often (A few times / week)
   - ○ Occasionally (A few times / month)
   - ○ Infrequently (Less than a few times / month)
   - ○ Never
12. Does anyone (or did in the past) have *sexually explicit images* or videos of yourself, either because they took them directly, or because they were shared with them?

○ Yes
○ No
○ I prefer not to answer

13. [Conditioned on the positive response to Q12] How did this(these) person(s) obtain *sexually explicit pictures or videos of you*?

□ They took them ▭

□ I took them and shared them ▭

□ Someone else took them and shared them ▭

□ I do not know ▭

14. Did *you* ever *have* access to *sexually explicit images or videos of someone else*, either because you took them directly or because they were shared with you?

○ Yes
○ No
○ I prefer not to answer

15. [Conditioned on the positive response to Q14] How did you obtain sexually explicit pictures or videos of someone else?

□ I took them ▭

□ They took them and shared them with me ▭

□ Someone else took them and shared them with me ▭

□ I do not remember ▭

16. Have *you* ever been the *victim of revenge porn*[13] (e.g., someone sharing sexually explicit pictures of you without your consent)?

○ Yes
○ No
○ I prefer not to answer

17. On average, how often do *you share* pictures or videos online?

○ Very often (Almost daily)
○ Somewhat often (A few times / week)
○ Occasionally (A few times / month)
○ Infrequently (Less than a few times / month)

[13] Revenge porn (sometimes lengthened to revenge pornography) is the sexually explicit portrayal of one or more people that is distributed without their consent via any medium. The sexually explicit images or video may be made by a partner of an intimate relationship with the knowledge and consent of the subject, or it may be made without his or her knowledge. The possession of the material may be used by the perpetrators to blackmail the subjects into performing other sex acts, to coerce them into continuing the relationship, or to punish them for ending the relationship. Halder and Jaishankar (2013) define Revenge porn as: “an act whereby the perpetrator satisfies his anger and frustration for a broken relationship through publicizing false, sexually provocative portrayal of his / her victim, by misusing the information that he may have known naturally and that he may have stored in his personal computer, or may have been conveyed to his electronic device by the victim herself, or may have been stored in the device with the consent of the victim herself; and which may essentially have been done to publicly defame the victim”. Source: https://en.wikipedia.org/wiki/Revenge_porn

○ Never

Part IV: Social networks usage

18. Do you currently have a Facebook account?
Please note that we will not ask you any questions revealing personal identifiable information. Facebook usage related questions are only for screening purposes.

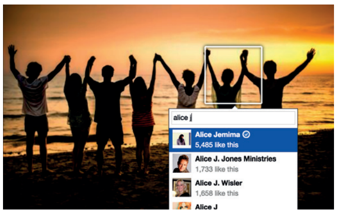
○ Yes
○ No

19. [Conditioned on the positive response to Q18] How often do *you tag or mention friends* on Facebook (either in a status or on a picture)?

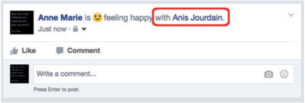
○ Very often (Almost on every one of my posts)
○ Somewhat often (On more than half of my posts)
○ Occasionally (On less than half of my posts)
○ Infrequently (Hardly ever or on very few of my posts)

20. [Conditioned on the positive response to Q18] When *you add content regarding others* on Facebook (e.g., a picture or video in which they appear), which of the following methods do you use?

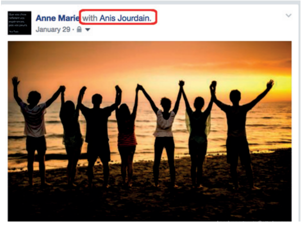
□ I tag their face (e.g., in pictures)



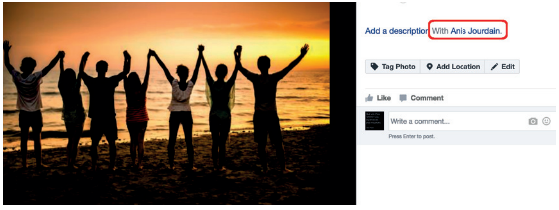
□ I use the “with” tags in the status or in the picture description


Anne Marie is feeling happy with Anis Jourdain.
Just now ·

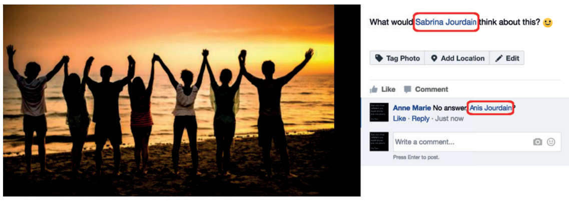
or


Anne Marie with Anis Jourdain.
January 29 ·

or


Add a descriptior With Anis Jourdain.

□ I link their name (e.g., by using "@") in a status, in a picture description or in comments



What would Sabrina Jourdain think about this? 🙂

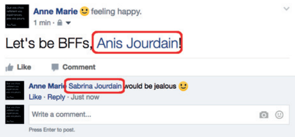
Anne Marie No answer, Anis Jourdain?

Like · Reply · Just now

Write a comment...

Press Enter to post.

or



Anne Marie 🙂 feeling happy.

1 min ·

Let's be BFFs, Anis Jourdain!

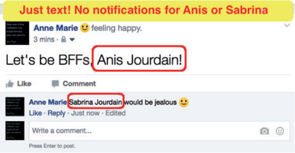
Anne Marie Sabrina Jourdain would be jealous 🙂

Like · Reply · Just now

Write a comment...

Press Enter to post.

□ I might not tag at all or link others' names when I post content about them



Just text! No notifications for Anis or Sabrina

Anne Marie 🙂 feeling happy.

3 mins ·

Let's be BFFs, Anis Jourdain!

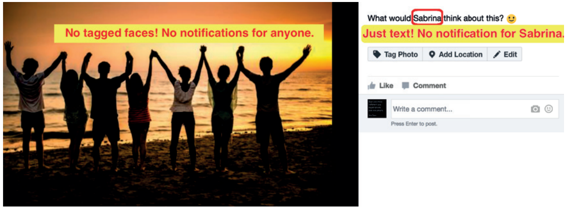
Anne Marie Sabrina Jourdain would be jealous 🙂

Like · Reply · Just now · Edited

Write a comment...

Press Enter to post.

or



No tagged faces! No notifications for anyone.

What would Sabrina think about this? 🙂

Just text! No notification for Sabrina.

Write a comment...

Press Enter to post.

21. [Conditioned on the positive response to Q18] How often do *your friends ask you to remove* status updates and/or pictures that you have shared of them on Facebook?

    ○ Very often (Almost every time)
    ○ Somewhat often (More than half the times)
    ○ Occasionally (Less than half the times)
    ○ Infrequently (Hardly ever or extremely rarely)

22. [Conditioned on the positive response to Q18] How often have you noticed *your friends untag themselves* from status updates and/or photos in which you had tagged them ?

    ○ Very often (Almost every time)
    ○ Somewhat often (More than half the times)
    ○ Occasionally (Less than half the times)
    ○ Infrequently (Hardly ever or very rarely)

23. [Conditioned on the positive response to Q18] How often does *someone else tag you* in a status update or a picture on Facebook?

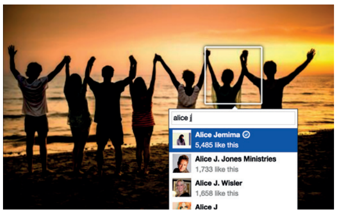
    ○ Very often (Almost daily)
    ○ Somewhat often (A few / week)

○ Occasionally (A few / month)

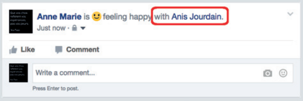
○ Infrequently (Less than a few / month)

24. [Conditioned on the positive response to Q18] When *others add content regarding you* on Facebook (e.g., a picture or video in which you appear), which of the following methods do they use?

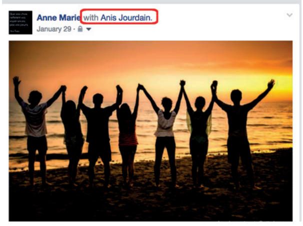
- □ They tag my face (e.g., in pictures)



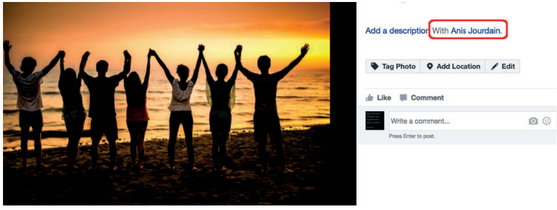
- □ They use the “with” tags in the status or in a picture description


Anne Marie is 😃 feeling happy with Anis Jourdain.
Just now ·

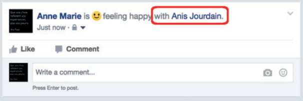
or


Anne Marie with Anis Jourdain.
January 29 ·

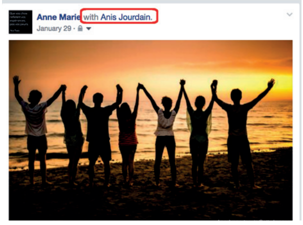
or


Add a description With Anis Jourdain.

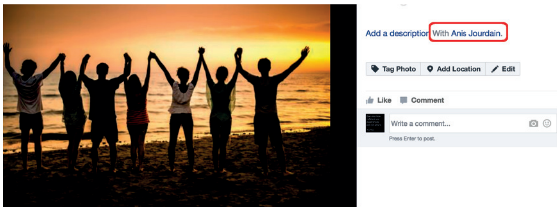
- □ They link my name (e.g., by using “@”) in a status, in a picture description or in comments


Anne Marie is 😃 feeling happy with Anis Jourdain.
Just now ·

or

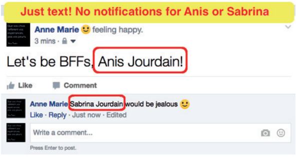

Anne Marie with Anis Jourdain.
January 29 ·

or



Add a description With Anis Jourdain.

