Towards Location Privacy Awareness on Geo-Social Networks

Fatma Alrayes
Cardiff School of Computer Science and Informatics
Cardiff University
Email: AlrayesFS@cardiff.ac.uk

Alia Abdelmoty
Cardiff School of Computer Science and Informatics
Cardiff University
Email: AbdelmotyAI@cardiff.ac.uk

***Abstract*—With the current trend of embedding location services within social networks, an ever growing amount of users' spatiotemporal tracks are being collected and used to generate user profiles. Issues of personal privacy and especially those stemming from tracking user location become more important to address. In this work, it is argued that support of location privacy awareness within social networks is needed to maintain the users' trust in their services. Current practices of pre-configuring location disclosure settings have been shown to be limited, where users' sense of location privacy dynamically change with context. In this paper, location privacy awareness is considered within a composite view of place, time and social data recorded in user profiles. The paper examines the possible threats to personal privacy from exposure of this data and the design of feedback tools to allow users to control their privacy. A user study is used to examine the impact of the feedback provided on users' perception of privacy and the link between their privacy concerns and their attitude towards using the geo-social network. Findings confirm the strong need for more transparent access to and control over user location profiles, and guide the proposal of recommendations to the design of more privacy-sensitive geo-social networks.**

***Keywords*—Location privacy; Privacy awareness; Geo-social networks; Usable privacy**

I. Introduction

Users are concerned about online privacy. This is a fact that has now been tested many times in recent years [1], [2]. Our interactions on the web and on social media are used for making predictive inferences about our personality and what we might want to buy, read, or listen to[1]. Our whereabouts is another layer of information that is now being carefully captured and added to our records. We are "switching on" location on our devices to tell our friends where we are and to search for the nearest good restaurant, but are we paying the price with our privacy? The importance of sharing our location data online is that it pulls together our virtual and physical existences, and thus raises critical questions about privacy in both worlds.

People trust the social networks services they use [3], they mostly do not read terms and conditions (and privacy policies) [4], and find privacy settings increasingly difficult to manage [5]. With the continuous stream of privacy-intrusion alerts, and the reality of current demands of surveillance of online information[2], people will eventually come to question their trust in social networking services, and may tend to be more rigorous and stringent with their online sharing behaviour. However, this trend may be reversed if individuals were able to recognise the information recorded in their online profiles and were given opportunities to control that information and its visibility.

In this paper we focus on user content awareness and in particular location content awareness in relation to privacy on Geo-Social Networks (GeoSNs)- online social networks with location-sharing facilities. The map of locations in a user's profile can be used as a key to define and relate other types of data elements, for example, their interests and activities can be related to the places they visit. Along with temporal semantics, the profile can provide a rich resource of personal and possibly sensitive information. The paper examines the dimensions of location data sharing on GeoSNs and how user awareness is situated in this information space. Levels of threat to privacy are proposed as means of assigning value to the information in this space. To enable location content awareness, design of feedback tools is considered that captures user's attention to both the content they are sharing and to the associated risks.

A user study, in the form of an online survey, is carried out to test the implications of location content awareness on user privacy concerns as well as on user's attitude and behaviour on geo-social networks. Results generally confirm initial assumptions with respect to the limitations of users' awareness and their need for more transparent access to their profile content. In addition, it was also clear that users need actionable feedback that allows them to control their content. Results also suggest, that transparency of content and the opportunity of control over the content may neutralise the adverse effects of privacy concerns and lead to more trust in the geo-social networking services. These findings are useful for designing more effective privacy-sensitive geo-social networks.

Few previous studies addressed location awareness and privacy concerns [6], and these were mainly questioning users' attitude with respect to sharing their current location information. These studies neglected revealing the privacy implications in terms of what personal information can be extracted based on location disclosure using limited location-sharing applications for their evaluations, which can be insufficient for reflecting the public GeoSNs features. This paper contributes a more detailed study that considers awareness with respect to extended user profiles on the space, time and social dimensions and provides an understanding of how users' perception of their location content influences their privacy concerns and

[1] www.eff.org/deeplinks/2009/09/online-trackers-and-social-networks

[2] www.theguardian.com/world/2014/jun/17/mass-surveillance-social-media-permitted-uk-law-charles-farr

2161-2897/16 $31.00 © 2016 IEEE
DOI 10.1109/NGMAST.2016.26

behaviour on GeoSNs. To the best of our knowledge, this is the first study to address location privacy in public GeoSNs through content awareness approach.

II. Related Work

Our work is related to two lines of research; methods for location-based inferences from mobility data sets to examine the potential sort of information user profile on GeoSNs, and the design of feedback to enable privacy awareness on these networks.

A. Location-Based Inference in GeoSNs

A significant interest can be witnessed in research studying the value and utility of location information on GeoSNs to understand users' behaviour. Some studies looked into accurate identification of user's location from their GPS trails [7], as well as using the location of the user's friends on Twitter [8]. Using the user's profile of visited places and socio-historical ties, Gao et al. demonstrated an accurate prediction of next check-in information [9].

Other works investigated the potential inference of social relationships between users of GeoSNs. For instance, co-occurrence between users, extracted from geo-tagged Flickr pictures, was sufficient for deducing their social ties with high probability [10]. In addition, users' interactions on the network coupled with their location information were used to predict friendship links [8].

Extracting spatiotemporal movement and activity patterns of users on GeoSNs attracted much research in recent years. Dearman et al. exploited location reviews on Yelp in order to identify a collection of potential activities promoted by the reviewed location [11]. Mobility patterns on Foursquare was the subject of study in [12], by considering popular places and transitions between place categories, and in [13], where displacement between consecutive check-ins and returning probability to venues were computed. More recently, a study showed that location information can be used to reveal personal details of users, with minimal amounts of check-in information [14].

The above sample of studies demonstrate the variety and amount of information that can be derived from user location information on GeoSNs, and consequently the need for examining threats to user privacy on these networks.

