Location Privacy Awareness on Geo-Social Networks

Enhancing Awareness with Feedback Solutions

Fatma Sulaiman AlRayes

2017

Cardiff University

School of Computer Science and Informatics

A thesis submitted in partial fulfilment of the

requirement for the degree of Doctor of Philosophy

Declaration

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed . . . . . (candidate)

Date . . . . . .

Statement 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of ..... (insert MCh, MD, MPhil, PhD etc, as appropriate)

Signed . . . . . (candidate)

Date . . . . . .

Statement 2

This thesis is the result of my own independent work/investigation, except where otherwise stated, and the thesis has not been edited by a third party beyond what is permitted by Cardiff University's Policy on the Use of Third Party Editors by Research Degree Students. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed . . . . . (candidate)

Date . . . . . .

Statement 3

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed . . . . . (candidate)

Date . . . . . .

To my role model, my mum Sua'ad
To my mentor, my dad Sulaiman
To my beloved husband, Ibrahim
To my little angel, Sarah
,
I hope I made you proud!

Abstract

Users of GeoSocial Networks (GeoSNs) share their personal location information with other users online. GeoSNs use spatiotemporal histories of users and other semantic information from their tags and comments to build location-based profiles and to offer personalised services and interaction experiences. However, such location-based profiles can potentially be used to extract private information about users, that they may not wish to disclose, and can thus pose a threat to their privacy. Users are generally unaware of the extent of data they are sharing and its potential implicit content. Studies have also shown that users are concerned about their location privacy and that current solutions offered by GeoSNs, namely privacy policies and privacy settings, do not effectively address their concerns.

The focus of this thesis is on addressing the problem of location privacy on GeoSNs through enhancing users' location privacy awareness of potential risks to allow them to make informed consent about their location disclosure. Therefore, this work firstly studies the link between location information disclosure and the risks to personal privacy and evaluates the level of user awareness and their attitude to privacy implications of sharing location information in GeoSNs. Factors contributing to the location privacy problem are identified, including those stemming from the nature of the data collection procedures and the modes of using the application by the users. Systematic user studies were carried out that showed the limitation in users' awareness of the extent of the data and information they are disclosing.

Thus, to enable location privacy awareness, a data-driven approach is undertaken to assessing the threat associated with the exposure of location-related personal information. Based on that, a privacy threat model is proposed that takes into account the types of shared data, its visibility by possible adversaries and the user's awareness of the disclosed information. In addition, privacy feedback solutions are proposed to address the gaps in user awareness by revealing the level of risk to their privacy associated with exposing different types of location-related personal information. These solutions allow users to view their geo-profiles collected and inferred based on their location-sharing actions on GeoSNs and notify them about who of the other users can see their information. User-based experiments were used to evaluate the effectiveness of the proposed solutions using surveys, interviews, and prototypes along with realistic users'

data. Results demonstrate clearly the significance of the proposed solutions on enhancing user awareness. Employing the methods proposed in this thesis will thus enable users to effectively manage their privacy and make informed decisions about their location disclosure on GeoSNs.

Acknowledgements

I would like to express my appreciation and gratitude for my main supervisor, Dr. Alia Abdelmoty for her support throughout my studies. I thank for your teaching me how to become a researcher, for your precious knowledge and advices you offered me, and for your encouragement to purse my PhD degree. I specially thank you for being understanding and supportive during my personal circumstances that I have been through while studying. I, indeed, have learned a lot from you.

Many thanks to my second supervisor, Dr. George Theodorakopoulos, for his motivation, ideas and kindness. I would like also to extend my thanks to all members of the School of Computer Science and Informatics for their support and helpful advices, especially to Dr. Rob Davies, and Mrs Helen Williams.

I am extremely thankful for Princess Nora bint Abdulrahman University for the scholarship, and for the Saudi Arabian Cultural Bureau for their continuous help and sponsor during my study years.

My deepest gratitude goes to my backbone, my family. Mum, I am thankful to be a daughter of a great lady like you. Words cannot describe how much I am grateful for your endless love, encouragement and support. You made me who I am. Special thank for my mentor, my dad, who have been caring and supporting throughout my studies. To my husband, I am appreciative for standing by me, motivating and believing in me. Many thanks to my siblings, Rayes, Omama, Sarah, Omar, Sahar and Shahad for encouraging me and being there for me. To my little angle, my baby Sarah, thank you for filling my heart with love and for being strong during the times I had to be away from you. I love you all and you mean the world to me.

Finally, I would like to especially thank my close friends Dr. Liqqa Nawaf and Dr. Hana Aldahawi for their care, help and encouragement. You made my PhD journey enjoyable and unforgettable experience. Many thanks also goes to Dr. Taimur, Dr. Soha, Dr. Nasser, Dr. Haya Dr. Shada, Aseelah, Ashwan, Lowri, Liam, Nyala and Wafa for their support and kindness.