□ They post content about me, but do not tag or link my name, so I don't get notified. I have to look at their activity to identify such content.



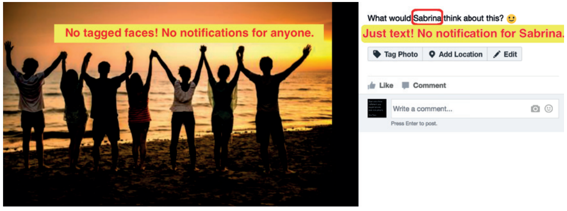
Just text! No notifications for Anis or Sabrina

Anne Marie feeling happy.
3 mins ·

Let's be BFFs, Anis Jourdain!

Anne Marie Sabrina Jourdain would be jealous
Like · Reply · Just now · Edited

or



No tagged faces! No notifications for anyone.

What would Sabrina think about this?

Just text! No notification for Sabrina.

25. [Conditioned on the positive response to Q18] How often do *you ask your friends to remove* status updates and/or pictures that they have shared of you on Facebook?

○ Very often (Almost every time)
○ Somewhat often (In more than half of the cases)
○ Occasionally (In less than half of the cases)
○ Infrequently (Hardly ever or in very few of the cases)

26. [Conditioned on the positive response to Q18] How often do *you untag yourself* from status updates and/or photos in which other people had tagged you?

○ Very often (Almost every time)
○ Somewhat often (In more than half of the cases)
○ Occasionally (In less than half of the cases)
○ Infrequently (Hardly ever or in very few of the cases)

Part V: A world of consent

27. We have designed a *system* that requires *consent from everyone appearing in a picture before* their faces are uploaded online (e.g., on Facebook).
Therefore, you will be asked to *accept* or *deny* that you are visible in every picture that your friends post of you. Note that this can be automated with automatic policies (e.g., always decline for sexually explicit content).
Our system guarantees *privacy* to the users with respect to each picture. This means that:

   – Facebook or our system itself would *not* be able to see the actual part of the picture in which you appear *before* you give consent.

– Our system would *not* be able to link your identity with any given picture *before* you give consent.

How useful would you find this system? (We would greatly appreciate any additional comments about the reasons for your choice).

○ Very useful
○ Useful
○ Neither useful nor unuseful
○ Somewhat useful
○ Not useful at all

Please enter your comment here:

□

28. In order to limit the *burden of manually reviewing* each and every one of the pictures requiring your consent, *you could specify automatic policies.* Which one of the following would you use:

□ Deny for sexually explicit images

□ Deny for partial nudity (e.g., photos in a swimsuit)

□ Deny for geo-tagged photos (where location is specified)

□ None of these

□ Other □

29. In order to limit the *burden of manually reviewing* each and every picture posted, another option would be to manually review a few pictures posted when first using the system. During this initial phase, *the system learns* your individual decision patterns (e.g., the friends you usually allow to post pictures or the type of pictures you usually deny or accept). After this initial phase, you *let the system make decisions on your behalf.* The system would be set up that when its confidence is not very high (a confidence level which you can specify, therefore *you control the level of automation you want to allow*), it would still prompt you for manual decisions and use them to further learn.
Would you use this feature?

○ Strongly disagree
○ Disagree
○ Neither agree nor disagree
○ Agree
○ Strongly agree

30. In order for our system to identify *you* in future pictures, you would have to *register* by providing *a few webcam or phone camera pictures.* Our system does *not store these pictures* themselves, but uses them to compute representations of your face. The computations are done locally on your smartphone or computer. *How comfortable would you be with this*? (We would greatly appreciate any additional comments about the reasons for your choice).

○ Very comfortable
○ Comfortable
○ Neither comfortable nor uncomfortable
○ Uncomfortable
○ Very uncomfortable

Please enter your comment here:

31. Consequently, when *you want to post* a picture in which *others appear*, a sanitized version of this picture is initially shared online, in which the faces of others are blurred for privacy reasons, but the background of the picture and yourself are fully visible. We emphasize that the sharing of your picture would not be slowed down by in any way. As soon as others give consent, their face is *automatically made visible* in the picture.
    Would you find this inconvenient? (We would greatly appreciate any additional comments about the reasons for your choice).

    ○ Strongly disagree
    ○ Disagree
    ○ Neither agree nor disagree
    ○ Agree
    ○ Strongly agree

    Please enter your comment here:

32. Considering the aforementioned benefits and functionality of our system, *would you consider using this system*? This is a general question, we will not contact you for further experiments, but we would greatly appreciate any additional comments about the reasons for your choice.

    ○ Yes
    ○ Maybe
    ○ No

    Please enter your comment here:

Chapter 5

Conclusion

As soon as I wake up [· · ·] I remember that everything is interrelated, the teaching of interdependence. So then I set my intention for the day: that this day should be meaningful. Meaningful means, if possible, serve and help others. If not possible, then at least not to harm others. That's a meaningful day.

Dalai Lama

In this thesis, we have exposed new privacy challenges that stem from the natural interdependencies in the data shared by individuals. We have analysed how users' divergent behaviors can affect the global privacy of OSN users, and we have identified the driving factors of their decisions to share location and co-location information. Finally, we have proposed mechanisms to be implemented by service providers – possibly enforced by regulatory entities: These mechanisms mitigate the privacy threats caused by others' sharing decisions, by improving users' awareness, and by giving them the option to control whether to permit the data to be shared, as a pre-emptive step to its misuse.

In Chapter 2, we have formally quantified the location privacy of users of location-based services, by relying on a Bayesian inference approach. We have proposed a probabilistic framework for inferring locations of users at times where it is unknown; this framework incorporates knowledge of user mobility profiles in the form of Markov chains, as well as user-reported locations and co-locations. We have formally shown that an optimal inference algorithm that considers co-locations is intractable due to the high complexity, and we have proposed polynomial-time approximate solutions (relying on the belief propagation algorithm) that converge to the optimal solution. In doing so, we have proven that an attacker can successfully exploit co-locations to localize users. Using a mobility dataset, we have quantified the privacy loss that stems from co-locations. Unsurprisingly, our results show that the more co-location the adversary has, the better he is able to localize users: Even when considering co-locations with only one friend of a target user, the target's location privacy is decreased by up to 62%, in a typical setting. In the case where a target user does not disclose any location information, her privacy can decrease by up to 21%, simply due to the location and co-location information reported by another user. This demonstrates that existing location-privacy protection-mechanisms fail to provide users full control over their privacy. We have proposed easy countermeasures for mitigating the effect of co-locations on location privacy and have evaluated their effectiveness. We have further suggested extensions of location-privacy protection mechanisms based on generalization or obfuscation of co-locations.

In Chapter 3, we have studied the problem of location and co-location information sharing on location-based OSNs. We have proposed the first game-theoretic framework to analyze the strategic behaviors of users in this setting. To enhance the practicality of our results, we have estimated the parameters of the utility functions from real users' actual preferences. Specifically, we have conducted a survey of 250 Facebook users and quantified their benefits of sharing vs. viewing (co)-location information and their preference for privacy vs. benefits through conjoint analysis. Our survey findings expose the fact there is a large variation, in terms of these preferences, among the users, which draws attention to the fact that conflictual situations can be frequent. We have simulated users' decisions, using our model, for various combinations of the estimated parameters, exposing situations where dangerous patterns can emerge (*e.g.,* a vicious circle of sharing, or an incentive to over-share), even when the users have similar preferences.

In Chapter 4, we have proposed a framework for sharing, in a consensual and privacy-preserving manner, various types of data that have privacy implications for subjects other than the uploader. We have identified the different challenges in the design of such a framework, the main building blocks, as well as the incentives for adoption for all the parties involved. We have designed, implemented and evaluated our proposed system for photos (ConsenShare) by using image processing and cryptographic techniques. The key property of ConsenShare is that it is privacy-preserving by design, not only with respect to other users of the system but also with respect to the service providers involved. We have experimentally demonstrated the feasibility of our approach, by using a Flickr dataset of 20k photos: We conclude that the overhead for ensuring privacy, while preserving existing features of photo sharing platforms (such as Flickr), is acceptable. Furthermore, we have conducted a user study of Facebook users: It reveals interest from users for a system such as ConsenShare, as well as their growing concern and limited awareness regarding photos shared on Facebook in which they appear. Our work constitutes an important first step in the design of privacy-preserving sharing of various types of interdependent and multi-subject data.

In conclusion, our hope is that this thesis will help raise awareness among end-users, service providers, and regulatory institutions alike, regarding the privacy risks of our era. We believe that a crucial step in ensuring privacy is to educate users about interactions with technology. We were saddened to observe, in our surveys, the repeated remarks of the type "*I do not need to be worried about my privacy because I have done nothing wrong and I have nothing to hide.*" We are confident that this thesis is a step forward in the direction of convincing such individuals that privacy is not about hiding things, but about protecting things and doing so, not only for themselves, but also for their peers. Furthermore, we hope that the threats we exposed might encourage people to express a certain curiosity and to practice caution when using services in their daily routine, as well as to consider the possible consequences of how the data they share could affect themselves or others in the future. As privacy will never again be evaluated or attained at an individual level, it must be protected as a global right, to which we all contribute. Last but not least, as our world becomes more connected, our technologies become more complex, and as the adversaries we face become smarter and faster, we trust that the problem of privacy will continue to be studied and its challenges continue to be overcome. To this end, we believe that our research has opened opportunities that could be further explored by ourselves and by other researchers, and that it has provided regulatory institutions with an enriched knowledge of the existing privacy problems and has presented them with adequate solutions. The implementation of which should be made mandatory to ensure the future of privacy.