B. Privacy Feedback and Notifications

Feedback and notifications tools are commonly used for warning users about security and privacy risks on the web. Studies are emerging to assess user awareness of privacy implications and the impact of such tools on the user attitude while interacting with systems [15].

Visualisation of privacy warnings was found to be effective in increasing user awareness of privacy threats, as demonstrated in [16], while users were more able to access their information and to manage it effectively, if provided with a view of how their profile appeared to other people [17].

The extent of users' awareness and its impact on their attitude and privacy concerns was the subject of many studies. Rader observed the links between limited awareness of possible privacy violations and the usefulness of policy-based privacy solutions [2], while other studies noted that increased awareness encourages users to utilise stricter accessibility options [18]. Similarly, Sadeh et al. found that methods that raise users' awareness about the way their data is used tend to stimulate users to produce more accurate preferences and increase the users' trust in the application [19].Tsai et al. developed a mobile location-sharing application to investigate how informing users of who can access their location might influence their privacy concerns and attitude [20]. Their findings show that informed users were more comfortable with sharing their location and had less privacy concerns. Recently, Patil et al. observed that immediate feedback about location disclosures without any ability to control the disclosures, evoked feelings of oversharing and recommended the use of proactive techniques for adjusting recommendations to disclosure settings especially in the case of socially distant users and visiting atypical locations [6].

Usability of privacy notices and feedback tools is of relevance to this work. The complexity of privacy polices and settings and the need for more accessible tools motivated much research in this area [5]. Of interest are studies into users' perception of privacy risk, where visual cues were shows to be useful [21], particularly when shown in-context [22].

The above studies generally assume that awareness of location information is confined to the visits the user make to places, but do not consider a holistic view of possible inferences that may be made in the network, as described in section II-A above. This paper attempts to fill this gap, by utilising both lines of research in studying user awareness in GeoSNs.

III. Modelling User Content Awareness on Geo-Social Networks

In this section, we examine the question of whether a user is aware of the information they are sharing on geo-scoial networks. To answer this question, we begin by considering the dimensions of location data, its properties and relationships that can be used to build a user profile (geo-profile) on these networks, and use this information space to analyse user awareness. We then propose a model of privacy threat levels that is related to the content implicit in the shared data, and finally suggest that design of feedback tools for location content awareness needs to project both relevant content and associated threat levels.

A. Dimensions of a User Geo-Profile

On GeoSNs, users intentionally declare their presence in a particular place at a particular time. In some applications, for example, Google and Foursquare, users are able to grant permission for continuous background collection of their spatiotemporal tracks (by ”switching on location” on devices). In this section, we examine the dimensions of the data being collected in such systems and the types of information that can be inferred to construct geo-profiles for users.

Three primary dimensions to user information on geo-social networks can be identified; 1. the spatial dimension, 2. the social dimension, and 3. the temporal dimension.

The *spatial dimension* refers to the geographic locations associated with the user. A spatiotemporal (ST) track is composed of a sequence of time-stamped geographic coordinates representing the user’s movement in geographic space over time. The coordinates may refer directly to specific identifiable places, when users explicitly define the place they visit, or a process of reverse geocoding can be used to infer the possible place identity from the point coordinates. Increasingly, geographic gazetteers are shared between applications to aid this process, for instance, Instagram allows users to geotag their pictures using the Foursquare API and Facebook Places[3] and Twitter uses the Google API for linking users’ selected place name with a location on a map.

Given the place identifiers on a ST track, other useful place properties can be extracted, for example, the type of the place, e.g. a school or a hospital, and the types of services (or human activities) a place provides, e.g. education or health-related services, etc. [23]

Based on the spatial dimension, a user geo-profile would be capable of supporting the following queries.

- Which particular places are the user associated with? Outline the neighbourhoods of the user activity?
- What types of places does the user visit?

The *social dimension* is compound and comprises two distinct dimensions: a) social links to other users, and b) shared content. Explicit links to other users, for example, as friends or followers, is an orthogonal dimension to both the spatial and temporal dimensions, where social ties are formed and maintained between users independently of their presence in geographic locations. Shared content on social networks refers to different types of data provided by the users, for example, text (tags, tips, reviews, tweets, etc.), images or videos. This dimension is dependent on the spatial and temporal dimensions, thus particular tags or images are shared in particular places at specific time points.

With the spatio-social dimensions, a user geo-profile would be capable of supporting the following queries.

- Which concepts are the user interested in?
- Where would the user be associated with some specific concepts?
- Who does the user share particular interests with?
- Where would the user share an interest in a specific concept with another user or group of users?

The *temporal dimension* is essentially the time line recording the time stamps of the user’s visits to locations. Frequency of visits to geographic places can be used as an indicator of the degree of association with the place, or with the related activities and concepts. A mapping of the time line can be made to cluster specific temporal intervals and study emerging patterns of user activity, e.g. daily patterns (mornings, afternoon, evenings and night), weekends and weekdays, seasons, etc.

With the spatial-social-temporal composite space, a user geo-profile would be capable of supporting the following queries.

- When did the user visit a place? How often? How much time did he spend there?
- Where would the user be on (weekday mornings)?
- Which concept/activity is of interest to the user at a particular time point?
- Which other users/friends is this user normally with on (weekends)?
- Where does the user practice a certain activity with (friends) on (weekday evenings)?

In addition to patterns of presence in a place, a user geo-profile can also be used to detect patterns of absence from places.

- When is the user normally absent from a particular place during the week?

B. Modelling Levels of Threat to Location Privacy

One aspect of user content awareness is related to “Social privacy”, which concerns how an individual manages self-disclosures, availability, and access to information about themselves by other people when using social-driven applications [2]. To manage social privacy, one needs to understand the level of threat implied by his information disclosure and be able to relate it to the scope of visibility granted for this information. Here, a possible model is proposed of the levels of privacy threats with respect to the user geo-profile.