Contents

List of Figures

List of Tables

List of Acronyms

AOI Areas of Interest

API Application United Interface

GeoSNs Geo-Social Networks

GPS Global Positioning System

LBSs Location-Based Services

LBSNs Location-based Social Networks

LESNs Location-enabled Social Networks

LPTLM Location Privacy Threat Levelling Model

LPPMs Location Privacy-Preserving Mechanisms

MTurk Amazon Mechanical Turk

Chapter 1

Introduction

1.1 Introduction and Motivation

An exponential evolution and user growth is evident on online social networks, where people can communicate through various means based on information-sharing. For example, the number of users of social media worldwide has reached 1.96 billion, and 81% of the US citizens had a social media profile [1]. The affordability of location-aware mobile devices has enabled these networks to seize location feature by supporting location-based services along with the other online social networking advantages. Online social networks which embrace geo-location technologies can be referred to as Geo-Social Networks (GeoSNs), such as Facebook, Twitter and Foursquare. GeoSNs can be defined as "a web-based or mobile-based service that allows users to (1) construct a profile containing some of their geolocated data (along with additional information), (2) connect with other users of the system to share their geolocated data, and (3) interact with the content provided by other users (for instance by commenting, replying or rating)" [2].GeoSNs basically allow for the sharing of location information with other users, while enjoying all other features of traditional online social networks [2]. Users of GeoSNs can, for example, record their visit to places (check-in), share their thoughts about a place (leave a comment), or geo-tag a picture of theirs. A social dimension is involved within these applications in which users are motivated to interact with others rather than interacting with the service provider. Users can geo-tag other users in their shared contents. Moreover, users can share a large amount of real-time information that can be accessed by wide audience, as opposed to traditional Location-Based Services (LBS) [2, 3]. They also share other contextual information along with their location data including social connections and reviews [4]. Interestingly, almost all of the modern online social networks are GeoSNs where they enable sharing location information or using services based on location. In general, the majority of users share their accurate location that they visited to take advantage of the services provided by the GeoSNs. For example, in the study presented in Chapter 7, participants confirmed that their spatiotemporal location history presented to them are correct and reflected their real mobility. However, some users tend to disclose a nearby or general location, or even share it after leaving the place

which can affect the spatial and temporal accuracy of the data if they felt the need to protect their privacy, as discussed in Section 3.2.1.6.

Sharing location information on GeoSNs can pose several threats to user privacy. Providing location-based services requires the knowledge of the users' location, and in some GeoSNs, this location information is very precise. Location is considered to be sensitive in comparison to other types of information. Location information typically includes the spatial (where) and temporal (when) factors that make it distinctive and dynamic in nature. Tracking a user's location over time can reveal their identity [5, 6]. In addition, users' historical location information can be linked to contextual and semantic information publicly available on GeoSNs and can be used for constructing comprehensive user profiles that include inferred personal information about users [7]. Derived information in such profiles can include user activities, habits, mobility patterns and relationships with others [8, 9, 10, 11, 12]. Such enriched location-based profiles can be considered to be useful if used to personalise and enhance the quality and usability of the applications. However, it can potentially be used for undesirable purposes and pose new threats to users' location privacy, which refers to a particular type of information privacy that supports users' right to be consent to all aspects of their location disclosure [13, 5]. Thus, location disclosure on GeoSNs can expand the implications to location privacy, compared to LBS.

Users' concerns about their location privacy are evident [14, 15, 16], yet the provided privacy solutions have shown to be ineffective. Privacy policies and privacy settings have been applied widely in online services as a means of providing privacy management to users by offering them some information about how their data are handled and basic access controls to their data. Nevertheless, privacy policies are relatively ineffective due to the vague and complex presentation of information [17]. Privacy settings also enable limited protection that has shown to be difficult to use [18, 19]. In fact, users are unable to effectively use them due to their lack of awareness of the potential privacy consequences of their data exposure [20, 21, 22, 23]. In addition, Location Privacy-Preserving Mechanisms (LPPMs) have been developed to provide protection in LBS which mainly utilise anonymity and obfuscation techniques [13, 24]. However, these are not fully robust in protecting location privacy against attacks [25]. They mainly work on reducing the spatial and temporal accuracy which restricts their applicability on GeoSNs and impacts the quality of service provided [7] (see Section 2.2.2). Hence, there is a need for effective location privacy solutions.

GeoSNs fail to provide privacy awareness related to users' information exposure and associated privacy implications. The design of these applications focuses on providing an enjoyable experience for users, and lacks support for awareness of data collection, accessibility, and utilisations [26, 4]. Therefore, users are not fully aware of their explicit or implicit location data collection and the related privacy threats, which can impact how they appreciate their location privacy

[22, 20, 27]. Users' ability to make informed decisions about their location information disclosure and manage their privacy according to their preferences are restricted as well. The extent of the privacy management actions carried out by users is related to their personal awareness and experience with how a system processes their data, which is shown to be limited [28, 29]. Thus, users need to be informed about the explicit and implicit profiles that can be constructed based on their location data shared on GeoSNs.