Bibliography

[1] A. F. Westin and O. M. Ruebhausen, *Privacy and freedom.* Atheneum New York, 1967, vol. 1. [cited at pages v and vii]

[2] D. Crandall, L. Backstrom, D. Cosley, S. Suri, D. Huttenlocher, and J. Kleinberg, "Inferring social ties from geographic coincidences," *Proceedings of PNAS*, vol. 107, 2010. [cited at pages 1, 39, and 44]

[3] R. Cheng, J. Pang, and Y. Zhang, "Inferring friendship from check-in data of location-based social networks," in *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*, ser. ASONAM '15. New York, NY, USA: ACM, 2015, pp. 1284–1291. [cited at p. 1]

[4] A. Noulas, M. Musolesi, M. Pontil, and C. Mascolo, "Inferring interests from mobility and social interactions," in *Proceedings of NIPS Workshops*, 2009. [cited at pages 1 and 44]

[5] J. Freudiger, R. Shokri, and J.-P. Hubaux, "Evaluating the privacy risk of location-based services," in *International conference on financial cryptography and data security.* Springer, 2011, pp. 31–46. [cited at p. 1]

[6] J. Krumm, "Inference attacks on location tracks," in *International Conference on Pervasive Computing.* Springer, 2007, pp. 127–143. [cited at pages 1 and 39]

[7] A. Pyrgelis, C. Troncoso, and E. De Cristofaro, "Knock knock, who's there: Membership inference on aggregate location data," in *Network and Distributed System Security Symposium (NDSS)*, 2018. [cited at p. 2]

[8] A. Chaabane, G. Acs, and M. Kaafar, "You are what you like! information leakage through users' interests," in *NDSS*, 2012. [cited at pages 2 and 74]

[9] R. Dey, C. Tang, K. Ross, and N. Saxena, "Estimating age privacy leakage in online social networks," in *INFOCOM*, 2012. [cited at pages 2, 39, 74, and 81]

[10] A. Mislove, B. Viswanath, K. P. Gummadi, and P. Druschel, "You are who you know: inferring user profiles in online social networks," in *Proceedings of WSDM.* ACM, 2010. [cited at pages 2, 39, 40, 74, and 81]

[11] M. Humbert, T. Studer, M. Grossglauser, and J.-P. Hubaux, "Nowhere to hide: Navigating around privacy in online social networks," in *European Symposium on Research in Computer Security.* Springer, 2013, pp. 682–699. [cited at pages 2 and 74]

[12] M. Humbert, E. Ayday, J.-P. Hubaux, and A. Telenti, “Quantifying Interdependent Risks in Genomic Privacy,” *ACM TOPS*, 2017. [cited at pages 2, 39, 81, 82, 106, and 107]

[13] “With genetic testing, i gave my parents the gift of divorce,” http://www: vox:com/2014/9/9/5975653/with-genetic-testing-i-gave-my-parents-the-gift-of-divorce-23andme, last visited: Dec. 2018. [cited at p. 2]

[14] “Military women ask Facebook to do more to stop revenge porn,” http://www.refinery29.com/2017/04/149960/military-women-facebook-revenge-porn, 2017, last visited: Dec. 2018. [cited at pages 2 and 81]

[15] “Revenge porn - the ugly side of social media,” Online, https://www.thetimes.co.uk/article/its-destroying-the-lives-of-17-year-olds-nbktc352n, 2017, last visited: Dec. 2018. [cited at pages 2 and 81]

[16] “Revenge porn: Image-based abuse hits one in five Australians,” Online, http://www.bbc.com/news/world-australia-39777192, 2017, last visited: Dec. 2018. [cited at pages 2 and 81]

[17] “Woman kills herself over the threat of revenge porn,” http://www.dailymail.co.uk/news/article-5630429/Woman-24-jumped-death-man-threatened-send-revenge-porn-video.html, last visited: Dec. 2018. [cited at p. 2]

[18] “Italian woman kills herself after revenge porn video goes viral,” https://metro.co.uk/2016/09/16/italian-woman-kills-herself-after-revenge-porn-video-goes-viral-6131294/, last visited: Dec. 2018. [cited at p. 2]

[19] A. Acquisti, L. Brandimarte, and G. Loewenstein, “Privacy and human behavior in the age of information,” *Science*, vol. 347, no. 6221, pp. 509–514, 2015. [cited at pages 2 and 46]

[20] “Govt. agencies, colleges demand applicants’ facebook passwords,” https://www.nbcnews.com/business/consumer/govt-agencies-colleges-demand-applicants-facebook-passwords-f328791, last visited: Dec. 2018. [cited at p. 2]

[21] A.-M. Olteanu, K. Huguenin, R. Shokri, and J.-P. Hubaux, “Quantifying the Effect of Co-locations on Location Privacy,” in *PETS’14: Proceedings of the 14th International Symposium on Privacy Enhancing Technologies*, 2014. [cited at p. 5]

[22] A.-M. Olteanu, K. Huguenin, R. Shokri, M. Humbert, and J.-P. Hubaux, “Quantifying interdependent privacy risks with location data,” in *IEEE Transactions on Mobile Computing*, 2016. [cited at pages 5, 25, 44, 55, 61, 74, and 78]

[23] A.-M. Olteanu, M. Humbert, K. Huguenin, and J.-P. Hubaux, “The (Co)-Location Sharing Game,” in *Privacy Enhancing Technologies Symposium (PoPETs)*, 2019. [cited at p. 5]

[24] A.-M. Olteanu, K. Huguenin, I. Dacosta, and J.-P. Hubaux, “Consensual and privacy-preserving sharing of multi-subject and interdependent data,” in *25th Network and Distributed System Security Symposium (NDSS)*, 2018. [cited at p. 5]

[25] "Facebook Messenger adds fast photo sharing using face recognition," The Verge, http://www.theverge.com/2015/11/9/9696760/facebook-messenger-photo-sharing-face-recognition, nov 2015, last visited: Dec. 2018. [cited at p. 8]

[26] S. Gnesi, I. Matteucci, C. Moiso, P. Mori, M. Petrocchi, and M. Vescovi, "My Data, Your Data, Our Data: Managing Privacy Preferences in Multiple Subjects Personal Data," in *Proceedings of APF*. Springer, 2014. [cited at pages 8, 81, and 108]

[27] C. Vicente, D. Freni, C. Bettini, and C. S. Jensen, "Location-related privacy in geo-social networks," *IEEE Internet Computing*, vol. 15, no. 3, pp. 20–27, 2011. [cited at pages 8 and 40]

[28] R. Shokri, G. Theodorakopoulos, J.-Y. Le Boudec, and J.-P. Hubaux, "Quantifying location privacy," in *S&P*, 2011, pp. 247–262. [cited at pages 9, 12, and 39]

[29] L. E. Baum and T. Petrie, "Statistical inference for probabilistic functions of finite state markov chains," *The Annals of Mathematical Statistics*, vol. 37, no. 6, pp. 1554–1563, 1966. [cited at p. 11]

[30] A. Narayanan and V. Shmatikov, "De-anonymizing social networks," in *S&P'09: Proceedings of the 30th IEEE Symp. on Security and Privacy*, 2009, pp. 173–187. [cited at p. 12]

[31] G. Ghinita, P. Kalnis, A. Khoshgozaran, C. Shahabi, and K.-L. Tan, "Private queries in location based services: Anonymizers are not necessary," in *SIGMOD*, 2008, pp. 121–132. [cited at p. 12]

[32] R. L. Stratonovich, "Conditional Markov Processes," *Theory of Probability & its Applications*, vol. 5, no. 2, pp. 156–178, 1960. [cited at p. 13]

[33] D. Koller and N. Friedman, *Probabilistic graphical models: principles and techniques*. MIT press, 2009. [cited at pages 13, 20, and 59]

[34] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann, 2014. [cited at p. 20]

[35] K. P. Murphy, Y. Weiss, and M. I. Jordan, "Loopy belief propagation for approximate inference: An empirical study," in *UAI*. Morgan Kaufmann Publishers Inc., 1999, pp. 467–475. [cited at p. 20]

[36] R. I. M. Dunbar, "Neocortex size as a constraint on group size in primates," *Journal of Human Evolution*, vol. 22, no. 6, pp. 469–493, 1992. [cited at p. 21]

[37] P. Ilia, I. Polakis, E. Athanasopoulos, F. Maggi, and S. Ioannidis, "Face/Off: Preventing Privacy Leakage From Photos in Social Networks," in *Proceedings of CCS*. ACM, 2015. [cited at pages 22, 43, and 109]

[38] M. E. Andrés, N. E. Bordenabe, K. Chatzikokolakis, and C. Palamidessi, "Geo-indistinguishability: differential privacy for location-based systems," in *CCS*, 2013, pp. 901–914. [cited at p. 24]

[39] K. Chatzikokolakis, C. Palamidessi, and M. Stronati, "A Predictive Differentially-Private Mechanism for Mobility Traces," in *PETS*, 2014, pp. 21–41. [cited at p. 24]

[40] D. Kifer and A. Machanavajjhala, "No free lunch in data privacy," in *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data.* ACM, 2011, pp. 193–204. [cited at p. 24]

[41] ——, "A rigorous and customizable framework for privacy," in *PODS*, 2012, pp. 77–88. [cited at p. 25]

[42] ——, "Pufferfish: A framework for mathematical privacy definitions," *ACM Transactions on Database Systems (TODS)*, vol. 39, no. 1, p. 3, 2014. [cited at p. 25]