To model the threat level to location privacy on GeoSNs, we propose mapping to appropriate threat levels based potential privacy risks resulted from linking between data visibility and disclosed dimensions. The model assumes that there are three levels of visibility: private (no access to other people), friends (access only to friends) and public (access to others whether inside or outside the social network). It also takes into account three abstract levels of threat to convey the risk associated with disclosing personal information: green (safe to disclose the information; the disclosed information is discrete and in line with visibility granted), amber (caution; disclosing the information can result in particular associations extracted and revealed beyond visibility granted), and red (danger; disclosing the information can result in implicit patterns extracted and revealed beyond visibility granted). Considering the potential value of the information disclosed and the scope of visibility granted by the user, we introduce a possible mapping of threat

[3] instagram.com/developer/endpoints/locations/

TABLE I. A possible mapping of privacy threat levels against the dimensions of data in a geo-profile.

Dimension	Visibility		
	Private	Friends	Public
Spatial	green	green	amber
Social	green	green	amber
Spatial-Social	green	amber	amber
Spatial-Social-Temporal	green	red	red

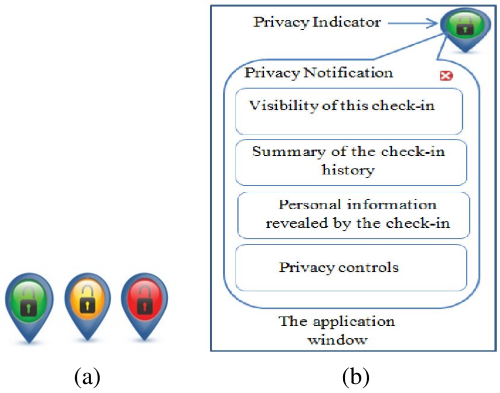

Privacy Indicator
Privacy Notification
Visibility of this check-in
Summary of the check-in history
Personal information revealed by the check-in
Privacy controls
The application window
(a)
(b)

Fig. 1. A potential design of the privacy-enhancing feedback and control tool showing the (a) icon design, (b) content of the privacy notification tool.

levels against the dimensions and visibility of data in a geo-profile as demonstrated in Table I.

Depending on the visibility scope, the threat level is increased as multiple dimensions are considered simultaneously (allowing links and patterns between data elements to be inferred). For instance, as the user gives explicit consent of visibility to the *Friends* group, access to data on the individual dimensions; spatial and social axes, are assumed to be granted (green), where as the threat level increases with the likelihood of disclosure of implicit data along composite dimensions.

Note that using the dimensions in Table I gives only an abstract model of the level of threats associated with those dimensions. Finer specification of the data elements, and relationships between data elements, shared or inferred, along those dimensions is possible and would give more insight to the threat levels inherent within the information space of a geo-profile.

C. Feedback Design for Location Awareness

To enable user content awareness in geo-social networks, privacy-enhancing feedback and control tools need to be designed and incorporated within the services. The development of such tools need to consider two requirements, a) Which content needs to be communicated to the user?, and b) how (and when) should the content be communicated to the user to satisfy (and enhance) their privacy awareness?

The first question involves considering the communication of three aspects related to a geo-profile. These are as follows.

1) Data content, both captured or constructed. Ultimately, a view of the whole geo-profile data space is possible, including historical data stored and inferred.
2) Visibility (or accessibility) of the geo-profile content to other users. The user needs to be able to know which other users in the network are able to gain access to their data, which types and how much volume of the data are visible.
3) Estimated threat level associated with the geo-profile. An indication of the link between content and visibility can be summarised as a degree of threat to user privacy. Some default estimation mechanism can be used to determine the levels of threat, such as the one described above, but this can be customized by the user, who may be able to indicate more accurately their perception of the value of their own data sets.

The second question is related to the usability of the design used for the feedback and control tools. Several research works have considered this issue and proposed design principles and frameworks for building privacy-friendly systems [24], [25], [26], [27], and highlighted important pitfalls [28] that privacy designers should avoid. In particular, content provided in location-awareness feedback tools need to be: a) relevant: so the user is able to recognize the relevance of the presented information to their location-sharing activity, b) transparent: with respect to system reasoning with the user data, c) timely: feedback should be presented at the point it is needed, d) actionable: so that the user can respond to the feedback with appropriate actions and e) comprehensible: so the user can make accurate interpretation of all elements of feedback provided.

As an application of the above design ideas, consider the possible design of a feedback tool for location awareness. Immediate feedback on location exposure consequences is assumed, where the system is able to use captured location information (where the user is at the present time) to project a report on privacy implications based on registering this location in the user geo-profile.

One possible design is shown in Figure 1(a). The figure shows an icon design for the feedback tool in the form of a location pin with a lock as an indicator for privacy threat. The colour of the icon is used to reflect the level of threat estimated by the system. The icon allows the user to explore their content to understand the basis for the threat indicated. This can be in the form of a concise pop-up window, as shown in Figure 1(b) that includes: a) a summary of the current location status, b) visibility permissions granted, c) a view of the geo-profile that lists possible constructed information based on this location. Note that this example is only given to illustrate a form of realisation of a feedback tool and will be used as a basis to measure some aspects of location awareness in the experiment described below. A more dedicated study of design issues is needed, but is out of the scope of this work.