Privacy feedback and control methods has been shown to be useful for enhancing users' privacy awareness, and ultimately allowing them to effectively manage their privacy in the area of mobile and social networking applications [30, 31, 33]. They also take into consideration the varying levels of personal privacy required by individuals, where they can decide upon their sharing actions according to their personal privacy preferences [35]. Thus, supporting users' awareness of what of their data are collected, how they can be used to extract personal information, and who can access their data, has great potentials for addressing location privacy in GeoSNs. This awareness includes informing users about potential privacy consequences, hence enabling them to provide informed consent about their information disclosure which matches their desired privacy preferences. Providing real-time privacy feedback has been shown to significantly impact users' data sharing decisions. There, users are notified of possible privacy risks just before they disclose their information [32, 34]. This type of feedback can be utilised for enhancing location privacy awareness in GeoSNs by allowing users to make information location-disclosure decisions through offering just-in-time privacy notification based on users' profiles and where a privacy awareness is needed.

Few previous studies have addressed the impact of location privacy awareness on users' sharing behaviour [36, 37, 38, 34]. The studies mainly offer rule-based privacy controls for managing location disclosure by specifying dynamic rules including place, time and recipient in geo-social applications. They generally assume that awareness of location information is confined to the visits the user makes to places, and focus on user awareness of the visibility and accessibility of their location information. However, these studies do not consider a holistic view of the possible personal information inferences that may be made in the network, which is another important aspect of privacy awareness. The studies' test environment was limited, as proprietary applications with limited features of interaction were used in their evaluation, which does not provide sufficient representation the public (commercial) GeoSNs environment. For example, the participants interact with few users and mainly share their location without any contextual information such as reviews, tags, or pictures. Thus, the key concerns and challenges of this work are to address the problem of location privacy in GeoSNs through improving location privacy awareness of both the content aspect in terms of data collected and inferred about a user, and the accessibility aspect in terms of who of the application's users can view a user's data and by using public GeoSNs that are used by and available for millions of users.

The aim of this research is to enable location privacy awareness by supporting privacy-oriented GeoSNs. In order to achieve this goal, firstly, the link between location information disclosure and risks to privacy need to be extensively studied. In addition, explicit presentation of location privacy threats is considered. In particular, possible inference of hidden personal information based on location-sharing activities on GeoSNs is investigated in order to understand information dimensions and relationships in geo-profile. User awareness of their information disclosure and location privacy implications when using these networks is also assessed. This assessment led to identifying the gaps in user awareness with respect to user geo-profiles. Feedback design solutions are proposed based on the gaps identified that provide explicit presentation of location information exposure and associated locations privacy risks resulting from location-sharing activities on GeoSNs. These solutions are evaluated for their impact on users' attitude (awareness and concerns) and behaviour.

1.2 Research Hypothesis and Questions

This thesis addresses the following hypothesis: *The lack of personal location privacy on GeoSNs can be addressed by enhancing user awareness of the information they share and its implication on their personal privacy. A framework for the storage and presentation of personal location information as well as its privacy implications needs to be supported by GeoSNs. The effectiveness of the framework can be evaluated by measuring its impact on users' attitude and behaviour when interacting on these networks.*

By attitude we mean their privacy awareness and concerns, and by behaviour we mean their information sharing actions. The lack of personal location privacy implies that the provided support for privacy in GeoSNs is not adequate for protecting users' data against undesirable collection, accessibility or personal information inferences that can pose threats to their privacy. Users tend to be aware only of what data they are currently sharing and fail to recognise how information resulting from their sharing activities can be linked and utilised in a way that lead to privacy implications (see Section 2.4). Generally, there are two levels for users' data exposure that can compromise their privacy, namely, the application and its third parties, and users of the application. This work considers users' friend and the general public (i.e. other users of the application) as the stake-holders where location privacy implication can take place in relation to exposing users' data to these stake-holders.

To test this hypothesis, this research aims to address the following:

a) Research the problem of location privacy awareness on GeoSNs by examining privacy implications of location information disclosure and users' attitude towards them.

b) Assess the sources and extent of the lack of awareness of privacy implications by users when using GeoSNs.

c) Identify methods for extending or modifying the design of GeoSNs to enable user awareness of their information and privacy threats associated with sharing this information.

d) Evaluate the effectiveness and usability of the proposed solutions by measuring their impact on user attitude and behaviour when sharing their location information.

