Identity Inference as a Privacy Risk in Computer-Mediated Communication

Sara Motahari, Sotirios Ziavras, Richard P. Schuler, Quentin Jones
New Jersey Institute of Technology
{sg262, sotirios.g.ziavras, rps22, quentin.jones}@njit.edu

Abstract

New Web 2.0 applications, with their emphasis on collaboration and communication, hold the promise of major advances in social connectivity and coordination; however, they also increase the threats to user privacy. An important, yet under-researched privacy risk results from social inferences about user identity, location, and activities. In this paper, we frame the 'social inference problem'. We then present the results from a 292 subject experiment that highlights: 1) the prevalence of social inference risks; 2) people's difficulties in accurately predicting social inference risks; and 3) the relation between information entropy and social inference. We also show how to predict possible social inferences by modeling users' background knowledge and calculating information entropy and discuss how social inference support systems can be deployed that protect user privacy.

1. Introduction

Considerable advances in addressing privacy threats have been made in recent years in terms of computer and network security [29], user control mechanisms [12], privacy policies, and ethical considerations [17]. However, changes in the technological environment are creating numerous new and unaddressed risks to user privacy. A key reason for this is the collaborative and often invasive nature of new mobile and web applications. Such applications can give users the ability to leverage background knowledge about the social environment/context to make unwanted inferences.

The term inference as used in the privacy literature is the process of deducing unrevealed information as a consequence of being presented with authorized information. A well known example of the inference problem relates to an organization's database of employees [3], where the relation <Name, Salary> is a secret, but user *u* requests the following two queries: "List the RANK and SALARY of all employees" and "List the NAME and RANK of all employees." None of the queries violates the security requirement because they do not contain the secured < NAME; SALARY > pair; however, since employees' ranks appear in both pairs, the querier can use it to map employees' names to their salaries and infer their salaries. Inference is mostly known as a security threat to databases [10] and sometimes as a privacy risk in data mining [25]. Although the inference problem as a threat to database confidentiality is discussed in many studies, mobile social computing raises new classes of more complicated inferences which we call *social inferences*. Social inferences are inferences that result from using social applications and are about user information associated with these applications such as identity, location, activities, social relations, and profile information. Previous inference prevention methods are not enough to address social inferences since in this domain:

- Users are able to learn information outside the application (background knowledge) and use this information as a premise to make inferences. Therefore, in contrast to the name-rank-salary example, their premises are not limited to the information revealed by the application.
- Most social inferences are partial inferences not absolute inferences, i.e., they don't logically result from the premises as in the name-rank-salary example, but they can be guessed as a result of low information entropy; and
- The sensitivity of user information may have a dynamic nature based on the context, such as time and location.

While numerous social computing applications aim to address privacy concerns through access control [4, 12] (e.g., Facebook enables users to set privacy preferences), direct access control models fail to prevent unwanted inferences.

In this paper, we analyze the social inference problem in social computing systems theoretically and empirically. We present results of our study on subjects chatting with unknown partners and use our study results to show how identity inferences happen and what comprises users' background knowledge that leads to identity inferences. We will also show that while many subjects have anonymity concerns, they are not able to realize what level of anonymity they are really maintaining. We then propose prediction methods and solutions taking consideration that most inferences in social computing are partial inferences based on background knowledge.

2. Social Computing and Privacy Risks

Social computing is the act of using social software. Social software describes any type of software that connects users and supports social interaction and sharing data [21, 30]. A wide variety of social applications and mobile social computing applications exist; email and instant messaging are simple ways of on-line communications. Many social applications such as Facebook and Myspace enable users to exchange messages, reveal their profile items, and find profile-based matches. Some of them such as LoveGety leverage location and mobility to provide innovative services [13].

The use and sharing of geo-temporal (user location histories) and personal information raises many serious privacy concerns. Seven categories of potential privacy invasions in mobile social computing systems were introduced in a previous study [24]. They are:

1. **Inappropriate use by Administrators**: For example, a system administrator may sell personal data without permission.
2. **Legal Obligations**: For example, a system administrator may be forced by an organization such as the police to reveal personal data.
3. **Inadequate Security**: For example, the server is not protected against intrusions or wireless transmission through the air is not secured.
4. **Lack of Control over Direct Revelations (Poor Features):** For example, a cell phone application that reveals my location to my friends, but does this without properly informing me or giving me control of this feature.
5. **Instantaneous Social Inferences:** For example, when my cell phone shows that Bob is nearby and I can only see two people with a similar cell phone around me. One of them must be Bob, thus increasing my chance of identifying him.
6. **Historical Social Inferences** through persistent user observation. For example, two nicknames are repeatedly shown on the first floor of the gym where the gym assistant normally sits. One of them must be the gym assistant.
7. **Social Leveraging of Privileged Data**: For example, David can't access my location, but Jane can. David asks Jane my location.

Social inferences include the fifth and sixth category. They are inferences about user information associated with mobile social applications such as identity, location, activities, social relations, and profile information.