[43] R. Chen, B. C. Fung, P. S. Yu, and B. C. Desai, "Correlated network data publication via differential privacy," *The VLDB Journal—The International Journal on Very Large Data Bases*, vol. 23, no. 4, pp. 653–676, 2014. [cited at p. 25]

[44] B. Yang, I. Sato, and H. Nakagawa, "Bayesian differential privacy on correlated data," in *Proceedings of the 2015 ACM SIGMOD international conference on Management of Data.* ACM, 2015, pp. 747–762. [cited at p. 25]

[45] T. Zhu, P. Xiong, G. Li, and W. Zhou, "Correlated differential privacy: hiding information in non-iid data set," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 2, pp. 229–242, 2015. [cited at p. 25]

[46] C. Liu, S. Chakraborty, and P. Mittal, "Dependence Makes You Vulnerable: Differential Privacy Under Dependent Tuples," in *NDSS*, 2016, to appear. [cited at p. 25]

[47] S. Song, Y. Wang, and K. Chaudhuri, "Pufferfish privacy mechanisms for correlated data," in *Proceedings of the 2017 ACM International Conference on Management of Data.* ACM, 2017, pp. 1291–1306. [cited at p. 25]

[48] Y. Zheng, L. Liu, L. Wang, and X. Xie, "Learning transportation mode from raw GPS data for geographic applications on the web," in *WWW*, 2008, pp. 247–256. [cited at pages 25, 45, and 60]

[49] Y. De Mulder, G. Danezis, L. Batina, and B. Preneel, "Identification via location-profiling in GSM networks," in *WPES*, 2008. [cited at p. 39]

[50] P. Golle and K. Partridge, "On the anonymity of home/work location pairs," in *Pervasive*, 2009, pp. 390–397. [cited at p. 39]

[51] B. Hoh, M. Gruteser, H. Xiong, and A. Alrabady, "Enhancing security and privacy in traffic-monitoring systems," *IEEE Pervasive Computing*, vol. 5, no. 4, pp. 38–46, 2006. [cited at p. 39]

[52] G. Ghinita, M. L. Damiani, C. Silvestri, and E. Bertino, "Preventing velocity-based linkage attacks in location-aware applications," in *GIS*, 2009, pp. 246–255. [cited at p. 39]

[53] M. L. Damiani, E. Bertino, and C. Silvestri, "The PROBE framework for the personalized cloaking of private locations," *Transactions on Data Privacy*, vol. 3, pp. 123–148, 2010. [cited at p. 39]

[54] L. Huang, H. Yamane, K. Matsuura, and K. Sezaki, "Silent cascade: Enhancing location privacy without communication QoS degradation," in *SPC*, 2006, pp. 165–180. [cited at p. 39]

[55] M. Srivatsa and M. Hicks, “Deanonymizing mobility traces: Using social network as a side-channel,” in *CCS*, 2012, pp. 628–637. [cited at p. 39]

[56] N. Eagle, A. Pentland, and D. Lazer, “Inferring Friendship Network Structure by Using Mobile Phone Data,” *Proceedings of the National Academy of Sciences*, vol. 106, pp. 15 274–15 278, 2009. [cited at p. 39]

[57] “How the NSA is tracking people right now,” http://apps.washingtonpost.com/g/page/national/how-the-nsa-is-tracking-people-right-now/634/, 2013, last visited: Feb. 2014. [cited at p. 39]

[58] C. A. Davis Jr, G. L. Pappa, D. R. R. de Oliveira, and F. de L Arcanjo, “Inferring the location of twitter messages based on user relationships,” *Transactions in GIS*, vol. 15, no. 6, pp. 735–751, 2011. [cited at p. 39]

[59] D. Xu, P. Cui, W. Zhu, and S. Yang, “Graph-based residence location inference for social media users,” *IEEE MultiMedia*, vol. 21, no. 4, pp. 76–83, 2014. [cited at p. 39]

[60] Y. Jia, Y. Wang, X. Jin, and X. Cheng, “TSBM: The temporal-spatial Bayesian model for location prediction in social networks,” in *WI-IAT*, vol. 2, 2014, pp. 194–201. [cited at p. 39]

[61] A. Uchiyama, S. Fujii, K. Maeda, T. Umedu, H. Yamaguchi, and T. Higashino, “UPL: Opportunistic localization in urban districts,” *IEEE Transactions on Mobile Computing*, vol. 12, no. 5, pp. 1009–1022, 2013. [cited at p. 39]

[62] E. Cho, S. A. Myers, and J. Leskovec, “Friendship and mobility: user movement in location-based social networks,” in *Proceedings of KDD*. ACM, 2011. [cited at p. 39]

[63] I. Polakis, G. Argyros, T. Petsios, S. Sivakorn, and A. D. Keromytis, “Where’s wally?: Precise user discovery attacks in location proximity services,” in *CCS*, 2015, pp. 817–828. [cited at p. 39]

[64] M. Li, H. Zhu, Z. Gao, S. Chen, L. Yu, S. Hu, and K. Ren, “All your location are belong to us: Breaking mobile social networks for automated user location tracking,” in *MobiHoc*, 2014. [cited at p. 39]

[65] M. Xue, C. Ballard, K. Liu, C. Nemelka, Y. Wu, K. Ross, and H. Qian, “You can yak but you can’t hide: Localizing anonymous social network users,” in *Proceedings of the 2016 Internet Measurement Conference*. ACM, 2016, pp. 25–31. [cited at p. 39]

[66] M. Xue, Y. Liu, K. W. Ross, and H. Qian, “I know where you are: thwarting privacy protection in location-based social discovery services,” in *Computer Communications Workshops (INFOCOM WKSHPS), 2015 IEEE Conference on*. IEEE, 2015, pp. 179–184. [cited at p. 39]

[67] B. Henne, C. Szongott, and M. Smith, “Snapme if you can: Privacy threats of other peoples’ geo-tagged media and what we can do about it,” in *WiSec*, 2013, pp. 95–106. [cited at pages 39 and 74]

[68] L. Backstrom, E. Sun, and C. Marlow, “Find me if you can: improving geographical prediction with social and spatial proximity,” in *Proceedings of WWW*. ACM, 2010. [cited at pages 39 and 74]

[69] N. Vratonjic, K. Huguenin, V. Bindschaedler, and J.-P. Hubaux, "A location-privacy threat stemming from the use of shared public ip addresses," *IEEE Transactions on Mobile Computing*, vol. 13, no. 11, pp. 2445–2457, 2014. [cited at pages 39 and 74]

[70] M. Gymrek, A. L. McGuire, D. Golan, E. Halperin, and Y. Erlich, "Identifying personal genomes by surname inference," *Science*, vol. 339, no. 6117, pp. 321–324, 2013. [cited at p. 39]

[71] G. Biczók and P. H. Chia, "Interdependent privacy: Let me share your data," in *FC*, 2013, pp. 338–353. [cited at pages 39, 44, 74, 75, and 82]

[72] Y. Pu and J. Grossklags, "An economic model and simulation results of app adoption decisions on networks with interdependent privacy consequences," in *GameSec*, 2014, pp. 246–265. [cited at pages 39, 44, 72, 74, and 75]

[73] M. Humbert, E. Ayday, J.-P. Hubaux, and A. Telenti, "On non-cooperative genomic privacy," in *FC*, 2015. [cited at pages 40, 74, and 75]

[74] "Device usage of Facebook users worldwide as of January 2018," https://www.statista.com/statistics/377808/distribution-of-facebook-users-by-device/, last visited: Dec. 2018. [cited at p. 43]

[75] H. Krasnova, S. Spiekermann, K. Koroleva, and T. Hildebrand, "Online social networks: Why we disclose," *Journal of information technology*, vol. 25, no. 2, 2010. [cited at p. 44]

[76] J. Krumm, "A survey of computational location privacy," *Personal and Ubiquitous Computing*, vol. 13, no. 6, 2009. [cited at p. 44]

[77] C. Riederer, D. Echickson, S. Huang, and A. Chaintreau, "Findyou: A personal location privacy auditing tool," in *WWW*, 2016. [cited at p. 44]

[78] L. González-Manzano, A. I. González-Tablas, J. M. de Fuentes, and A. Ribagorda, "Cooped: Co-owned personal data management," *Computers & Security*, vol. 47, 2014. [cited at p. 44]

[79] J. M. Such and N. Criado, "Resolving multi-party privacy conflicts in social media," *IEEE KDE*, 2016. [cited at pages 44, 74, 82, and 108]

[80] J. M. Such, J. Porter, S. Preibusch, and A. Joinson, "Photo privacy conflicts in social media: A large-scale empirical study," in *Proceedings of CHI*, 2017. [cited at pages 44, 74, 81, 82, 105, and 108]

[81] J. Von Neumann and O. Morgenstern, *Theory of games and economic behavior*. Princeton university press, 2007. [cited at pages 44 and 45]

[82] R. B. Myerson, *Game theory*. Harvard university press, 2013. [cited at pages 44 and 45]

[83] D. Fudenberg and J. Tirole, *Game theory*. MIT press, 1991. [cited at pages 44 and 45]

[84] P. E. Green and V. Srinivasan, "Conjoint Analysis in Consumer Research: Issues and Outlook," *Journal of Consumer Research*, vol. 5, no. 2, 1978. [cited at pages 44, 46, and 54]

[85] A. C. Squicciarini, M. Shehab, and F. Paci, "Collective privacy management in social networks," in *Proceedings of the 18th International Conference on World Wide Web*, ser. WWW '09. New York, NY, USA: ACM, 2009, pp. 521–530. [cited at p. 44]

[86] R. Mason, R. Gunst, and J. Hess, *Statistical design and analysis of experiments with applications to engineering and science.* J. Wiley, 2003. [cited at p. 46]