The above objectives can be expanded into the following research questions:

RQ1: What are the privacy implications of sharing location on GeoSNs? (Work of Chapter 3)

a) What are the factors that contribute to location privacy threats with respect to how data are collected and handled by GeoSNs?

b) What sorts of implicit information can be derived from the data collected?

This question is answered by examining how users' data are handled in two distinctive GeoSNs as examples from a privacy perspective, followed by analytically exploring the location data content in terms of the range of possible inference that can be made from using a representative data set.

RQ2: What is the degree of awareness of location privacy of users of GeoSNs? (Work of Chapter 4)

a) Are users aware of the potential privacy risks associated with personal location exposure?

b) How concerned are users about their location privacy?

This question is addressed using an online user survey that covers different aspects related to measuring users' privacy awareness.

RQ4: How can providing real-time privacy feedback improve users' location privacy awareness and influence their privacy attitude and behaviour? (Work of Chapter 5)

a) What are the main data dimensions of a user location-based profile derived from GeoSNs and how can it be mapped to possible privacy risks and gaps in users' privacy awareness?

b) Can privacy feedback methods address the gap in user awareness of their location privacy?

c) Does enhancing user awareness have an impact on users' attitude and behaviour when sharing information on GeoSNs?

This question is answered by proposing and evaluating a real-time privacy feedback design through a user-based experiment that test its impact on users' privacy attitude and location-sharing behaviour towards realistic scenarios.

RQ5: Can the levels of threat to personal privacy be modelled? (Work of Chapter 6)

a) What factors influence the user's appreciation of privacy risk on GeoSNs?

b) Can these factors be used to model the level of threat to personal location privacy?

An in-depth investigation was conducted using a large user-based experiment to explore users' behaviour and perception in diverse location-sharing scenarios that reflect factors related to location disclosure.

RQ6: How does providing appropriate access to a user's profile impact their privacy awareness and their privacy attitude and behaviour? (Work of Chapter 7)

a) What are the components of a system that support location privacy awareness?

b) How can the interface to personal location information be enhanced to improve user awareness of location privacy?

c) Can improved access to personal information affect user's attitude and behaviour on GeoSNs?

This question is addressed by developing and evaluating a privacy-aware geo-profile visualisation system by interviewing users of GeoSNs in a lab experiment to examine whether it influences their privacy attitude and behaviour, using their own profiles.

1.3 Research Contribution

The main contributions of this research work are outline as the followings:

1) **Identifying the main factors contributing to privacy risks as a consequence of personal location sharing on GeoSNs (see Section 3.5)** using realistic applications and application data sets .

2) **Evaluating the level of and gaps in user awareness and attitude to privacy implications of sharing location information in GeoSNs** using representative samples of users of these applications.

3) **Proposing and evaluating a novel design of privacy notification and feedback tool to address the gaps in user awareness in GeoSNs** by providing a real-time notification of users' disclosed information and related privacy risks for a specific location-sharing task.

4) **Proposing a novel data-driven model of threat levels associated with disclosure of personal location information on GeoSNs (see Section 6.4.2)** through understanding users' privacy perception and sharing behaviour in regards to three factors; the dimensions of the exposed data, data visibility to others, and users' awareness of potential privacy implications resulting from data disclosure.

5) **Designing and evaluating a novel privacy awareness-oriented interface for accessing a user location profile on GeoSNs** by providing a privacy-oriented interface for accessing users' geo-profiles derived from their accounts on these applications.

The novelty of the three later contributions is clearly marked and discussed in Section 2.7.

1.4 List of Publications

Work conducted in this thesis has contributed to the following publications:

Journal Papers

- F. Alrayes and A. Abdelmoty, "Towards Understanding Location Privacy Awareness on Geo-Social Networks," *ISPRS International Journal of Geo-Information*, vol. 6, no. 4 ,p. 109, 2017.
- F. Alrayes and A. Abdelmoty, "No Place to Hide: A Study of Privacy Concerns due to Location Sharing on Geo-Social Networks," *International Journal On Advances in Security*, vol. 7, no. 3 and 4, pp. 62-75, 2014.

Conference Papers

- F. Alrayes and A. Abdelmoty, "Towards Location Privacy Awareness on Geo-Social Networks," *in Next Generation Mobile Applications, Security and Technologies (NGMAST)*, IEEE, 2016.
- F. Alrayes and A. Abdelmoty, "Privacy concerns in location-based social networks," *in GEOProcessing 2014: The Sixth International Conference on Advanced Geographic Information Systems, Applications, and Services*, IARIA, pp. 105-114, 2014.

1.5 Thesis Outline

Figure 1.1 shows a map of the work carried out in this thesis. A summary of the rest of the thesis is as follows:

Chapter 2: gives an overview of location privacy and emerging challenges within the domain of social networks, and a review of the literature related to the topics addressed in this work.