Based on information theory, as we collect more information about a user, such as his/her contextual situation, our uncertainty about other aspects such as his/her identity may be reduced, thus increasing our probability of correctly guessing these aspects. This uncertainty is measured by information entropy in information theory. As we will see, this entropy also depends on the number of entities (e.g., users) that match our collected information. Collected information includes not only the information that our system provides to users, but the information available outside of database or background knowledge. Therefore, social inferences happen when collected information reduces the inferrer's uncertainty about an attribute to a level that she/he could deduce that attribute's value. The social inference problem can include a wide range of issues. However, any inference that results from using social applications can be made in one of the following two ways: 1) the inferrer uses only the current state of the system. This type of inference is based only on the current observation of the system and is called *instantaneous inference* (our fifth category) or, 2) the inferrer uses the history of her/his observations (or the history of the answers to previous queries), which is called *historical inference*.

Based on the nature of mobile social applications, social inferences are either the result of accessing location-based information or the result of social communications, or both. We call the first type, *location-related inferences* and the second type *inferences in computer-mediated communications*. The following three scenarios illustrate this point: 1) *Instantaneous social inferences in online communications* - Cathy chooses a nickname for her profile and hides her real name, but her profile shows that she is a female football player. Since there are only a few female football players at my school, there is a high chance that I can identify her; 2) *Instantaneous location-related social inferences* - My cell phone shows few nicknames in a room and I know (or I find out from the school's website) that the room is Professor Smith's office. Therefore, there is a high probability that Prof. Smith is in his office and one of those few nicknames belongs to him; 3) *Historical location-related inferences* – Two users identified as Superman2 and Prof. Johnson are repeatedly shown in a room, which I know is Prof. Johnson's office. I also know (or I find out from the school's website) that David is his PhD student. Therefore, Superman2 is probably David and he is currently at Prof. Johnson's office.

3.0 Instantiated Privacy Management Solutions

Extensive research focuses on helping computer users protect their privacy. Researchers have looked at various aspects of privacy enhancement such as ethics of information management, system features, access control systems, and security and database confidentiality protection. We classify their effort into four sections as discussed below.

1. Ethics, principles, and rules: To properly respond to concerns of ethics and principles, and to protect the user privacy, researchers have made various suggestions. In particular, Langheinrich [17] defines the principles of fair information practices as openness and transparency, individual participation, collection limitation, data quality, use limitation, reasonable security, accountability, and explicit consent. He then sets principles for privacy in mobile computing, that consist of notice, choice, proximity, anonymity, security, and access.

2. Access control systems: Access control systems provide the user with an interface to set their privacy preferences. They directly control people's access to the user's information based on his privacy settings. Access control systems with an interface to protect user privacy started with internetworking. Later, they were extended to context-aware and then ubiquitous computing. The earliest work within this area is P3P [5]. P3P enables users to regulate their settings based on different factors including consequence, data-type, retention, purpose, and recipient. Ackerman [1] implemented critic-based agents, called Privacy Critics, for online interactions. These agents watch the user's actions and make appropriate privacy suggestions. Access control mechanisms for mobile and location-aware computing were introduced later [12, 15, 18, 27].

3. Security protection: Security protection handles the following aspects:

- Availability (services are available to authorized users).
- Integrity (free from unauthorized manipulation).
- Confidentiality (only the intended user receives the information).
- Accountability (actions of an entity must be traced uniquely).
- Assurance (assure that the security measures have been properly implemented).

The inference problem is mostly known as a security problem that targets system-based confidentiality. Confidentiality protection is the area that includes most of the previous research on the inference problem. Therefore, suggested solutions often deal with secure database design. There are also methods that evaluate the queries to predict any inference risks. Both methods are explained in the next section.

4. Inference management: Two different techniques have been proposed to identify and remove inference channels. One makes use of semantic data modeling methods to locate inference channels in the database design and then to redesign the database to remove these channels. The other technique evaluates database queries to understand whether they lead to unauthorized inferences. Each technique has its drawbacks, the former is vulnerable to false positives and negatives and denial of service attacks. The latter usually has a high computational complexity. Additionally, a dynamic social environment limits the usability of such systems because they can restrictively limit a user's access to information. Both techniques have been studied for statistical databases [19], multilevel secure databases [14, 28], and general purpose databases [3, 7]. A few researchers have also addressed this problem via data mining [26]. Since in the domain of mobile social computing user information and preferences are dynamic, the first technique cannot be used in such systems.

Denning and Morgenstern, pioneers in calculating the partial inference risk, employed classical information theory to measure the inference chance [8, 23]. Given two data items *x* and *y*, let *H(y)* denote the entropy of *y* and $H_x(y)$ denote the entropy of *y* given *x*, where entropy is as defined in information theory. Then, the reduction in uncertainty of *y* given *x* is defined as follows:

$$Infer(x \rightarrow y) = \frac{H(y) - H_x(y)}{H(y)}$$

The value of *Infer* (*x*→ *y*) is between 0 and 1, representing how likely it is to derive y given *x*. If the value is 1, then *y* can be definitely inferred given *x*. Denning and Morgenstern did not know how to use this formulation in real situations and they mention the serious drawbacks of using this technique [8], Firstly, it is difficult, if not impossible, to determine the value of $H_x(y)$; secondly, the computational complexity that is required to draw the inference is ignored [8].