[87] A. Acquisti and J. Grossklags, "Privacy and rationality in individual decision making," *IEEE Security Privacy*, vol. 3, no. 1, 2005. [cited at pages 46 and 106]

[88] E. M. Redmiles, M. L. Mazurek, and J. P. Dickerson, "Poster: Do users make rational security decisions?" 2018. [cited at p. 46]

[89] ——, "Dancing Pigs or Externalities?: Measuring the Rationality of Security Decisions," in *Proceedings of the 2018 ACM Conference on Economics and Computation.* ACM, 2018, pp. 215–232. [cited at p. 46]

[90] A. Acquisti, "Privacy in electronic commerce and the economics of immediate gratification," in *Proceedings of the 5th ACM conference on Electronic commerce.* ACM, 2004, pp. 21–29. [cited at pages 49 and 50]

[91] R. S. Laufer, H. M. Proshansky, and M. Wolfe, "Some analytic dimensions of privacy," in *Proceedings of the Lund Conference on Architectural Psychology.* Lund, Sweden, 1973. [cited at pages 49 and 74]

[92] R. S. Laufer and M. Wolfe, "Privacy as a concept and a social issue: A multidimensional developmental theory," *Journal of social Issues*, vol. 33, no. 3, pp. 22–42, 1977. [cited at pages 49 and 74]

[93] A. F. Westin, "Social and political dimensions of privacy," *Journal of social issues*, vol. 59, no. 2, pp. 431–453, 2003. [cited at p. 50]

[94] "Xlstat statistical software for microsoft excel," https://www.xlstat.com/en/, 2016, last visited: Dec. 2018. [cited at p. 54]

[95] A. Acquisti and J. Grossklags, *Privacy Attitudes and Privacy Behavior*, 2004. [cited at pages 56 and 73]

[96] S. Kokolakis, "Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon," *Computers Security*, 2015. [cited at pages 56 and 73]

[97] X. Page, B. P. Knijnenburg, and A. Kobsa, "FYI: Communication Style Preferences Underlie Differences in Location-sharing Adoption and Usage," in *UbiComp*, 2013. [cited at p. 56]

[98] F. Benevenuto, T. Rodrigues, M. Cha, and V. Almeida, "Characterizing user behavior in online social networks," in *IMC*, 2009. [cited at p. 56]

[99] F. Xu, Z. Tu, Y. Li, P. Zhang, X. Fu, and D. Jin, "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, 2017. [cited at p. 58]

[100] D. Meier, Y. A. Oswald, S. Schmid, and R. Wattenhofer, "On the Windfall of Friendship: Inoculation Strategies on Social Networks," in *EC*, 2008. [cited at pages 72 and 74]

[101] Y. Pu and J. Grossklags, "Valuating friends' privacy: Does anonymity of sharing personal data matter?" in *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*. USENIX Association, 2017. [cited at pages 72 and 74]

[102] ——, "Towards a model on the factors influencing social app users' valuation of interdependent privacy," *PoPETS*, 2015. [cited at pages 72 and 74]

[103] ——, "Using conjoint analysis to investigate the value of interdependent privacy in social app adoption scenarios," in *ICIS*. Assoc. for Information Systems, 2015. [cited at pages 72 and 74]

[104] D. Koller and B. Milch, "Multi-agent influence diagrams for representing and solving games," *Games and economic behavior*, vol. 45, no. 1, 2003. [cited at p. 73]

[105] E. Toch, J. Cranshaw, P. H. Drielsma, J. Y. Tsai, P. G. Kelley, J. Springfield, L. Cranor, J. Hong, and N. Sadeh, "Empirical models of privacy in location sharing," in *UbiComp*, 2010. [cited at p. 74]

[106] J. Wiese, P. G. Kelley, L. F. Cranor, L. Dabbish, J. I. Hong, and J. Zimmerman, "Are you close with me? are you nearby?: Investigating social groups, closeness, and willingness to share," in *UbiComp*, 2011. [cited at p. 74]

[107] Z. D. Ozdemir, H. J. Smith, and J. H. Benamati, "Antecedents and outcomes of information privacy concerns in a peer context: An exploratory study," *European Journal of Information Systems*, vol. 26, no. 6, pp. 642–660, 2017. [cited at p. 74]

[108] K. Raynes-Goldie, "Aliases, creeping, and wall cleaning: Understanding privacy in the age of facebook," *First Monday*, vol. 15, no. 1, 2010. [cited at p. 74]

[109] G. Wang, S. Y. Schoenebeck, H. Zheng, and B. Y. Zhao, "“will check-in for badges": Understanding bias and misbehavior on location-based social networks," in *ICWSM*, 2016. [cited at p. 74]

[110] M. Yang, Y. Yu, A. K. Bandara, and B. Nuseibeh, "Adaptive sharing for online social networks: A trade-off between privacy risk and social benefit," in *TrustCom*, 2014. [cited at p. 74]

[111] R. Shokri, G. Theodorakopoulos, C. Troncoso, J.-P. Hubaux, and J.-Y. Le Boudec, "Protecting location privacy: optimal strategy against localization attacks," in *CCS'12: Proceedings of the 19th ACM Conf. on Computer and Communications Security*, 2012. [cited at p. 74]

[112] R. Shokri, "Privacy games: Optimal user-centric data obfuscation," *PoPETS*, vol. 2015, no. 2, 2015. [cited at p. 74]

[113] I. Bilogrevic, K. Huguenin, B. Ağır, M. Jadliwala, M. Gazaki, and J.-P. Hubaux, "A Machine-Learning Based Approach to Privacy-Aware Information-Sharing in Mobile Social Networks," *Pervasive and Mobile Computing (PMC)*, Nov. 2015. [cited at p. 74]

[114] M. Humbert, E. Ayday, J.-P. Hubaux, and A. Telenti, "Addressing the Concerns of the Lacks Family: Quantification of Kin Genomic Privacy," in *CCS'13: Proceedings of the 20th ACM Conf. on Computer and Communications Security*, 2013. [cited at p. 74]

[115] Y. Shoshitaishvili, C. Kruegel, and G. Vigna, "Portrait of a privacy invasion; detecting relationships through large-scale photo analysis," *PoPETS*, 2015. [cited at p. 74]

[116] J. Chen, J. W. Ping, Y. C. Xu, and B. C. Tan, "Information privacy concern about peer disclosure in online social networks," *IEEE Transactions on Engineering Management*, vol. 62, no. 3, pp. 311–324, 2015. [cited at p. 74]

[117] H. Cho and A. Filippova, "Networked privacy management in facebook: A mixed-methods and multinational study," in *Proceedings of CSCW*, 2016. [cited at pages 74 and 82]

[118] H. Jia and H. Xu, "Autonomous and interdependent: Collaborative privacy management on social networking sites," in *Proceedings of CHI*, 2016. [cited at pages 74 and 82]

[119] A. Lampinen, V. Lehtinen, A. Lehmuskallio, and S. Tamminen, "We're in it together: interpersonal management of disclosure in social network services," in *Proceedings of CHI*. ACM, 2011. [cited at pages 74 and 82]

[120] P. Wisniewski, H. Lipford, and D. Wilson, "Fighting for my space: Coping mechanisms for sns boundary regulation," in *Proceedings of CHI*. ACM, 2012. [cited at pages 74 and 82]

[121] H. Xu, "Reframing privacy 2.0 in online social network," *U. Pa. J. Const. L.*, 2011. [cited at pages 74 and 82]

[122] A. Besmer and H. Richter Lipford, "Moving beyond untagging: photo privacy in a tagged world," in *Proceedings of CHI*. ACM, 2010. [cited at pages 74 and 82]

[123] J. M. Such and N. Criado, "Multiparty privacy in social media," *Communications of the ACM*, vol. 61, no. 8, pp. 74–81, 2018. [cited at pages 74, 82, and 108]

[124] G. Misra and J. M. Such, "Pacman: Personal agent for access control in social media," *IEEE Internet Computing*, vol. 21, no. 6, pp. 18–26, 2017. [cited at p. 74]

[125] R. L. Fogues, P. K. Murukannaiah, J. M. Such, and M. P. Singh, "Sharing policies in multiuser privacy scenarios: Incorporating context, preferences, and arguments in decision making," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 24, no. 1, p. 5, 2017. [cited at p. 74]

[126] S. Rajtmajer, A. Squicciarini, J. M. Such, J. Semonsen, and A. Belmonte, "An ultimatum game model for the evolution of privacy in jointly managed content," in *International Conference on Decision and Game Theory for Security*. Springer, 2017, pp. 112–130. [cited at p. 75]

[127] A. Acquisti, "Nudging privacy: The behavioral economics of personal information," *IEEE Security Privacy*, 2009. [cited at p. 75]

[128] A. Laszka, M. Felegyhazi, and L. Buttyan, “A survey of interdependent information security games,” *ACM Computing Surveys*, 2015. [cited at p. 75]

[129] S. Ahern, D. Eckles, N. S. Good, S. King, M. Naaman, and R. Nair, “Over-exposed?: privacy patterns and considerations in online and mobile photo sharing,” in *Proceedings of CHI*. ACM, 2007. [cited at p. 81]

[130] K. Walker and E. Sleath, “A systematic review of the current knowledge regarding revenge pornography and non-consensual sharing of sexually explicit media,” *Aggression and violent behavior*, vol. 36, pp. 9–24, 2017. [cited at p. 81]

[131] “Facebook warned over legal action after revenge porn case,” http://www.bbc.com/news/uk-northern-ireland-42675036, last visited: Dec. 2018. [cited at p. 81]

[132] H. Nissenbaum, *Privacy in context: Technology, policy, and the integrity of social life*. Stanford University Press, 2009. [cited at pages 82 and 107]