Chapter 3: presents a study on the location privacy of users in the domain of GeoSNs by identifying the factors that contribute to the location privacy problem, and analysing possible derived information from typical data sets collected by these applications for different types of users.

Chapter 4: introduces a user-based study for investigating users' privacy attitude and behaviour towards potential location privacy implications on GeoSNs .

Chapter 6: proposes a Privacy Feedback Design for Location Awareness based on investigating level of awareness with respect to extended user profiles that aims for providing real-time and in-task privacy notification. It then evaluates its impact on privacy attitude and location-sharing behaviour through a user-based experiment.

Chapter 7: presents an in-depth investigation of users' privacy perception and sharing behaviour on GeoSNs though a large user-based study, and utilises its outcomes to model threat levelling to location privacy on such applications.

Chapter 8: proposes a Geo-profile Visualisation System for supporting location privacy awareness in GeoSNs by enabling users to access their geo-profile extracted from their accounts, and view the related risks to their privacy. It then evaluates its impact on privacy attitudes and location-sharing behaviour through a user-based experiment.

Chapter 9: concludes this thesis by summarising the main contributions in relation to the hypothesis and discussing possible future work.

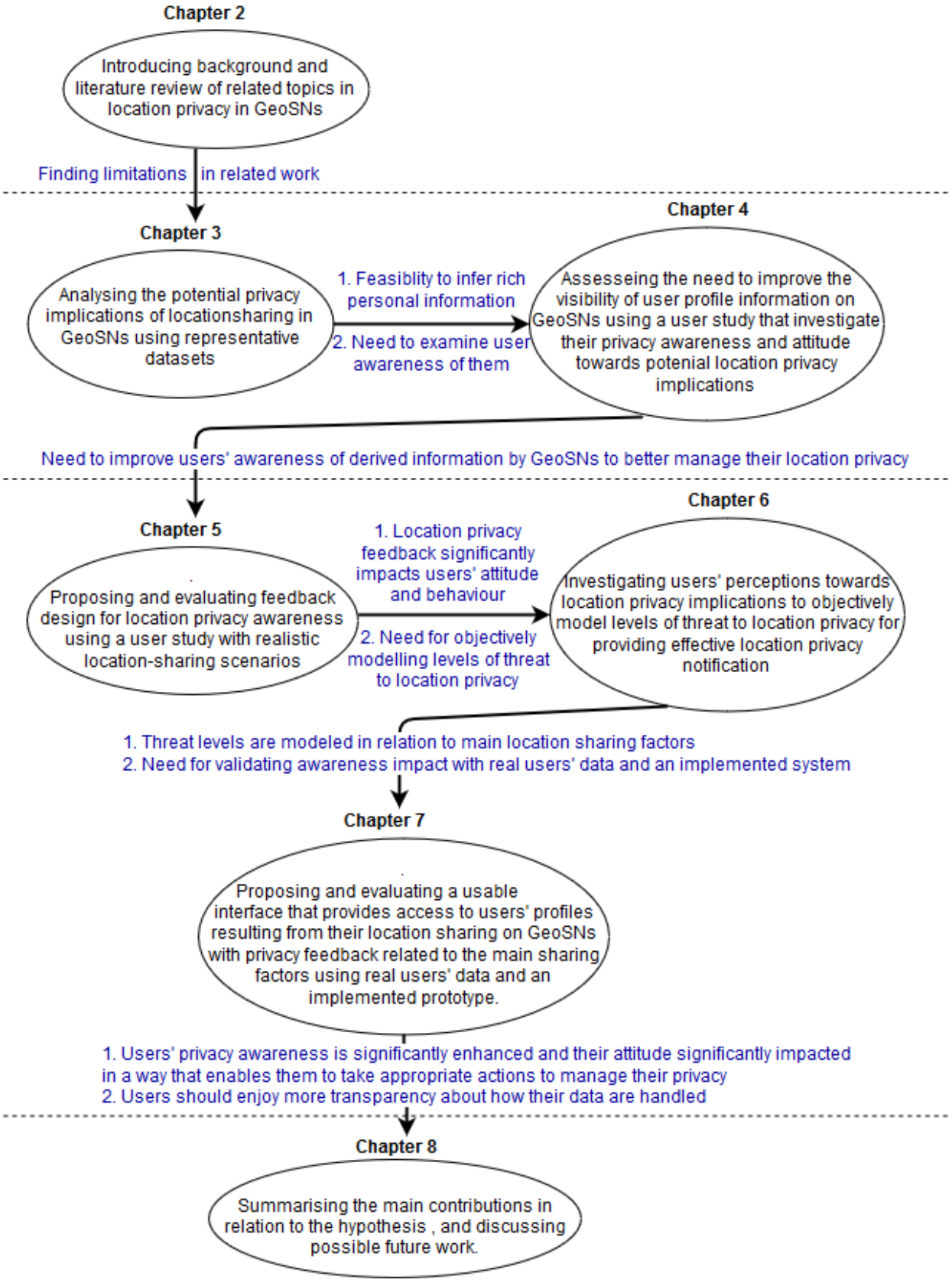


Figure 1.1: A map of the work carried out in this thesis..