Nevertheless, this formulation has the advantage of presenting the probabilistic nature of inference (i.e., inference is a relative not an absolute concept).

Although the inference problem includes a wide range of issues, studies and polls suggest that identity is the most sensitive piece of users' information [24] and anonymity preservation is an important aspect of application designs and research [11, 16]. Anonymity is defined as *"not having identifying characteristics* *such as a name or description of physical appearance disclosed so that the participants remain unidentifiable to anyone outside the permitted people promised at the time of informed consent"* [9]. Recently, new measures of privacy called *k*-anonymity and L-diversity have gained popularity [20, 30]. *k*-anonymity is suggested to manage identity inference, while L-diversity is suggested to protect both identity inference and attribute inference in databases. In a *k*-anonymized dataset, each record is indistinguishable from at least *k*−1 other records with respect to certain "identifying" attributes. These techniques can be broadly classified into generalization techniques, generalization with tuple suppression techniques, and data swapping and randomization techniques. Nevertheless, *k*-anonymized datasets are vulnerable to many inference attacks including attacks based on background knowledge. Identity inferences in next generation mobile social computing systems cannot be addressed by the above techniques because:

- The sensitivity of user information is dynamic in nature based on the context, such as time and location.
- Information such as life patterns, physical characteristics, and the quality of social relations that are not kept in the database can be inferred from information available to the user. Therefore, inferences in such systems are not limited to attribute disclosures.
- Users' background knowledge (the information users learn outside the database) is a premise in many inferences.

We will demonstrate how by assuming an ideal perfect model of the background knowledge in mobile social computing systems and focusing on anonymity protection, our modeling and use of inference chance based on information entropy can sometimes simplify into a dynamic *k*-anonymity problem.

4. Aims and Research Questions

An important role of social applications is to help users enhance their social network and make new social ties. Social networking sites such as Facebook and Orkut allow you to view users' profile information based on their privacy settings and to exchange messages. A popular way of making new social ties is social matching [31]. For example, friend-of-friend systems provide users access to various parts of their friends' profiles [2]. In existing social matching systems such as OKCupid and Orkut users usually utilize a combination of synchronous and asynchronous private messages, profile comments, and friend requests to move the introduction forward. While messaging and exchanging profile items help the introduction process, revealing profile items during this introduction increases the chance of social inferences and in particular identity inference, which lead to anonymity violation. We were interested in answering the following questions regarding the social inference risks associated with social computing applications:

- How and how often do users make social inferences?
- What are the best predictors of the chances of a social inference?
- What is the impact of conversation parameters and user interface on the chance of a social inference?
- What level of anonymity is desired by the users and are they able to reach that?
- What background knowledge do individuals routinely use to make social inferences?

As a first step in addressing these questions we conducted a user experiment exploring identity inferences in computer-mediated communication between unknown chat partners.

5. Method

5.1. Subjects

All subjects were recruited through desks situated in the campus center of a medium sized urban university with the offer of $20 for: 1) entering a personal profile online; 2) participating in a short chat experiment; followed by 3) completing a short post chat survey. Five hundred and thirty students completed the profile entry portion, 304 participated in the chat session of which 292 subjects completed all three study components. Subjects were exclusively university students representative of the various majors offered on campus and ranging from 18 to 57 years old. Twenty three percent of the subjects were female, and 66% of the subjects were commuters (living off campus).

5.2. Procedure

Phase I – Online Profile Entry – The online profile consisted of 67 individual profile fields (items), clustered into 5 broad categories (basic information, personal information, education information, contact information, and interests). Completing the profile entry web page took subjects approximately 10 minutes.

Phase II – Chat Experiment – Subjects that completed the profile entry were able to select various time slots to participate in the chat experiment. These subjects were then randomly assigned to one of four experimental conditions (described below) differentiated by the introduction mechanism used.

Phase III – Post Chat Survey – After completing the chat session subjects answered a survey which asked about their chat partner, the level of anonymity they had desired and achieved at the end of the chat session and their willingness/desire to meet their chat partner.

5.3. Introduction Mechanism

Participants used a custom-developed software application designed to aid in communication and exchange of personal profile information. The application was exclusively intended for use in a larger study on strangers' communication, of which the study described in this paper, was a part.

After both participants authenticate, each participant sees his/her own profile on the left side of the screen, a chat box for typing in the center, and his/her chat partner's profile (all information hidden by default) on the right side, and action buttons (e.g., end chat, etc.) at the center bottom.