It is assumed that the application stores a basic user location profile that records their spatiotemporal track of place visit and related contextual data, and that the application is

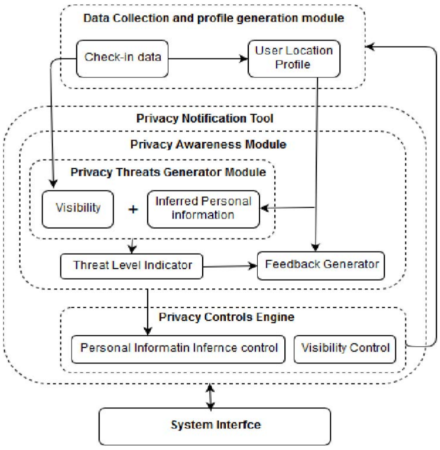

Data Collection and profile generation module
Check-in data
User Location Profile
Privacy Notification Tool
Privacy Awareness Module
Privacy Threats Generator Module
Visibility
+
Inferred Personal information
Threat Level Indicator
Feedback Generator
Privacy Controls Engine
Personal Informatin Infernce control
Visibility Control
System Interfce

Fig. 2. Components of the Location Privacy Awareness Tool.

also capable of deriving implicit information from this data, including for example, the strength of the user relationship to specific places, and his visit patterns to different places. Figure 2 is an overview of the envisaged privacy awareness system and its components. The system essentially comprises a profile analysis unit that dynamically analyses the user track data collected by the application to determine an appropriate privacy threat level to be displayed on the interface. A privacy control unit then allows the user to adjust the visibility and content of their stored data.

IV. Experiment

This experiment is designed to evaluate the relationship between the location awareness model presented and users' privacy perception and behaviour on GeoSNs. In particular, the experiment will aim first to explore the impact on a user's perception of privacy due to providing privacy feedback including a) The presenation of the geo-profile content, and b) the use of a privacy threat level indicator. Secondly, the experiment will study the impact on a user's behaviour when sharing his location data online due to his perception of privacy, resulting from the introduction of the privacy feedback with and without offering privacy controls.

A. Method

The Location-Based Social Network (LBSN) Foursquare was chosen as a platform for this study. It is a fairly popular LBSN that provides a typical example of GeoSNs, and as such has been used in several previous studies in the literature [7], [12]. Using a public GeoSN for evaluation provides more accurate insights of the general user's privacy attitude and bahaviour rather than using restricted location-sharing applications (e.g.[6], [20], [19]). Foursquare offers place discovery and recommendation services based on users' location and previous visits to places (check-ins). User's friends have access to his place profile, and the user is also able to grant access to other users who visit the same places in his profile.

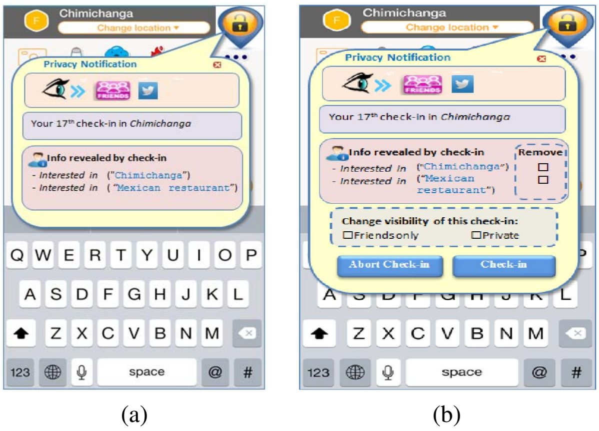
TABLE II. Summary of the check-in scenarios used.

Privacy level	Visibility	Inferred information
Amber (Moderate)	Friends and Twitter	Interests
	Friends only	Interests; Friendship relation
Green (Safe)	Friends only	None
Red (Risky)	Friends, Facebook and Twitter	Pattern of visit; Place type; Friendship relation
	Friends and Twitter	Private place (Home); Friendship relation
	Friends only	Pattern of activity; Absence from (Home); Predicted next check-in

The experiment took the form of an online user study that utilises realistic scenarios of using the Foursquare checking-in application. The scenarios were designed for checking-into places to cover different patterns of data exposure along the spatial, social and temporal axes. Feedback is provided "just-in-time" when needed during task execution. We used scenarios since we need to capture users' privacy attitude and bahaviour when presenting with potenial privacy risks rather than capturing whether they would check-into the particular places introduced in the scenarion in the real-life situation. Hypothetical requests and scenarios are exploited for gaining generalisable outcomes in considerable location-sharing studies (e.g. [19], [6], [29]). On the spatial axis, patterns of presence as well as absence from places were used and on the social axis, patterns of co-location with friends as well as of interest in certain concepts and activities, that may be inferred as a consequence of visiting the place or sharing a tip in the place, were used.

Six scenarios were designed based on our proposed threat level model, as shown in Table II, where in green-level scenarios, the check-in can be shared only with friends and just the current spatiotemporal information is revealed, the check-in can be public in amber-level scenarios and the users' association with the places,concepts and social ties can be extracted, and finally in red-level scenarios, check-ins can also be public and the disclosed spatial-social data constructs temporal patterns. The scenarios are presented in two conditions. First, the scenarios are presented with feedback only and then presented again with actionable controls over the information disclosed. We opted for within-subjects design since we were interested in capturing the impact of privacy awareness with and without controls besides its advantage of reducing errorvariance associated with individual differences (e.g. [30], [31]).

Perception of privacy is dependent on the user's ability to comprehend the information being disclosed. This study is not intended to measure comprehension, and thus it was important to reduce the effect of this variable on the result of the experiment to ensure its validity. To address this issue, the scenarios included an initial section that enforced (and simultaneously checked) the participants' comprehension, by repeating the displayed information as a list of statements and asking the participants to check their correctness in the



Your 17th check-in in *Chimichanga*

Info revealed by check-in
- *Interested in* ("Chimichanga") □
- *Interested in* ("Mexican restaurant") □

Change visibility of this check-in:
□Friends only □Private

(a)

(b)

Fig. 3. A sample of the location privacy awareness notice shown in the (a) feedback-only scenarios, (b) feedback and control scenarios.

scenario presented. For example, a participant would need to indicate whether visibility is set to friends only or whether the place type is displayed, before proceeding with the questionnaire.