Chapter 2

Background and Literature Review

2.1 Introduction

This chapter introduces a background and overview of work in five main areas of research that are relevant to the scope of work presented in this thesis. The topics covered in this chapter are as follows:

1) Challenges of location privacy in GeoSNs and the effectiveness of existing privacy management methods.

2) User profiling elements resulting from disclosing location data in GeoSNs in order to understand the potential personal information inferences and hence privacy risks.

3) Privacy attitude in the area of social and location applications to assess the role of privacy awareness in shaping users' privacy perceptions and behaviour.

4) Privacy modelling for investigating to what extent these models can be applied in the GeoSN domain.

5) Evaluation of the privacy-enhancing technologies in the context of their effectiveness and applicability for improving location privacy awareness in GeoSNs, starting with classification of the approaches and experimental methods used, and moving to reviewing them in more detail.

The scope of this literature survey is to examine location privacy threats in relation to other users of the application (friends or public).

2.2 Location Privacy on the Social Web

Advances in mobile devices, wireless communication technologies and more recently location sensing capabilities have enabled the development and use of location based services (LBS). Al-

though users can benefit from taking advantage of their location to obtain personalised services, disclosing location data can pose risks to their privacy, especially since location is distinctive and sensitive in nature, and can lead them revealing more personal information than they wish. Duckham and Kulik defined location privacy as "a special type of information privacy which concerns the claim of individuals to determine for themselves when, how, and to what extent location information about them is communicated to others." [5]. This definition acknowledges three main aspects, namely, when; whether the data collection is on demand or continuous, how; whether location data collection is a manual or automated process and whether it is done with the users' consent or without consent, and finally, to what extent; which refers to the granularity of the location data collected and the data users prefer to share with others [39].

Online social networks have shown an exponential user growth that has been evident in recent years. These networks provide convenient and interesting ways of communicating and interacting with other people which mainly involves sharing information. A seven-year long study showed that users disclose information that is increasing in amount and scope, and at the same time the network is collecting more data over time which can also be accessed by third parties [40]. A study into users' information disclosure patterns in online social networks using field data from 4000 students on Facebook showed users' willingness to share a large volume of their personal information [41]. Such extensive sharing and collection of personal information raises privacy concerns in regard to who can access user data, what types of data are collected and for what purposes they are used [42].

The affordability of location-aware mobile devices has led to the support of location-based services in social networks that embrace geo-location technologies. These are referred to as GeoSNs. GeoSNs in general are social networks that enable sharing geo-location information, yet location data might not be necessarily the core of the service. In other words, users' location is supplementary identification of other primary data such as Facebook, Twitter and Flickr, which can be referred to as Location-Enabled Social Networks (LESNs) [7]. The other type of GeoSNs are Location-Based Social Networks (LBSNs), where location is essential to providing the service, for example, Foursquare, Yelp and Google Places, [16]. In fact, there are two ways in which GeoSNs acquire their users' location: by users visiting predefined venues such as Foursquare, or by locating users both periodically and continuously using their devices' position-aware capabilities, mainly GPSs, or third parties positioning systems such as Twitter [7]. The mode of sharing location information in GeoSNs can be classified into two categories. The first is a check-in by disclosing the whereabouts of users by registering their visits to predefined venues such as Foursquare, and the second is geo-tagging by referencing users' digital content with location information, where users can also associate other users with location, which in this case is called user-tagging [7]. In particular, LBSNs as opposed to LESNs, enable sharing and collection of fine-granularity and detailed personal location information. They

also provide significant semantic data associated with location, such as place name, type and address, as well as allowing users to express their opinions and experience in terms of reviews and tips.

2.2.1 Challenges for Location Privacy in Social Networks

These GeoSNs have witnessed great interest due to their usefulness and entertainment value, attracting millions of online users. For instance, Foursquare has over 60 million users across the world [1]), and 28% of American adults use mobile and social location-based services [43]. At the same time, the privacy gaps and risks of traditional online social networks are also extended to GeoSNs, in addition to the new threats posed from disclosing location information.

GeoSNs have some unique characteristics and powerful capabilities that differentiate them from traditional LBSs. These GeoSNs have expanded the potential privacy threats and implications of using location-based services. These aspects can be listed as follows:

- GeoSNs revolve around the social factor, which considerably impacts how users disclose their location information with others. Users of such services share and interact with many other users, as opposed to the one-to-one sharing seen in traditional LBSs [15]. Tang, Lin and Hong recognised location-aware services that thrive on the social factor of sharing as 'social-driven' applications, whereas traditional LBSs are referred to as 'purpose-driven', since they utilise users' location for a particular purpose [15].