By default, a participant's profile information is completely hidden from her/his partner and vice versa. To reveal a field in her/his profile to her/his partner, the participant simply clicks on the desired field. When a participant reveals a field, it appears on the right side of her/his partner's screen next to the appropriate field in a yellow highlight on both screens. Similarly, a participant can request parts of his/her partner's profile by clicking on the item which makes it appear in a green highlight. Participants also have the option of requesting and revealing whole sections of a profile by clicking on each of 5 broad category labels.

Depending on the condition, participants could perform 16 actions such as reveal, request, chat, or use the action buttons.

5.4. Experimental Conditions

Condition 1 – Chat Only (75 subjects) - In this condition, profiles are disabled and although subjects can see their own profile on the left side of the screen, they cannot reveal and request profile items. They can only chat by typing in the chat box.

Condition 2 – Introduction Mechanism only (72 subjects) - In this condition, participants could only click on profile fields to reveal or request them, but could not type in the chat box and use it for a regular chat.

Conditions 3 (63 subjects) and 4 (82 subjects) - Chat and Introduction Mechanism - Conditions 3 and 4 enabled participants to use both profiles and the simple typing mechanism of the chat box. They could reveal their previously entered profile information by clicking on it or typing it, they also had the option of general text based chat where it was possible for them to reveal any kind of information. In condition 3, revealed profile fields of a subject's partner would disappear from the subject screen after 5 seconds and to see it again, the subject had to click on it, or scroll the middle box up to find it. However, in condition 4, revealed information stayed visible until the end of the session.

5.5. Post Experimental Chat Survey

Following questions related to social inferences were asked in the post chat survey.

1. Without your chat partner revealing her/his complete name, do you think you could work out your partner (such as guess his name or remember his physical appearance)?

○ *I am able to guess exactly who my partner is now, although s/he didn't introduce.*

○ *I was able to guess exactly who my partner was during the chat and I was right.*

○ *I guessed exactly who my partner was during the chat, but I was wrong.*

○ *I was/am able to guess that I had seen my partner before, but I only knew her/his physical appearance.*

○ *I was/am able to narrow my partner down to a few possible individuals before a complete introduction, but I didn't know exactly who s/he was. (If you're able to narrow her/him down to a few people, how many people would that be?)*

2. Please elaborate on the details you deduced or guessed about her/his identity without your partner revealing it. If you guessed some possible names, please mention the names or if you think you have seen her/ him, elaborate on her/his physical characteristics such as height, weight, facial features, etc.

3. How did you make the above guesses on the person's identity? What knowledge did you use to do this?

4. After this conversation, given enough time and other information sources available to you, how well do you think you can identify your chat partner?

5. Please mention all the information sources you may use to do so.

6. How anonymous did you want to be to your chat partner before you decided to introduce yourself?

7. What does your partner know about you?

In questions 4, 6, and 7, subjects were given options similar to question 1.

5.6. Dependent and Independent Variables

In addition to answers directly collected from the survey, the following parameters were calculated for each subject:

- *Duration* of the introduction phase (chat);
- Software *condition* used;
- Number of profile items revealed;
- Profile items revealed in the conversations. All information revealed was coded into the field number. Fields revealed by using the mechanism were automatically coded and typed information was hand coded.
- *Name-revelation*: defined as 1 if the complete name was revealed;
- *Immediate-introduction*: defined as 1 if complete name was the first thing to be revealed;
- *Maintained degree of anonymity*: the number of people to who the subject's partner could correctly narrow him down. For example, if subject ***A*** can correctly narrow subject ***B*** down to 2 individuals, subject ***B***'s maintained degree of anonymity will be 2;
- *Desired degree of anonymity*: indicates what degree of anonymity a subject wished to maintain. For example, if subject ***B*** does not want to be exactly identified, her desired degree of anonymity will be 2;
- *Perceived degree of anonymity*: shows what degree of anonymity the subject thought he maintained;
- *Estimated degree of anonymity*: A measure of information entropy, which shows our estimation of what degree of anonymity a subject really maintained based on what they revealed. We estimated this variable based on profiles of 530 students that registered for the study. Details of calculation will be explained in section 7.2; and
- *Identity inference incidents*: to model the inference risk, we defined this variable, which was set to 1 if subjects could narrow their partner down to less than 50 people.

6. Results

6.1. Identity Inferences

Eighty percent of the subjects said they had no idea who their partners were before a complete introduction. Among the 20% who said they could guess something, 55% were correct (11% of total subjects) and the rest were wrong or not clear enough about what they inferred. In addition, 8.5% of the subjects guessed exactly their partner's identity. Fifty-six percent of the subjects said that having access to information sources, they could at least narrow their partner down to a few individuals.

As expected, all inferences happened when a subject's revealed information combined with their partner's background knowledge resulted in low information entropy which led to uniqueness of revealed information. For example, a subject revealed her gender, ethnicity, and group memberships to be female, Hispanic, and soccer team member, respectively. Her partner, who had seen the women's soccer team playing, knew there was only one Hispanic player on the team and was able to infer who she was. Other examples included revelation of gender, ethnicity, and courses taken (white girl attended a chemistry class), ethnicity and occupation (Indian resident assistant in the dormitories), etc.