B. Procedure

The user study is an online survey with four main sections. The first section collects the participants' demographics, captures their experience using social networks, and observes their location privacy concerns, awareness and behaviour when using them. This is to allow a comparison to be made of those variables after the experiment.

The second section is the feedback-only section. Six check-in scenarios are presented to the user. For example, *You are about to check into "Chimichanga"- a mexican restaurant. You choose to share this check-in on Twitter account*. In every scenario, two screen-shots are displayed to the participants. The first is the normal Swarm[4] check-in screen that presents the check-in task details and a location privacy icon displayed on the top left corner. The second screen shows the location privacy awareness pop-up window that would appear if the user were to click on the location privacy icon, as presented in Figure 3(a). The set of questions in this section are designed to capture the impact of the information on their privacy awareness and concern, their agreement with the choice of level threat associated with information and finally whether they would modify their check-in in any way if given the option to do so.

In the third section, the same check-in scenarios are used, but with the location privacy window now providing control options as well. Participants were offered the opportunity to delete any of the information elements presented from their geo-profile (by modifying the user data in a way that makes deriving the presented information element is impossible), to change the visibility of the check-in or opt to abort this check-in all together, as shown in Figure 3(b).

The final section examines the participants' perception with regards to their personal privacy, their need to be aware of the contents of their geo-profile and their need to control access to their data, as a consequence of participating in the study. Moreover, questions were used to gauge their reaction towards the location awareness tool proposed and its usability.

Pilot tests were conducted on three research students in the school and three Amazon Mechanical Turk workers who met the participation criteria (discussed in next section) in order to ensure the clarity and coherence of the user study. The tests provided valuable feedback on the structure and wording of the survey.

C. Recruitment and Participants

The experiment was conducted in June 2015. Participants were recruited using Amazon Mechanical Turk (MTurk) and were confined to those who use the Foursquare/Swarm application and check-in frequently (not less than three times a week on average). This was necessary to enable the participants to realistically relate themselves to the scenarios presented and to use their experience with the application when commenting on privacy implications. The MTurk workers also need to have 95% or more approval rate for at least 500 tasks to be able to participate to make sure that they provide valid feedback according to the study instructions.

Of the 363 who entered the study, 25 workers were excluded with the qualification test. We also ensured that a MTurk worker can only participant once in our study by monitoring the worker ID. 338 participants undertook the study, completed it in 23 minutes on average and were compensated $1.5 each. The demographics questions revealed that most of the participants were young people (mean= 30.29, SD= 6.45), with slightly more male participants (57%) than female ones (43%). Furthermore, the majority were from North America (59.2%) and Asia (34.6%).

V. Findings

Analysis of the survey data and presentation of the results were achieved using R statistical programming language, and SPSS was used for applying Friedman, McNemar-Bowker, Cochran's Q and Spearman's rank correlation tests. An overview of participants' social networking experience and pre-study privacy concerns is presented first, followed by analysis of the results from the check-in scenarios section and finally the post-study reflection on privacy perception and evaluation of the location awareness tool.

A. Pre-Study Phase

A pre-study evaluation of the participants' privacy concerns on the application was conducted to understand the relationship between their level of experience with the application, their location sharing behaviour and their privacy concerns.

[4] Swarm is the checking-in application for Foursquare.

Most of the participants were moderate users (check-in several times per week) (57.6%), while the rest were frequent users (check-in once or more per day) (42.4%). In addition, most participants would enable location services on their mobile devices (52% enable them frequently (always on) and 43% enable them moderately (when required by an application)).

Accessibility to the user personal data by other users is a primary privacy concern. This is commonly controlled by defining visibility of one's profile in the privacy settings within the network. 'Friends' on Foursquare are granted access to the full location history and thus can potentially have access to a complete geo-profile. However, it is interesting to note that people will accept friendship requests from strangers and in fact may not be fully aware of their friendship links. This idea was examined in the questionnaire where participants were asked if they actually know all of their friends (or would accept friendships with users whom they do not know), and revealed that only 31.7% of users know all their friends (44.4% know most of them, 23.4% know some of them). While 64% of participants think check-ins can be dangerous, most stated that they currently feel safe using Swarm (87%). Moreover, 71% of participants thought that the privacy settings provided were sufficient to protect their privacy, but many (46.15%) also admitted to not checking their privacy settings for a long time.

It is interesting to note the seemingly contradicting findings, where no evident link can be observed between the extent of visibility of location data and the sense of risk associated with disclosure of personal location with privacy concerns (feeling safe). One possible explanation is that user's awareness is related directly to the needs of the task being executed. Thus, awareness is limited to the location data a user is sharing at any point in time while using the application and hence his privacy concerns are also limited to only this part of his data set. This observation is supported by examining responses to a question on which aspects of their location history were they able to recall, where about 47% of were able to recall only one aspect and 2.7% remember nothing of thier history.

B. Check-in Scenarios Phase

Here the results of the questions from sections II (feedback only) and III (feedback and control) of the study are presented.

1) Impact of Content on Privacy Perception:

a) Sufficiency of The Content Provided: Following every scenario, two questions were used to gather users' perception of the sufficiency of the information content provided to convey privacy risk and the effect of the information on their privacy concerns. Most of the participants reported that the tool sufficiently indicated the privacy risks associated with the check-in scenarios, as shown in Figure 4. The agreement was highest in the red level scenarios, followed by Amber and green (representing 77%, 68%, and 63% respectively).

The content presented have a clear impact on the participants' privacy concern based on the threat level of the check-in scenario (Friedman Chi-Square = 91.227, $p = .000$), where participants were mostly concerned about their privacy in the red level scenarios as expected, followed by Amber and green (representing 72%, 55%, and 45% respectively). There is also a positive correlation between the participants' concern level with the threat level of the check-in scenario (Spearman rank correlation = .245, p = .000). Hence, the more threat the location disclosure poses, the more concerned the participants are on their privacy.