- GeoSNs enable users to share geo-located contents on a relatively massive scale [2]. In fact, there is no limit to the number of user updates that contain location information, such as tweets on Twitter or check-ins on Foursquare. In addition, the usefulness and convenience factors of the mobile applications of these services tend to stimulate users to use them more frequently and share more of their data.

- Geo-social resources can be accessed by a large number of users inside and outside the GeoSN, and are exploited by third party applications that might be connected to such networks [3]. Most GeoSNs support interconnectivity with other social networks, so users can share their information with a wider audience. Although users of these networks can control some level of accessibility to their shared information by others, it will still be available online, and hence might be exposed to potential adversaries, in addition to the legal right of the application and their APIs user to access users' information.

[1] Foursquare, About Us `https://foursquare.com/about`[Accessed: 27-Jul-2016]

- GeoSNs are unprecedented in how they offer location information along with voluminous and diverse user information such as relationships and opinions on one platform. Some of these networks collect large amounts of personal information in order to provide the service [4]. In fact, LBSNs provide significant metadata along with location information including place details such as name, type and address, as well as user comments on this place, pictures uploaded by the user and what places the user likes; contrary to the traditional LBS, which has access only to users' location without any semantic information [8].

- Most GeoSNs support user geo-tagging. Some of these services allow unconsented geo-tagging of their users. In other words, users can be tagged by others in their geo-located resources without their control or consent of these users [7]. Some GeoSNs also enable co-locating multiple users, where users are geo-tagging in the same place and time [7]. These practices allow for implicit derivation of user information based on their social activities (sharing of information) on the network .

- GeoSNs rely on providing real-time user updates, hence high temporal resolution is attached with the shared location information. That is to say, user movement in terms of the current location can be instantly provided online, which can reveal highly privacy-invasive information [7].

As a result, geo-locating users' resource over time opens the door to building and linking users' history [7]. Users' historical location information can be related to contextual and semantic information publicly available online, be used to infer personal and sensitive information, and for constructing comprehensive user profiles which would not be plausible if utilising either social or geolocated data alone [2]. Although user location data collected on GeoSNs can be used to enhance the services provided by those applications, it can potentially be utilised for undesirable purposes that can compromise users' privacy, including advertisements, location-based spam, damaging users' reputation and blackmailing or even physical harm [44]. Disclosing location information on GeoSNs can reveal users' mobility tracks and movement patterns [8, 9, 45]. Users' shared data on these networks can also be exploited to derive more contextual information about their activities, interests and relationships by mapping users' locations to more personal information shared on these services [10, 11, 12]. Revealing users' identity is also feasible when tracing their location over time [5, 25]. Furthermore, absence privacy, introduced by Freni et al., can be inferred as well, such as information regarding users' absence from a particular location (e.g. their home) over a given period of time [3]. A more detailed account on location-based inference studies is presented in Section 2.3.

2.2.2 Effectiveness of Existing Privacy Management Methods

Location Privacy-Preserving Mechanisms (LPPMs) were developed to protect users' privacy, particularly in LBS. A well-known LPPM relies on the concept of anonymity, which is simply isolating the user's identity from their personal information. In the case of location information, methods include the mix-zones model [13] (employing pseudonyms, instead of traceable user identities, that change dynamically when users enter geographic zones that are not pre-registered to particular applications), k-anonymity [46, 47] (a spatial and temporal cloaking technique whereby users disclose an adequate level of anonymous location information that can be mapped to k number of users within a region), and fake locations [39] (injecting a fake location into a user's location data to reduce the possibility of re-identification). Obfuscation is another method, which refers to "the means of deliberately degrading the quality of information about an individual's location in order to protect that individual's location privacy" [24, 47, 25]. Both anonymity and obfuscation methods have proved to be sufficiently effective for protecting location privacy in the case of sporadic location disclosure [48, 25].

However, none of these mechanisms have proven to be fully robust in protecting location privacy against attacks [25]. Anonymity is also ineffective in resisting location attacks when an adversary has access to the location history of a user [5, 39]. Some other work has considered protecting users' location privacy in social location-sharing applications by hiding or encrypting their location from the service provider (the application) [49, 50, 51]. Nevertheless, most public GeoSNs require collecting high spatial (precise location) and temporal (real-time) accuracy for users' location, which restricts the applicability of such mechanisms and impacts the quality of service provided[7]. Moreover, the social dimension adds another level of complexity to the spatiotemporal user data. Users of GeoSNs publicise their location intentionally to share their experience with other users. Hence, this work is concerned with increasing user awareness of the privacy consequences of their location-disclosure actions to enable them to make informed decisions about the information they share.