6.2. Predictor of the Identity Inference

A binary stepwise Logistic Regression test was performed on the *identity-inference-incident* (dependent variable) and the independent variables *duration, condition, number of revealed items, name-revelation, immediate-introduction, desired-degree-of-anonymity, perceived degree of anonymity*, and *estimated degree of anonymity*. Only *estimated degree of anonymity* showed was a reliable predictor of the *identity inference incident* (Wald χ^2=55.1, degree of freedom: df=8, p=0.016). This suggests that our *estimated degree of anonymity* is the only strong predictor of social inference incidents.

As we will see in section 7.2, in a simplistic case, identity entropy of person ***B*** in this study is determined by *estimated degree of anonymity*. Therefore, our estimation of information entropy was the best predictor of the chances of a correct social inference.

6.3. The Impact of Conversation Features

Although more inferences were made in condition 1 (chat-only) than in condition 2 (introduction mechanism-only), no significant statistical difference was found between experimental conditions. Users reveal many more profile fields in condition 2 and less in condition 1 (the mean number of items revealed was 4.39 for condition 1 and 31.58 for condition 2). This is probably because profile exchange and profile revelation requests are the only actions possible in condition 2. However, as we saw in section 6.2, revealing more profile items is not necessarily equivalent to maintaining less entropy and revealing more identifiers. A Spearman correlation test shows a small correlation between the number of revealed items and our *estimated degree of anonymity* (n=288, Spearman rho=0.21, p=0.005). There was no significant correlation between the *maintained degree of anonymity* and number of revealed items (n=288, Spearman's rho =0.018, p=0.76).

6.4. Subjects' Ability to Maintain Their Desired Degree of Anonymity

The level of anonymity desired varied greatly, 72% of the subjects who had anonymity concerns did not want to be exactly identified by their name or face, 6.3% of them did not want to be narrowed down to two people or less and the rest desired a higher degree of anonymity.

Although *perceived degree of anonymity* for a subject and *maintained degree of anonymity* for their partner had a small correlation (n=254, Spearman's rho=0.155, p=0.031), they were equal only 48% of the time. This means 52% of participants could not guess what their partner knew about their identity. Spearman correlation test showed a small correlation between *estimated degree of anonymity* and *maintained degree of anonymity*. However, *estimated degree of anonymity* was smaller than *maintained degree of anonymity* for 20% of the subjects, which means 20% of the subjects revealed identifiers that put them at the risk of *unwanted* identity inference. Fifteen percent of the subjects wished to be anonymous even though they revealed their full name during the conversation.

6.5. Sources of Background Knowledge

During their chats subjects did not have access to search tools, personal profiles, or other information sources. Hence, they only inferred their partners' identities if revealed information helped them realize that they had seen their partner before or if they already had enough knowledge to be able to narrow their partner down to a few people.

When subjects were asked to mention the information sources they would use, they mentioned sites that include personal profiles such as Facebook, Myspace, and Orkut. Other sources included the university directory and website, Yahoo and AIM to search email addresses, and phone directories to search for phone numbers.

We also noticed that subjects tend to guess their partners' gender and home country from their "chat style". Subjects' responses show that the chat style includes the way someone answers a question, slang and abbreviations they use, and their English language fluency. When participants where asked to elaborate on details they guessed about their partner's identity, 11% percent of them guessed their partners' gender and 6% guessed their partners' country/region. They were correct around 94% of the time about gender and 87% of the time about home country. This result suggests that gender and region are two types of probable background information with probabilities of 10.4% and 5.2% respectively.

7. Social Inference Prediction and Theoretic Frame

The results of our study suggest the need for an inference control system in social computing. We saw that social inferences happen as a result of low information entropy. An inferrer usually reaches this low entropy by combining the information he receives from the application combined with his background knowledge.

Cuppons and Trouessin [6] formulate inference control as follows;
If ***A*** is permitted to know Q and ***A*** can do the derivation (Q=>Φ), then ***A*** should be permitted to know Φ. Consequently, if we want Φ to be forbidden for ***A*** and Φ can be inferred from Q, Q should be forbidden for ***A*** as well.

To understand what determines Q => Φ, we need to remember that considering partial inferences in a mobile social computing system, Φ may not be logically deduced from Q as indicated by Q => Φ, but, as Morgenstern mentions, Φ may belong to the Sphere Of Influence of Q ($SOI(Q)$) [23]. We define that in a mobile social computing system, Φ belongs to the Sphere Of Influence of Q ($\Phi \in SOI(Q)$) if knowing Q reduces the uncertainty about Φ enough to make an inference. Accordingly, we modify Cuppons' formulation as follows: We define Q to be the information included in the query, its answer, and background knowledge that is modeled as described below. Q is safe to be completely known by A if $\forall\Phi$, $[(\Phi \in SOI(Q) \wedge PK_A(Q)) \Rightarrow PK_A(\Phi)]$, which means ***A*** is permitted to know Q only if he is permitted to know everything in its Sphere Of Influence).