[133] “Face recognition app taking russia by storm,” https://www.theguardian.com/technology/2016/may/17/findface-face-recognition-app-end-public-anonymity-vkontakte, last visited: Dec. 2018. [cited at p. 82]

[134] “Reddit policy regarding revenge pornography,” Reddit, https://reddit.zendesk.com/hc/en-us/articles/205704725 and https://www.reddit.com/help/contentpolicy, last visited: Dec. 2018. [cited at pages 83 and 108]

[135] “Facebook announces new ways to prevent revenge porn,” NY Times, https://www.nytimes.com/2017/04/05/us/facebook-revenge-porn.html, last visited: Dec. 2018. [cited at pages 83 and 108]

[136] “Facebook asks users for nude photos in project to combat revenge porn,” The Guardian, https://www.theguardian.com/technology/2017/nov/07/facebook-revenge-porn-nude-photos, last visited: Dec. 2018. [cited at p. 83]

[137] “The twitter rules,” https://support.twitter.com/articles/18311, last visited: Dec. 2018. [cited at pages 83 and 108]

[138] “Facial recognition for porn stars is a privacy nightmare waiting to happen,” https://motherboard.vice.com/en_us/article/a3kmpb/facial-recognition-for-porn-stars-is-a-privacy-nightmare-waiting-to-happen, last visited: Dec. 2018. [cited at p. 83]

[139] “Pornhub revenge porn content removal,” https://www.pornhub.com/content-removal, last visited: Dec. 2018. [cited at p. 83]

[140] B. Thomee, D. A. Shamma, G. Friedland, B. Elizalde, K. Ni, D. Poland, D. Borth, and L.-J. Li, “YFCC100M: The new data in multimedia research,” *Commun. ACM*, 2016. [cited at pages 83 and 95]

[141] H. Wang, X. Bao, R. R. Choudhury, and S. Nelakuditi, “Insight: Recognizing humans without face recognition,” in *Proceedings of HotMobile*. ACM, 2013. [cited at pages 86 and 88]

[142] H. Wang, X. Bao, R. Roy Choudhury, and S. Nelakuditi, "Visually fingerprinting humans without face recognition," in *Proceedings of MobiSys.* ACM, 2015. [cited at pages 86 and 88]

[143] A. Frome, G. Cheung, A. Abdulkader, M. Zennaro, B. Wu, A. Bissacco, H. Adam, H. Neven, and L. Vincent, "Large-scale privacy protection in google street view," in *Proceedings of ICCV.* IEEE, 2009. [cited at p. 87]

[144] Y. Li, N. Vishwamitra, B. P. Knijnenburg, H. Hu, and K. Caine, "Effectiveness and users' experience of obfuscation as a privacy-enhancing technology for sharing photos," *Manuscript submitted for publication*, vol. 4, 2017. [cited at p. 87]

[145] "Engineering Safety with Uber's real-time ID check," https://eng.uber.com/real-time-id-check/, 2017, last visited: Dec. 2018. [cited at p. 87]

[146] "Selfies and Security," https://newsroom.uber.com/securityselfies/, 2017, last visited: Dec. 2018. [cited at p. 87]

[147] "Introducing Airbnb verified ID," http://blog.atairbnb.com/introducing-airbnb-verified-id/, 2017, last visited: Dec. 2018. [cited at p. 87]

[148] "Facebook captcha test," https://www.wired.com/story/facebooks-new-captcha-test-upload-a-clear-photo-of-your-face/, last visited: Dec. 2018. [cited at p. 87]

[149] "Windows Hello: They can guess your password - not your face," https://www.microsoft.com/en-us/windows/windows-hello, 2017, last visited: Dec. 2018. [cited at p. 87]

[150] C. Holz, S. Buthpitiya, and M. Knaust, "Bodyprint: Biometric user identification on mobile devices using the capacitive touchscreen to scan body parts," in *Proceedings of CHI.* ACM, 2015. [cited at p. 87]

[151] C. Mallauran, J.-L. Dugelay, F. Perronnin, and C. Garcia, "Online face detection and user authentication," in *Proceedings of MM.* ACM, 2005. [cited at p. 87]

[152] K. Niinuma and A. K. Jain, "Continuous user authentication using temporal information," in *Proceedings of SPIE*, 2010. [cited at p. 87]

[153] "Apple's FaceID," https://www.apple.com/lae/iphone-xs/face-id/, last visited: Dec. 2018. [cited at p. 87]

[154] "Facebook can unlock your account with facial recognition," https://techcrunch.com/2017/09/29/facebook-face-id/, last visited: Dec. 2018. [cited at p. 87]

[155] R. Bost, R. A. Popa, S. Tu, and S. Goldwasser, "Machine learning classification over encrypted data." in *Proceedings of NDSS*, 2015. [cited at pages 88 and 109]

[156] D. J. Wu, T. Feng, M. Naehrig, and K. Lauter, "Privately evaluating decision trees and random forests," *PoPETs*, 2016. [cited at pages 88 and 109]

[157] R. Shokri and V. Shmatikov, "Privacy-preserving deep learning," in *Proceedings of the 22Nd ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS '15. New York, NY, USA: ACM, 2015, pp. 1310–1321. [cited at p. 88]

[158] M. Abadi, A. Chu, I. Goodfellow, H. B. McMahan, I. Mironov, K. Talwar, and L. Zhang, "Deep learning with differential privacy," in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security.* ACM, 2016, pp. 308–318. [cited at p. 88]

[159] P. Li, J. Li, Z. Huang, T. Li, C.-Z. Gao, S.-M. Yiu, and K. Chen, "Multi-key privacy-preserving deep learning in cloud computing," *Future Generation Computer Systems*, vol. 74, pp. 76–85, 2017. [cited at p. 88]

[160] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in *Proceedings of CVPR*, 2015. [cited at pages 88 and 94]

[161] "Facebook can recognise you in photos even if you're not looking," https://www.newscientist.com/article/dn27761-facebook-can-recognise-you-in-photos-even-if-youre-not-looking/, last visited: Dec. 2018. [cited at p. 88]

[162] K. Olejnik, I. Dacosta, J. Soares Machado, K. Huguenin, M. E. Khan, and J.-P. Hubaux, "Smarper: Context-aware and automatic runtime-permissions for mobile devices," in *Proceedings of S&P.* IEEE, 2017. [cited at p. 91]

[163] L. Yuan, J. Theytaz, and T. Ebrahimi, "Context-dependent privacy-aware photo sharing based on machine learning," in *Proceedings of IFIP SEC.* Springer, 2017. [cited at p. 91]

[164] P. Wijesekera, A. Baokar, L. Tsai, J. Reardon, S. Egelman, D. Wagner, and K. Beznosov, "The feasibility of dynamically granted permissions: Aligning mobile privacy with user preferences," *arXiv preprint arXiv:1703.02090*, 2017. [cited at p. 91]

[165] A. Nguyen, J. Yosinski, and J. Clune, "Deep neural networks are easily fooled: High confidence predictions for unrecognizable images," in *Proceedings of CVPR*, 2015. [cited at p. 92]

[166] J. Määttä, A. Hadid, and M. Pietikäinen, "Face spoofing detection from single images using texture and local shape analysis," *IET biometrics*, vol. 1, no. 1, pp. 3–10, 2012. [cited at p. 92]

[167] B. Amos, B. Ludwiczuk, and M. Satyanarayanan, "OpenFace: A general-purpose face recognition library with mobile applications," CMU-CS-16-118, CMU School of Computer Science, Tech. Rep., 2016. [cited at pages 93, 94, and 100]

[168] D. E. King, "Dlib-ml: A machine learning toolkit," *Journal of Machine Learning Research*, vol. 10, no. Jul, pp. 1755–1758, 2009. [cited at p. 94]

[169] "OpenFace - Free and open source face recognition with deep neural networks," https://cmusatyalab.github.io/openface/, last visited: Dec. 2018. [cited at p. 94]

[170] H. Jégou, R. Tavenard, M. Douze, and L. Amsaleg, "Searching in one billion vectors: re-rank with source coding," in *Proceedings of ICASSP*, 2011. [cited at p. 94]

[171] "PyNaCl - Python binding to the Networking and Cryptography library," https://github.com/pyca/pynacl, 2017, last visited: Dec. 2018. [cited at p. 95]

[172] “Recommendation for Keymanagement, Part 1: General, sp 800-57 Part 1 Rev. 4.” http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-57pt1r4.pdf, January 2016, last visited: Dec. 2018. [cited at p. 95]

[173] L. Yuan and T. Ebrahimi, “Image transmorphing with jpeg,” in *Proceedings of ICIP*. IEEE, 2015. [cited at p. 96]

[174] “How many public photos are uploaded to Flickr every day, month, year?” https://www.flickr.com/photos/franckmichel/6855169886, last visited: Dec. 2018. [cited at p. 98]

[175] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, “Web-scale training for face identification,” in *Proceedings of CVPR*, 2015. [cited at p. 100]

[176] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, “Labeled faces in the wild: A database for studying face recognition in unconstrained environments,” No 07-49, University of Massachusetts, Amherst, Tech. Rep., 2007. [cited at p. 100]

[177] “New report shows that 4% of U.S. internet users have been a victim of "revenge porn",” https://datasociety.net/blog/2016/12/13/nonconsensual-image-sharing/, last visited: Dec. 2018. [cited at p. 103]

[178] “Does a ’public figure’ have a right to privacy? well...” https://www.nytimes.com/1983/06/12/weekinreview/does-a-public-figure-have-a-right-to-privacy-well.html, last visited: Dec. 2018. [cited at p. 105]