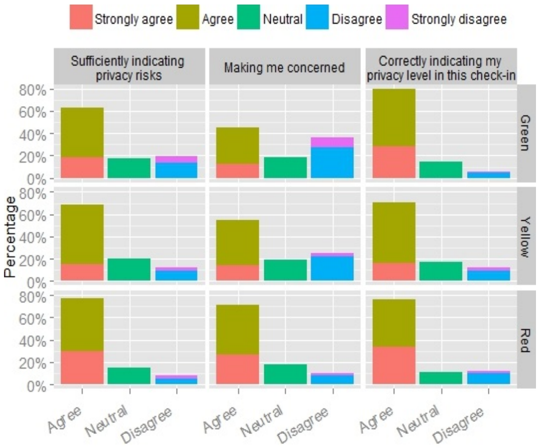


Fig. 4. Measure of effectiveness, grouped by threat level.

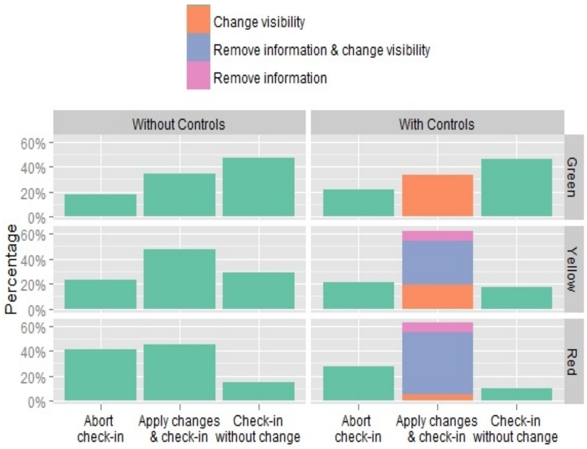


Fig. 5. Check-in decision with and without the privacy controls grouped by the threat level indicator.

b) Perception of Threat Level Estimation : A high level of agreement ($> 75\%$ overall) is reported by participants with the threat level indicator presented in every scenario (green: 80%, Amber: 76%, and red: 71%), whereas on average only 10% think the tool should indicate a different threat level. Of the 10% who disagreed with the threat level indicated, some thought that the threat is understated (it should be higher), as explained in their comments ("This seems like a fairly high degree of access to information" and "The application is profiling me and allowing any random person to know these things about me. That's extremely scary"), while others

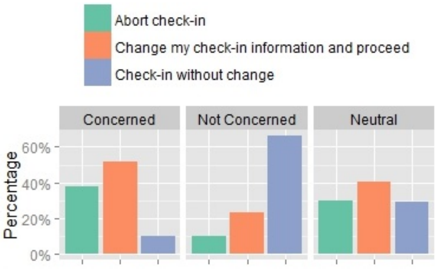


Fig. 6. Check-in attitude grouped by level of privacy concern.

felt that the privacy setting provided by the application were enough to neutralise the threat (“Only my friends will see my details” and “I am protected by my privacy settings”).

2) Impact of Content on User Behaviour: Figure 5 demonstrates the effect of content awareness on the attitude of users to modify their behaviour. On average, over 50% of users chose to modify their check-in action in some way, whereas the rest either chose to abort the check-in completely (28%) or would proceed without making changes (22%).

Scenarios with actionable control options significantly impact check-in behaviour (McNemar-Bowker=91.495, p = .000), where tendency to modify the check-in increased by 14% in the control scenarios (feedback only: 44%, feedback and control: 58%). In addition, with the control options, users were less likely to abort the check-in (by 7%), presumably as they were given more options to modify their information content. Users were rather conservative when choosing the control options, with 63% choosing to both remove the inferred information from their profile and change the visibility of their check-in, and the remaining group chose to either change the visibility (25%) or to remove the inferred information (12%).

a) Impact of the Threat Level Indicator on Behaviour : The threat level presented has a significant impact on the participants’ check-in behaviour (Cochran’s Q=33.566, p = .000). In particular, participants were equally willing to apply changes to their check-ins in the red (54%) and Amber (55%) threat levels, and less so with the green level (34%). Similarly, aborting a check-in was mostly evident with the red level scenarios 34%, followed by Amber and green (22%, and 20% respectively). As would be expected, ‘proceed with no changes’ option was more evident with the green level scenarios, followed by Amber and red (representing 47%, 23%, and 12% respectively).

b) Privacy Concern and Behaviour: It is useful to observe the impact of the level of privacy concern on the actions participants chose to perform (Cochran’s Q=254.628, p = .000), as presented in Figure 6. Participants who reported concern about their privacy were the most willing to modify their check-in information or to abort the check-in (52% and 38% respectively), followed by the group who were neutral about the privacy concerns (41% and 30% respectively). Note that the group who reported no privacy concern were still willing to modify their check-ins and abort the check-in scenarios (23% and 10% respectively). A positive correlation was noted between the participants’ level of concern and their check-in attitude, where higher levels of concern resulted in an increased tendency towards modifying the check-in information or aborting the check-in (Spearman rank correlation= .405, p = .000).

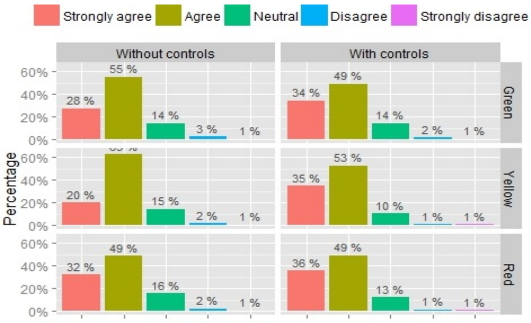


Fig. 7. The tool support for decision-making based on the availability of privacy controls, grouped by threat level.

c) Support in Decision-Making: Here we question how the participants’ decided to modify their check-in actions as a response to the feedback and control conditions. Control scenarios were found to more significantly influence the decision to take action (McNemar-Bowker Test=19.466, p = .000), where 41% (compared to 33%) of participants strongly agree that control scenarios were helpful in decision-making compared to the feedback condition. The difference was more pronounced in the red threat level scenarios as shown in Figure 7.