7.1. Modeling Background Knowledge

Most previous inference control frameworks are vulnerable to attacks based on background knowledge. Background knowledge is the information available to users outside the database. This information should be assumed to be known by all the users just like answers to their queries are assumed to be known by them. Thuraisingham [32] points out that to preserve database integrity, sometimes we need to model user knowledge in the outside world. As Jajodia and Midows [14] say, "We have no way of controlling what data is learned outside of the database, and our abilities to predict it will be limited. Thus, even the best model can give us only an approximate idea of how safe a database is from illegal inferences".

The results of our user study suggest that subjects not only use searching profiles and the campus directory (which is an organization's website) to identify their chat partner, but also tend to guess their partner's gender and home country from what they usually term as their "chat style". We categorize background information associated with computer-mediated communication as follows: the organization's website and public information, personal profiles, phone and address directories, gender, and home country

In the examples illustrated in section 2, we showed how in a location-aware system there exist meanings and implications associated with location, such as visual information of nearby people, ownership, common usage, and purpose of places, etc. That indicates background knowledge associated with location which includes the official manager of the place and people related to her/him, official use and schedule of the place, and visual information of the inferrer's vicinity.

Having all the information modeled in *Q*., we can estimate *SOI*(*Q*).

7.2. Calculating the Entropy for Instantaneous Inferences (Category 5)

As mentioned in Section 3, Morgenstern [8, 22] formulated for the first time partial inferences based on the entropy of information, but he didn't know how to calculate it in the general sense. We modified his formulation for social computing applications to calculate the instantaneous inference risk.

We want to calculate the risk that an attribute Φ is inferred from revealed information, *Q*. We define the inference function as follows:

$$INF1\ (Q \to \Phi) = \frac{H_{\max} - H_c}{H_{\max}}\ , \qquad \textbf{(1)}$$

where $H_{\max}$ represents the maximum information entropy for the environment and is fixed for any given application; H_c is the information entropy under the current conditions and is dynamic based on the situation. $H_{\max}$ is calculated as follows:

$$H_{\max} = -\sum_{1}^{X} P.\log_2 P\ , \qquad \textbf{(2)}$$

where P=1/X and X is the maximum number of entities (users) related to the application. H_c is calculated as follows:

$$H_c = -\sum_{i=1}^{V} P1(i).\log_2 P1(i)\ , \qquad \textbf{(3)}$$

where V is the number of possible values for attribute Φ. *P*1(i) is the probability that the ith possible value is thought to be the correct one by the inferrer. If the inferrer didn't know the collected information, *Q*, Φ could be any of the possible values with the prior probability of 1/ *X*. This results in maximum entropy, $H_{\max}$. However, after having access to *Q*, *P*1(i) is the posterior probability of each value and equals *P* given *Q*. When *INF1* is too high, say larger than *C*, an appropriate action needs to be taken. The appropriate action can be rejecting the query, blurring the answer, or sending a warning to the owner of the information. We will provide further examples to explain this model.

For example to calculate the identity inference risk of user ***A*** in an on-line communication similar to our above study:

Q: (profile items matching ***A***'s profile items that are already revealed).

Φ_1: Partner's identity at name or face granularity.

X: total number of potential users of the application.

V: number of users that satisfy *Q*.

We define:

Group F: users that are the same sex as ***A*** and satisfy *Q*.

Group G: users that come from the same country/region as ***A*** and satisfy *Q*.

*X*1: number of users in group F.

*X*2: number of users in group G.

*X*3: number of users in the intersection of F and G (F∩G).

ς: probability of guessing the right gender from the partner's chat style (which was shown to be 10.4% in our user study).

σ: probability of guessing the right home country from the partner's chat style (which was shown to be 5.2% in our user study).

- If users use their guesses on gender and home country as their background knowledge:

$$P1 = \begin{cases} \varsigma.\sigma/X3 + \varsigma.(1\text{-}\sigma)/(X1) + (1-\varsigma).\sigma/(X2) & \text{for F}\cap\text{G} \\ \varsigma.(1\text{-}\sigma)/(X1) + (1-\varsigma).(1\text{-}\sigma)/V & \text{for F-F}\cap\text{G} \\ (1-\varsigma).\sigma/(X2) + (1-\varsigma).(1\text{-}\sigma)/V & \text{for G-F}\cap\text{G} \\ (1-\varsigma).(1\text{-}\sigma)/\mathrm{V} & \text{for the rest of users} \end{cases} \qquad \textbf{(4)}$$

  This means that users of the same gender and ethnicity (F∩G) have the highest probability and the users, who don't fall in F or G have the lowest probability.