[179] “Now you see us, now you don’t,” https://www.theguardian.com/world/2001/jan/08/law.media, last visited: Dec. 2018. [cited at p. 105]

[180] J. Ross, L. Irani, M. Silberman, A. Zaldivar, and B. Tomlinson, “Who are the crowdworkers?: shifting demographics in mechanical turk,” in *Proceedings of CHI*, 2010. [cited at p. 106]

[181] A. Sotirakopoulos, K. Hawkey, and K. Beznosov, “On the challenges in usable security lab studies: lessons learned from replicating a study on ssl warnings,” in *Proceedings of SOUPS*, 2011. [cited at p. 106]

[182] S. Egelman and E. Peer, “Scaling the security wall: Developing a security behavior intentions scale (sebis),” in *Proceedings of CHI*, 2015. [cited at p. 106]

[183] S. Egelman, M. Harbach, and E. Peer, “Behavior ever follows intention?: A validation of the security behavior intentions scale (sebis),” in *Proceedings of CHI*, 2016. [cited at p. 106]

[184] P. G. Kelley, “Conducting usable privacy & security studies with amazon’s mechanical turk,” in *Proceedings of SOUPS*, 2010. [cited at p. 106]

[185] D. A. Reynolds, “An overview of automatic speaker recognition technology,” in *Proceedings of ICASSP*, vol. 4. IEEE, 2002. [cited at p. 106]

[186] L. Muda, M. Begam, and I. Elamvazuthi, “Voice recognition algorithms using mel frequency cepstral coefficient (mfcc) and dynamic time warping (dtw) techniques,” *arXiv preprint arXiv:1003.4083*, 2010. [cited at p. 106]

[187] "Does your family have a right to your genetic code?" https://www.technologyreview.com/s/602946/do-your-family-members-have-a-right-to-your-genetic-code/, last visited: Dec. 2018. [cited at pages 106 and 108]

[188] E. Ayday and M. Humbert, "Inference attacks against kin genomic privacy," *IEEE Security & Privacy*, no. 5, 2017. [cited at p. 106]

[189] S. R. Savage, "Characterizing the risks and harms of linking genomic information to individuals," *IEEE Security & Privacy*, no. 5, 2017. [cited at p. 106]

[190] E. Ayday, Q. Tang, and A. Yilmaz, "Cryptographic solutions for credibility and liability issues of genomic data," *IEEE Transactions on Dependable and Secure Computing*, 2017. [cited at p. 106]

[191] F. Tramèr, Z. Huang, J.-P. Hubaux, and E. Ayday, "Differential privacy with bounded priors: reconciling utility and privacy in genome-wide association studies," in *Proceedings of CCS*, 2015. [cited at p. 106]

[192] N. S. Good, "Designing for informed consent: A multi-domain, interdisciplinary analysis of the technological means to provide informed consent, in order to manage users' privacy and security," Ph.D. dissertation, University of California at Berkeley, 2008. [cited at p. 107]

[193] "The OECD privacy framework," http://www.oecd.org/internet/ieconomy/privacy-guidelines.htm, 2013, last visited: Dec. 2018. [cited at p. 107]

[194] "Consumer data privacy in a networked world," https://www.whitehouse.gov/sites/default/files/privacy-final.pdf, 2013, last visited: Dec. 2018. [cited at p. 107]

[195] "Laws LJ Wood v Commissioner of Police for the Metropolis EWCA Civ 414," http://www.5rb.com/case/wood-v-commissioner-of-police-for-the-metropolis-ca/, 2009, last visited: Dec. 2018. [cited at p. 107]

[196] "Deepfake," https://en.wikipedia.org/wiki/Deepfake, last visited: Dec. 2018. [cited at p. 108]

[197] "Tumblr bans non-consensual creepshots and deepfake porn," https://www.bbc.com/news/technology-45323287, last visited: Dec. 2018. [cited at p. 108]

[198] "Reddit bans deepfakes AI porn communities," https://www.theverge.com/2018/2/7/16982046/reddit-deepfakes-ai-celebrity-face-swap-porn-community-ban, last visited: Dec. 2018. [cited at p. 108]

[199] "Reddit, Pornhub ban videos that use AI to superimpose a person's face over an X-rated actor," https://www.cnbc.com/2018/02/08/reddit-pornhub-ban-deepfake-porn-videos.html, last visited: Dec. 2018. [cited at p. 108]

[200] P. Aditya, R. Sen, P. Druschel, S. Joon Oh, R. Benenson, M. Fritz, B. Schiele, B. Bhattacharjee, and T. T. Wu, "I-Pic: A Platform for Privacy-Compliant Image Capture," in *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services*, ser. MobiSys '16. New York, NY, USA: ACM, 2016, pp. 235–248. [cited at p. 108]

[201] F. Beato and R. Peeters, "Collaborative joint content sharing for online social networks," in *Proceedings of PERCOM*. IEEE, 2014. [cited at p. 108]

[202] A. C. Squicciarini, M. Shehab, and F. Paci, "Collective privacy management in social networks," in *Proceedings of WWW*. ACM, 2009. [cited at p. 108]

[203] A. C. Squicciarini, M. Shehab, and J. Wede, "Privacy policies for shared content in social network sites," *Proceedings of VLDB*, 2010. [cited at p. 108]

[204] A. Ratikan and M. Shikida, "Privacy protection based privacy conflict detection and solution in online social networks," in *Proceedings of HAS*. Springer, 2014. [cited at p. 108]

[205] H. Hu and G.-J. Ahn, "Multiparty authorization framework for data sharing in online social networks," in *Proceedings of CODASPY*, 2011. [cited at p. 108]

[206] H. Hu, G.-J. Ahn, and J. Jorgensen, "Detecting and resolving privacy conflicts for collaborative data sharing in online social networks," in *Proceedings of ACSAC*. ACM, 2011. [cited at p. 108]

[207] ——, "Enabling collaborative data sharing in google+," in *Proceedings of GLOBECOM*. IEEE, 2012. [cited at p. 108]

[208] ——, "Multiparty access control for online social networks: model and mechanisms," *IEEE TKDE*, 2013. [cited at p. 108]

[209] P. Ilia, B. Carminati, E. Ferrari, P. Fragopoulou, and S. Ioannidis, "Sampac: Socially-aware collaborative multi-party access control," in *Proceedings of CODASPY*. ACM, 2017. [cited at p. 108]

[210] F. Li, J. Yu, L. Zhang, Z. Sun, and M. Lv, "A privacy-preserving method for photo sharing in instant message systems," in *Proceedings of ICCSP*. ACM, 2017. [cited at p. 108]

[211] Y. Guo, L. Zhang, and X. Chen, "Collaborative privacy management: mobile privacy beyond your own devices," in *Proceedings of MobiCom*, 2014. [cited at p. 108]

[212] E. De Cristofaro, C. Soriente, G. Tsudik, and A. Williams, "Hummingbird: Privacy at the time of Twitter," in *Proceedings of S&P*. IEEE, 2012. [cited at p. 108]

[213] F. Beato, I. Ion, S. Čapkun, B. Preneel, and M. Langheinrich, "For some eyes only: protecting online information sharing," in *Proceedings of CODASPY*. ACM, 2013. [cited at p. 109]

[214] A. J. Feldman, A. Blankstein, M. J. Freedman, and E. W. Felten, "Social networking with frientegrity: privacy and integrity with an untrusted provider," in *Proceedings of USENIX*, 2012. [cited at p. 109]

[215] L. Yuan, "Privacy-friendly photo sharing and relevant applications beyond," Ph.D. dissertation, EPFL Switzerland, 2017. [cited at p. 109]

[216] M. Ziad, A. Alanwar, M. Alzantot, and M. Srivastava, "Cryptoimg: Privacy preserving processing over encrypted images," *arXiv preprint arXiv:1609.00881*, 2016. [cited at p. 109]

[217] J. He, B. Liu, D. Kong, X. Bao, N. Wang, H. Jin, and G. Kesidis, "Puppies: Transformation-supported personalized privacy preserving partial image sharing," *The Pennsylvania State University Technical Report, CSE–2015–007*, 2015. [cited at p. 109]

Index

Alexandra-Mihaela OLTEANU

EPFL IC ISC LCA1
BC 250 (BC Building)
Station 14
CH-1015 Switzerland

+41 21 69 34657
alexandramihaela.olteanu@epfl.ch
alexandra.olteanu@gmail.com
http://people.epfl.ch/alexandramihaela.olteanu?lang=en

PROFILE

Privacy and security researcher, specialized in quantifying privacy, identifying security and privacy threats and proposing privacy-enhancing solutions without loss of usability, notably in the context of interdependent and multi-subject data; machine learning enthusiast; former (and still) a software engineer at heart. A versatile scientist, passionate about many aspects of technology, but also about how and what technology can do.