C. Post-Study Phase

1) Location Awareness and Privacy Concern: The overall effect of location awareness on privacy concern was measured in post-scenarios questions (Cronbach’s α = .78) and results are shown in 8. The figure confirms the assumptions made at the start of this study, where a significant portion of participants (66%) were not aware of the possible information content in their geo-profiles and (71%) underestimated the privacy risk associated with their check-in activity. Similarly, (76%) reported that they are now more concerned about their location privacy (47% of those were strongly concerned), and 8% were not concerned.

Comparing privacy concern before and after the study (check-in scenarios with the privacy feedback) , it was clear that the tool has a significant impact on the level of privacy concern of participants (McNemar-Bowker Test=284.520, p = .000), where a strong negative correlation between the concern level before the scenarios was noted (Spearman rank correlation= -.829, p = .000). As a consequence, most participants (84%) also suggested that the experiment will impact the way they use Swarm in the future (“will be more cautious”).

2) Usability of The Location Awareness Tool: Finally, the overall impression of whether participants consider the tool

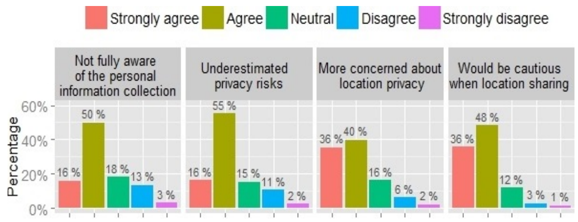


Fig. 8. General privacy perception after the check-in scenarios.

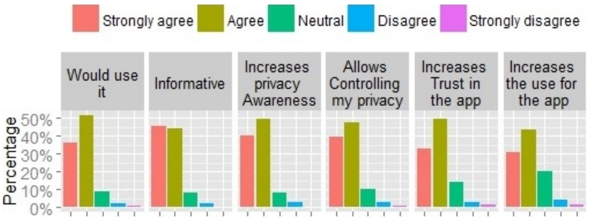


Fig. 9. Overall impression of the utility of the location awareness tool.

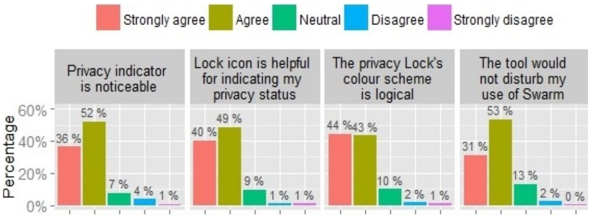


Fig. 10. Overall evaluation of tool design.

useful was measured (Cronbach's α = .892). Figure 9 gives a summary of the results. Note in particular, how the results imply that such a tool (providing location content awareness and control) can lead to an increased trust in the application and consequently more frequent use of the application.

Participants were also supportive of the design of the threat level indicator, in terms of the icon choice and the colour scheme used and suggested that the design will not interrupt their use of Swarm, as shown in figure 10. Comments to an open-ended question on their views on the tool are in line with the results ("It looks good and does not disrupt the current format", "I think the tool is an innovative idea, especially for those who are uncertain about the personal data that is being shared." and "I really love this tool, and it would make me feel much safer when using Swarm."). However, this conclusion needs to be verified by realistic experience sampling methods in the future.

VI. Conclusions

User's awareness of the consequences of sharing their location online is rather limited. The reason is twofold; firstly, due to the limitations in our abilities as humans to attend to and recall information that are not needed directly to the task at hand, thus we will not seek to recall details of our spatiotemporal profiles when checking-in a place, and secondly, due to the limited support offered by the social networks to enable users' perception of their information content.

This paper addresses this problem by, a) analysing the scope of privacy threats on geo-social networks, along the spatial, temporal and social dimensions of data in geo-profiles, b) proposing the design of feedback tools that project a view of the level of threat associated with the disclosure of location information, and c) testing the implication of presenting the feedback on users' perception of privacy concerns and their attitude towards sharing their location data on social networks. Findings from the user study conducted are summarised in the form of recommendations for the design of more effective privacy-sensitive geo-social networks. Future work will look further into the design aspects of the proposed feedback and control tools, in particular, the scope of information to be revealed and its timing with respect to task performance, and will seek in-depth evaluation of the usability aspects of the design.

References

[1] F. Alrayes and A. Abdelmoty, "Privacy concerns due to location sharing on geo-social networks," *International Journal On Advances in Securityl*, vol. 7, no. 3 and 4, pp. 62–75, 2014.

[2] E. Rader, "Awareness of behavioral tracking and information privacy concern in facebook and google," in *Proc. of Symposium on Usable Privacy and Security (SOUPS), Menlo Park, CA, USA*, 2014.

[3] D. Fisher, L. Dorner, and D. Wagner, "Short paper: location privacy: user behavior in the field," in *Proceedings of the second ACM workshop on Security and privacy in smartphones and mobile devices*. ACM, 2012, pp. 51–56.

[4] I. Liccardi, J. Pato, D. J. Weitzner, H. Abelson, and D. De Roure, "No technical understanding required: Helping users make informed choices about access to their personal data," in *Proceedings of the 11th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2014, pp. 140–150.

[5] N. Wang, J. Grossklags, and H. Xu, "An online experiment of privacy authorization dialogues for social applications," in *Proceedings of the 2013 conference on Computer supported cooperative work*. ACM, 2013, pp. 261–272.

[6] S. Patil, R. Schlegel, A. Kapadia, and A. J. Lee, "Reflection or action?: how feedback and control affect location sharing decisions," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2014, pp. 101–110.