- If there are no guesses involved and background knowledge only includes deterministic information (which is publicly available)

$$P1 = 1/\mathrm{V} \qquad \textbf{(5)}$$

In the second case, according to equations 6 and 3, *Hc* equals Σ(1/**V**).log(1/**V**). Therefore, if we assume that all the information available to the users is deterministic (which means if they are able to access the information source, they are either able to know the exact answer or not) and assume that all information available outside the database is included in *Q*, for any given application and known condition, *INF*1 is only determined by the number of users satisfying that condition (**V**). That is how we calculated *estimated degree of anonymity* in section 6 (number of users who share the same revealed profile items). We used it to perform the regression test in our user study and we saw that *estimated degree of anonymity* was the only strong predictor of the inference chance.

Furthermore, to have an *INF*1 value smaller than the associated threshold in the second case, at least **U** indistinguishable users are needed in the situation. This satisfies *k*-anonymity with *k*=**U**. However, our entropy control method is more general than the *k*-anonymity solution. In particular, it can model a probabilistic model of background knowledge, such as guesses on gender and home country. It can also be used to calculate the risk of inferring other attributes such as location.

7.3. Calculation of Entropy for Historical Inferences (Category 6).

Our model in the previous category applies here as well, but in this category *Q* also includes answers to past queries. The inference chance will be calculated from an inference formula similar to (1), as described below.

$$INF_i2(Q \to \Phi) = \begin{cases} \dfrac{H_{\max}-H_i}{H_{\max}}, if & \dfrac{H_{\max}-H_i}{H_{\max}} < C \\ \lambda \qquad\quad , if & \dfrac{H_{\max}-H_i}{H_{\max}} \geq C \end{cases} \quad \textbf{(6)}$$

where

$H_{\max} = -\sum_{1}^{Y} P.\log_2 P$. ($H_{\max}$ is again fixed for each application.)

λ = number of queries that involve the attribute and were sent after $(H_{\max} - H_i)/H_{\max}$ reached the threshold value *C*.

$H_i = -\sum_{1}^{V} P1(i).\log P1(i)$,

where *P*1 is the probability that a value is thought to be the correct attribute value by the inferrer. **V** is the number of possible values repeated in previous queries starting at the current time and going back an amount of time equal to T(given).

In this formulation, *INF$_i$2* also considers the history of the queries. Unlike *INF1*, *INF$_i$2* is calculated for any time slot *i*. *INF$_i$2* is between 0 and 1 until ($H_{\max}$-H_i)/ $H_{\max}$ reaches the threshold value C. When *INF$_i$2* equals C, entropy of Φ is too low and sending multiple queries involving Φ can lead to an inference. Therefore, at this time *INF$_i$2* starts counting the new queries. After *INF$_i$2* passes a number of queries, say K, the system takes an appropriate action such as dynamic blurring. We consider discrete finite duration T for past queries since humans don't have a perfect memory. Thus, we assume they forget the answers to queries sent more than T time units ago. However, to protect the system against inference attacks, the calculation and results can be extended for $T \rightarrow \infty$. Obviously, if time slots *i* and *j* overlap, rejecting a query based on *INF$_i$2* affects the value of *INF$_i$2*.

8. Conclusion

This article highlights the significant privacy threats associated with the use of social software resulting from social inference. The nature of this 'social inference problem' was illustrated through a 292 subject experiment, which showed that:

- Identity inferences are frequently made in on-line communication;
- Different users desire different levels of anonymity;
- Even when users are in complete control of the information they reveal, they are not able to maintain their desired degree of anonymity. This is because individuals are unable to correctly judge inference risks; and
- Information entropy, calculated by construction of an idealized model of background knowledge, was the only measure found to strongly predict identity inference;

These results validate our methods for calculating information entropy. They also highlight the need for social inference protection systems. Such protection systems could improve user inference-risk judgments through techniques, such as risk visualizations or warning messages. Alternatively inference protection systems could modify information exchange, for example, by lowering the granularity of revealed information.

9. Acknowledgments

This research is partially supported by the National Science Foundation Grant NSF IIS DST 0534520 and NSF CNS 0454081. The opinions expressed are those of the authors and may not reflect those of the NSF.

10. References

1. Ackerman, M. and Cranor, L., Privacy Critics: UI Components to Safeguard Users' Privacy. in *ACM Conference on Human Factors in Computing Systems (CHI99)*, (1999).

2. Boyd, D., Friendster and publicly articulated social networking. in *Conference on Human Factors and Computing Systems*, (Vienna, April 24-29, 2004), ACM.

3. Brodsky, A., Farkas, C. and Jajodia, S. Secure databases: constraints, inference channels, and monitoring disclosures. *IEEE Transactions on Knowledge and Data Engineering*, *2* (6). 900-919.

4. Cornwell, J., Fette, I., Hsieh, G., Prabaker, M., Rao, J., Tang, K., Vaniea, K., Bauer, L., Cranor, L., Hong, J., McLaren, B., Reiter, M. and Sadeh, N., User-controllable security and privacy for pervasive computing. in *WMCSA '07: Proceedings of the 8th IEEE Workshop on Mobile Computing Systems & Applications*, (2007).