WORK EXPERIENCE

Teaching at École Polytechnique Fédérale de Lausanne, Switzerland 2011 – 2019

Teaching assistant for the courses: "Privacy Protection", "Information Security and Privacy" (UNIL), "Algorithms", "Software Design Engineering", "Mobile Networks", "General Physics I", "Linear Algebra"

Volunteer at VP Bali, Indonesia April 2013 – July 2013

Help to set up a new volunteer program for teaching English and taught children ages 13 to 16

Research Assistant at National University of Singapore September 2012 – April 2013

Research assistant in the Communications and Internet Group

Programmer at MFS South Africa http://mfsafrica.com/ July – September 2011

Developed the backend of an online remittance portal for international payments to mobile money accounts across Africa http://www.mtnmmo.com. Responsibilities: functional specification review, design and implementation

Software Development Engineer at Microsoft – Redmond, WA, USA and Vancouver, BC, Canada March 2010 – February 2011
October 2008 – February 2010

Worked on a large-scale service - the Customer Experience Improvement Program - that collects and manages customer data for all Microsoft products. Ownership of existing modules (maintenance), as well as new ones (design and implementation)

Programmer at S.A.I.A research foundation (Solutions of Artificial Intelligence Applications) – Cluj-Napoca, Romania January 2006 – July 2008

Contribute to medical research projects on topics of artificial intelligence and data mining

Developer at Experter Sp. z o.o – Warsaw, Poland July – October 2007

Worked on “Smart 2”, a project management application. Responsible with the design of the database, implementation of the data access layer, design and implementation of the workflow engine

Trainee at ISDC – Cluj-Napoca, Romania Fall 2006 and Fall 2007

Designed and developed a website for a tourist agency, a student portal and a quiz engine

Programmer at Siemens PSE – Brasov, Romania July – September 2006

Research project exploring the migration of existing web sites from Tomcat Web Server to Websphere

EDUCATION

École Polytechnique Fédérale de Lausanne, Switzerland September 2013 – March 2019

Ph.D. in Computer, Communication and Information Sciences
Thesis: “Interdependent and Multi-Subject Privacy: Threats, Analysis and Protection”
Advisor: Prof. Jean-Pierre Hubaux, Laboratory for Communications and Applications 1, EPFL
Co-advisor: Prof. Kévin Huguenin, Information Security and Privacy Lab, UNIL-HEC Lausanne
Courses: “Games for Crowds and Networks”, “Advanced Topics in Algorithmic Game Theory and Mechanism Design”, “Privacy Protection”, “Cellular biology and biochemistry for engineers”

École Polytechnique Fédérale de Lausanne; National University of Singapore February 2011 – April 2013

MSc in Computer Science, Internet Computing Specialization; excellence scholarship throughout the program
Courses: PhD level course “Machine Learning”, PhD level course “Recommender Systems”, “Pattern Classification and Machine Learning”, “Computational Molecular Biology”, “Distributed Information Systems”, “Intelligent Agents”, “Advanced Algorithms, Cryptography”, “Computational Linguistics”, “Multi-agent Systems”, “Statistical Analysis of Genetic Data”, “Methods and Models for Random Networks”, “Computer Vision”, “Philosophy of Biology”
Master thesis at National University of Singapore, “Evolution of the Internet: An Economic Perspective”

University of Plovdiv, Bulgaria July 2007

Summer school on project management

ISDC – Cluj-Napoca, Romania 2006 – 2007

Oracle and PL/SQL training; .NET training

Workshops on consulting, project management, business development and soft skills

Extreme Programming workshop with Reinier Tang, CTO in ISDC B.V.

Babes-Bolyai University of Cluj-Napoca, Romania October 2004 – July 2008

BSc in Computer Science; bachelor thesis on Windows Communication Foundation (GPA 10/10; excellence scholarship)

HONORS AND AWARDS

LSRS "Student of the Year" award – Europe, university level December 2012

Winner of the LSRS (Romanian Students Abroad League) award, granted each year in collaboration with the Romanian government to the best Romanian students abroad (http://gala.lsrs.ro/gala-lsrs/gala-2013/castigatori/)

HC^2 – Swiss national programming competition April 2012

1st place in the "just for fun" category, 7th place overall (out of 42 teams)

Google Anita Borg Scholar April 2011

The aim of this scholarship is to award excellence among women studying Computer Science and to encourage them to become active role models and leaders. The scholarships were awarded based on the strength of the candidates' academic performance, leadership experience and passion for technology. (https://www.womentechmakers.com/scholars/previous)

Imagine Cup by Microsoft, Software Design section January – April 2008

Developed *"Ecotraveler"*, a framework for real time public transportation information made available to the general public via web, mobile apps or SMS. Finished in 2nd place at the university level competition; 3rd place at the national phase

Imagine Cup by Microsoft, Software Design section January – April 2007

Developed *"Globalpedia"*, a crowdsourcing educational application providing quick access to information. An objective (any place or item of interest) is tagged with a RFID label, to which information is associated. This can be then accessed and updated by anyone who owns a mobile device. Finished in 3rd place at the university phase

ECN-Sapientia international programming contest March 2007

Selected to represent Babes-Bolyai University; 2nd prize out of 21 teams competing

Campus Match of Imagine Cup, Algorithm Competition January 2007

3rd place in the international competition

Imagine Cup by Microsoft, Software Design section Spring 2006

Developed *"The end of stress as we know it"*, a platform aimed to help people eliminate stress through music, color and reflex-o-therapy. This project involved building a device that coordinates a web of motors, which massage energy centers in the sole. A stress assessment module determines the exact location, intensity and duration of the massage automatically

"Grigore Moisil" Algorithm Contest by "Babes-Bolyai" University April 2005

2nd prize (1st prize among the first-year students competing)

ACSL (American Computer Science League) 2003 – 2004

Attended the Worldwide Finals in Chicago in the Senior #3 Category; 2nd place out of 19 teams competing

Individual qualification, based on four international online contests, with a perfect score (1st place worldwide)

ACSL (American Computer Science League) 2002 – 2003

Attended the Worldwide Finals competing in the Intermediate #5 Category; obtained 5th place out of 18 teams competing

Individual qualification based on four international online contests, awarded 2nd place worldwide

PUBLICATIONS AND MAIN PROJECTS

A. M. Olteanu, K. Huguenin, M. Humbert, and J.-P. Hubaux. *The (Co)-Location Sharing Game*. The 19th Privacy Enhancing Technologies Symposium (PoPETs'19) https://infoscience.epfl.ch/record/218755?ln=en

A. M. Olteanu, K. Huguenin, I. Dacosta, and J.-P. Hubaux. *Consensual and Privacy-Preserving Sharing of Multi-Subject and Interdependent Data.* The Network and Distributed System Security Symposium 2018 (NDSS'18). DOI 10.14722/ndss.2018.23002. https://infoscience.epfl.ch/record/232563?ln=en

Media coverage

http://wp.unil.ch/hecimpact/kept-in-the-picture-a-consent-based-system-for-sharing-data-online/
https://www.ictjournal.ch/articles/2018-10-08/des-chercheurs-lausannois-inventent-un-systeme-empechant-la-publication-non

A. M. Olteanu, K. Huguenin, R. Shokri, M. Humbert, and J.-P. Hubaux. *Quantifying Interdependent Privacy Risks with Location Data*. IEEE Transactions on Mobile Computing 2016 (TMC'16). DOI 10.1109/TMC.2016.2561281. https://infoscience.epfl.ch/record/222783?ln=en

A. M. Olteanu, K. Huguenin, R. Shokri and J.-P. Hubaux. *Quantifying the Effect of Co-location Information on Location Privacy*. 14th Privacy Enhancing Technologies Symposium 2014 (PETS'14). DOI 10.1007/978-3-319-08506-7_10. https://infoscience.epfl.ch/record/198297?ln=en

Research project (4 months) on quantifying the loss of online privacy through Google AdSense behavioral advertising http://infoscience.epfl.ch/record/212899

Research project (4 months) focusing on constraint optimization problems: analysis and comparison of existing algorithms, implementing and evaluating the Asynchronous Forward Bounding algorithm http://infoscience.epfl.ch/record/188532

User study on recommender systems: design and validation of a psychometric model to understand the motivations of the subjects when adopting a recommender system

A. G. Floares, R. Badea, C. Floares, I. Coman, L. Neamtiu, V. Aron, *A. Olteanu*, M. Ciornei (ro). *Liver and Prostate i-Biopsy and the Related i-Scoring Systems.* Automation Computers Applied Mathematics Scientific Journal, vol 17, No 1, 2008, pages 63-68, ISSN: 1221-437X http://acam.tucn.ro/pdf/ACAM17%281%292008-abstracts.pdf

A. Floares, C. Floares, L. Neamtiu, *A. Olteanu*, R. Badea, I.Coman, V. Aron, M. Ciornei (ro). *RODES – Algorithm for Automatic Mathematical Modeling Complex Biological Networks via Knowledge Discovery in Data.* Automation Computers Applied Mathematics Scientific Journal, vol 17, No 1, 2008, pages 69-73, ISSN: 1221-437X http://acam.tucn.ro/pdf/ACAM17%281%292008-abstracts.pdf

TECHNICAL SKILLS

Programming Languages: Java, C#, C/C++, Python, Scala, Pascal, Matlab
Technologies: .NET Framework, ASP.NET, Web Services, HTML, CSS, XML, J2EE
Database Systems: Microsoft SQL Server, Oracle, MySQL
Software Development: Object Oriented Analysis and Design, Design Patterns
Version Control: Git, Svn
Fields of expertise and interest: Privacy and security, game theory and mechanism design, graphical models, Bayesian inference, machine learning, graph theory, algorithms, data mining, genetic algorithms

ORGANISATIONS

Vice president of the Romanian Student Association at EPFL February 2014 – December 2017

Our mission is to facilitate the integration of Romanian students in Switzerland and setup educational partnerships between the two countries, in collaboration with the Romanian consulate in Geneva

AIESEC member October 2006 – July 2008

The Centre for Health Policy and Public Health Cluj-Napoca, Romania March 2006 – July 2008

Microsoft Student Partner within the Microsoft Academic Program December 2005 – July 2008

ACM member October 2005 – present

FOREIGN LANGUAGES

Native Romanian • Fluent English (C2; Cambridge C1 Certificate grade A, TOEFL Internet Based Test score 117/120, GRE) • Advanced French (C1; B1 attestation from EPFL) • Intermediate Spanish (B1) • Intermediate Italian (B1) • Basic German (A1/A2)

Ce document a été imprimé au Centre d’impression EPFL,
imprimerie climatiquement neutre, certifiée myClimate.