[7] T. Pontes, M. Vasconcelos, J. Almeida, P. Kumaraguru, and V. Almeida, "We know where you live?: privacy characterization of foursquare behavior," in *UbiComp '12 Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, 2012, pp. 898–905.

[8] A. Sadilek, H. Kautz, and J. Bigham, "Finding your friends and following them to where you are," in *Proceedings of the fifth ACM international conference on Web Search and Data Mining, WSDM '12*, 2012, pp. 723–732.

[9] H. Gao, J. Tang, and H. Liu, "gSCorr: modeling geo-social correlations for new check-ins on location-based social networks," in *Proceedings of the 21st ACM international Conference on Information and Knowledge Management, CIKM '12*, 2012, pp. 1582–1586.

[10] D. Crandall, L. Backstrom, D. Cosley, S. Suri, D. Huttenlocher, and J. Kleinberg, "Inferring social ties from geographic coincidences," in *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 52, 2010, pp. 22 436–22 441.

[11] D. Dearman and K. Truong, "Identifying the activities supported by locations with community-authored content," in *Proceedings of the 12th ACM international conference on Ubiquitous computing*, 2010, pp. 23–32.

[12] A. Noulas, S. Scellato, C. Mascolo, and M. Pontil, “An empirical study of geographic user activity patterns in foursquare,” in *ICWSM*, 2011, pp. 70–73.

[13] D. Preotiuc-Pietro and T. Cohn, “Mining user behaviours: a study of check-in patterns in location based social networks,” *Web Science*, 2013.

[14] Y. Zhong, N. J. Yuan, W. Zhong, F. Zhang, and X. Xie, “You are where you go: Inferring demographic attributes from location check-ins,” in *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*. ACM, 2015, pp. 295–304.

[15] D. Malandrino, V. Scarano, and R. Spinelli, “Impact of privacy awareness on attitudes and behaviors online,” *SCIENCE*, vol. 2, no. 2, pp. pp–65, 2013.

[16] D. Christin, M. Michalak, and M. Hollick, “Raising user awareness about privacy threats in participatory sensing applications through graphical warnings,” in *Proceedings of International Conference on Advances in Mobile Computing & Multimedia*. ACM, 2013, p. 445.

[17] M. Anwar and P. W. Fong, “A visualization tool for evaluating access control policies in facebook-style social network systems,” in *Proceedings of the 27th Annual ACM Symposium on Applied Computing*. ACM, 2012, pp. 1443–1450.

[18] M. Fire, D. Kagan, A. Elishar, and Y. Elovici, “Social privacy protector-protecting users privacy in social networks,” in *SOTICS 2012: Second International Conference on Social Eco–Informatics*, 2012, pp. 46–50.

[19] N. Sadeh, J. Hong, L. Cranor, I. Fette, P. Kelley, M. Prabaker, and J. Rao, “Understanding and capturing people’s privacy policies in a mobile social networking application,” *Personal and Ubiquitous Computing*, vol. 13, no. 6, pp. 401–412, 2009.

[20] J. Y. Tsai, P. Kelley, P. Drielsma, L. F. Cranor, J. Hong, and N. Sadeh, “Who’s viewed you?: the impact of feedback in a mobile location-sharing application,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2009, pp. 2003–2012.

[21] B. Zhang, M. Wu, H. Kang, E. Go, and S. S. Sundar, “Effects of security warnings and instant gratification cues on attitudes toward mobile websites,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2014, pp. 111–114.

[22] M.-E. Maurer, A. De Luca, and S. Kempe, “Using data type based security alert dialogs to raise online security awareness,” in *Proceedings of the Seventh Symposium on Usable Privacy and Security*. ACM, 2011, p. 2.

[23] A. Alazzawi, A. Abdelmoty, and C. Jones, “What can i do there? towards the automatic discovery of place-related services and activities,” *International Journal of Geographical Information Science*, vol. 26, no. 2, pp. 345–364, 2012.

[24] V. Bellotti and A. Sellen, “Design for privacy in ubiquitous computing environments,” in *Proceedings of the Third European Conference on Computer-Supported Cooperative Work 13–17 September 1993, Milan, Italy ECSCW93*. Springer, 1993, pp. 77–92.

[25] M. Langheinrich, “Privacy by designprinciples of privacy-aware ubiquitous systems,” in *Ubicomp 2001: Ubiquitous Computing*. Springer, 2001, pp. 273–291.

[26] B. Friedman, P. Lin, and J. K. Miller, “Informed consent by design,” *Security and Usability*, pp. 495–521, 2005.

[27] A. Adams and M. A. Sasse, “Privacy issues in ubiquitous multimedia environments: Wake sleeping dogs, or let them lie,” in *Proceedings of INTERACT*, vol. 99, 1999, pp. 214–221.

[28] S. Lederer, J. I. Hong, A. K. Dey, and J. A. Landay, “Personal privacy through understanding and action: five pitfalls for designers,” *Personal and Ubiquitous Computing*, vol. 8, no. 6, pp. 440–454, 2004.

[29] K. P. Tang, J. I. Hong, and D. P. Siewiorek, “Understanding how visual representations of location feeds affect end-user privacy concerns,” in *Proceedings of the 13th international conference on Ubiquitous computing*. ACM, 2011, pp. 207–216.

[30] R. Balebako, J. Jung, W. Lu, L. F. Cranor, and C. Nguyen, “Little brothers watching you: Raising awareness of data leaks on smartphones,” in *Proceedings of the Ninth Symposium on Usable Privacy and Security*. ACM, 2013, p. 12.

[31] H. Almuhimedi, F. Schaub, N. Sadeh, I. Adjerid, A. Acquisti, J. Gluck, L. F. Cranor, and Y. Agarwal, “Your location has been shared 5,398 times!: A field study on mobile app privacy nudging,” in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 2015, pp. 787–796.