5. Cranor, L., Langheinrich, M., Marchiori, M. and Reagle, J. The Platform for Privacy Preferences 1.0 (P3P1.0) Specification *W3C recommendation*, 2002.

6. Cuppens, F. and Trouessin, G. Information Flow Controls vs Inference Controls: An Integrated Approach *Third European Symposium on Research in Computer Security (ESORICS 94)*, 1994.

7. Dawson, S., Capitani, S.D. and Samarati, d.V.P. Specification and Enforcement of Classification and Inference Constraints *IEEE Symposium on Security and Privacy*.

8. Denning, D.E. and Morgenstern, M. Military database technology study: AI techniques for security and reliability *SRI Technical report*, 1986.

9. Doyle, D. Trans-disciplinary Inquiry – Researching with Rather than Researching on. in Campbell, A. and Groundwater-Smith, S. eds. *An Ethical Approach to Practitioner Research: Dealing with Issues and Dilemmas in Action Research*, Routledge 2007.

10. Farkas, C. and Jajodia, S. The inference problem: a survey. *SIGKDD Explorer Newsletter*, *4* (2). 6-11.

11. Heinrich, E. Electronic Repositories of Marked Student Work and their Contributions to Formative Evaluation. *Educational Technology & Society,*, *7* (3). 82-96.

12. Hong, D., Yuan, M. and Shen, V.Y., Dynamic Privacy Management: a Plugin Service for the Middleware in Pervasive Computing. in *ACM 7th international conference on Human computer interaction with mobile devices & services* (2005), 1-8.

13. Iwatani, Y. Love: Japanese Style, http://www.wired.com/culture/lifestyle/news/1998/06/12899.

14. Jajodia, S. and Meadows, C. *Inference Problems in Multilevel Secure Database Management Systems*. IEEE Computer Society Press, Los Alamitos, California, USA 1995.

15. Jendricke, U., Kreutzer, M. and Zugenmaier, A. Pervasive Privacy with Identity Management *Workshop on Security in Ubiquitous Computing - Ubicomp*, 2002.

16. Kobsa, A. and Schreck, J. Privacy through pseudonymity in user-adaptive systems. *ACM Transactions on Internet Technology (TOIT)*, *3* (2). 149 - 183

17. Langheinrich, M., Privacy by Design – Principles of Privacy-Aware Ubiquitous Systems. in *Third International Conference on Ubiquitous Computing (UbiComp 2001).* , (2001), 273-291.

18. Lederer, S. Designing Disclosure: Interactive Personal Privacy at the Dawn of Ubiquitous Computing *Computer Science Division*, University of California at Berkeley, 2003.

19. Lunt, T.F. Current Issues in Statistical Database Security. *IFIP Transactions, Results of the IFIP WG 11.3 Workshop on Database Security V: Status and Prospects A-6*. 381-385.

20. Machanavajjhala, A., Gehrke, J. and Kifer, D., ℓ-Diversity: Privacy Beyond k-Anonymity. in *Proceedings of the 22nd IEEE International Conference on Data Engineering (ICDE 2006)*, (2006).

21. Madan, A. and Pentland, A.S. VibeFones: Socially Aware Mobile Phones *10th IEEE International Symposium on Wearable Computers* 2006.

22. Morgenstern, M., Controlling Logical Inference in Multilevel Database Systems. in *Proc of IEEE symp on security and privacy*, (1988), 245-255.

23. Morgenstern, M., Security and Inference in Multilevel Database and Knowledge-Base Systems. in *Proceedings of the 1987 ACM SIGMOD international conference on Management of data* (1987).

24. Motahari, S., Manikopoulos, C., Hiltz, R. and Jones, Q., Seven privacy worries in ubiquitous social computing. in *ACM International Conference Proceeding Series; Proceedings of the 3rd symposium on Usable privacy and security* (2007), 171-172.

25. Narayanan, A. and Shmatikov, V., Obfuscated Databases and Group Privacy. in *12th ACM conference on Computer and communications security* (2005), 102-111.

26. O'Leary, D.E. Some Privacy Issues in Knowledge Discovery: The OECD Personal Privacy Guidelines. *IEEE Expert: Intelligent Systems and Their Applications 10* (2). 48-52.

27. Osbakk, P. and Ryan, N., Context, CC/PP, and P3P. in *UbiComp 2002 Adjunct Proceedings*, (Göteborg, Sweden, 2002).

28. Stachour, P.D. and Thuraisingham, B. Design of LDV: A Multilevel Secure Relational Database Management. *IEEE Transactions on Knowledge and Data Engineering*, *2* (2). 190-209.

29. Stallings, W. *Cryptography and Network Security Principles and Practices*. Pearson Prentice Hall.

30. Sweeney, L. Achieving k-Anonymity Privacy Protection Using Generalization And Suppression. *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, *10* (5). 571-588.

31. Terveen, L. and McDonald, D. Social Matching. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *12* (3). 401 - 434.

32. Thuraisingham, B. privacy constraint processing in a privacy-enhanced database management system. *data and knowledge engineering*, *55* (2). 159-188.