Next, we review two main methods: Policies and Privacy Settings; within these settings are direct methods (basic access controls) and rule-based methods. Privacy policies are relatively ineffective for online privacy due to the inefficient presentation of information [17]. A study including an online survey followed by interviews with Android users concluded that only 17% of the users actually take notice of the app's access permissions, and only 3% of the online survey participants have shown a comprehensive understanding of the permission screen [52]. Furthermore, privacy policies are static in nature, offering no control over the malicious and illegitimate violation of users' location privacy [5, 53]. Another problem is that the existence of these privacy policies is not enough evidence that GeoSNs are strictly complying with them. For instance, Willis et al.[54] demonstrated a violation of Google's published policy by carrying

out an investigation to examine the scope of personalised results that Google search provides. They revealed that non-contextual ads are shown according to inferred interest from earlier interactions, while Google's policy at that time states that ads shown along with search results are contextual ads based on results' information. Privacy policies tend to be deceptive, where they state general terms about how they handle or use users' data that imply hidden practices that would trigger users' concerns if they were clearly stated [55].

Privacy settings which are mainly in the form of access controls are generally difficult to use and provide limited protection of privacy [18]. Liu et al.[56] showed that users are having trouble correctly configuring their privacy settings and called for new tools to protect privacy. In addition, a US-based survey revealed that 50% of the participants expressed difficulties with their privacy settings management [19]. It seems that users are unable to effectively use protection methods even if they are concerned about their location privacy because they lack understanding of how they work [57]. The offered privacy settings in social networks were shown to tolerate data access by strangers due to permissive default settings [27]. Another study into GeoSNs found that the provided privacy settings in GeoSNs are rather limited as well as difficult to reach; the default settings for personal information are set to be public, and most users never check their privacy settings [4]. Some studies show that location information is leaked by mobile applications without the users being aware of it, due to inadequate privacy settings or the user's inability to employ them effectively [21, 58, 31]. Using privacy settings to restrict profile visibility is a narrow approach to privacy protection since it does not prevent privacy risks resulting from the amount, persistence and quality of data, which most users are not aware of [59]. In addition, social networks provide automated information disclosure such as revealing a user's location in a shared post. This situation makes privacy rather complex to manage since it is insensitive to situational factors where a user is unable to control disclosure at particular times or places, offers inadequate control over the details of what is disclosed, as well as lacking the disclosure control to service providers and third parties [60]. Therefore, privacy needs to be managed using dynamic boundaries balancing, but not through enforcing static rules, as Palen and Dourish argued [61].

Rule-based privacy preferences are an emerging location protection approach where users can manage their location disclosure through specifying dynamic rules. Barkhuus, Brown and Bell [62] demonstrated the most basic form of this method in their application, Connecto, where users are able to set their location manually based on their own privacy preferences . In addition, Poolsappasit and Ray [63] proposed a scalable location privacy model where a logical location will be sent by the system instead of providing the exact coordinates. These are levelled in a hierarchical manner, descending from "the universe" to more specific places names that aid users in choosing location granularity according to their privacy preferences. In this model, users' location disclosing decisions depend on who the requester is, and the time and place

of request. In LOCYOUTION, developed by Tsai et al. [37], users were capable of specifying time-based rules under a 'My Rules' tab that can set the day and hour range during week days for location sharing . PeopleFinder, by Sadeh et al.[38], also provided rule-based privacy controls where users can define their privacy preferences by determining the location (places on a map), time (days and hours) and requestor (individuals or groups) of location sharing. Similarly, Patil and Lai [64] asked users of MySpace to define permissions for their personal information sharing by setting various boundaries and linking them to four user categories: global, team, groups, and individuals. Furthermore, Kelley et al. [65], as a way of examining users' behaviour regarding disclosing their location to mobile advertisement systems, examined users' sharing time with these systems when using six preference-dependant privacy techniques, namely: opt-in, time, time with weekends, location, location and time, and location and time with weekends. Jagtap [66] developed a policy-based framework for enabling users to control their context and location privacy in terms of managing their personal information collected by their mobile devices, where any information sharing request received is reasoned against the user's specified privacy preferences.

Nevertheless, rule-based privacy policies are not fully effective since they have no consideration for contextual factors of being in an atypical location at a certain time, where about 65% of the participants' responses represent a mismatch between the pre-defined privacy preferences and the in-situ sharing attitude [36]. Moreover, it is relatively difficult for users to express effective privacy preferences [38]. Thus, users of GeoSNs are still vulnerable to privacy attacks. One study showed that users' data, including location, shared on social networks is not protected against malicious exploitation, even from other users of LBSNs [67]. Besides currently lacking the support of effective location privacy management approaches in GeoSNs, users' awareness of the potential privacy threat resulting from location disclosure has an important role in stimulating the use of the available privacy controls. Another study concluded that users are not able to make use of these solutions if they are not aware of the violations that are actually taking place [29]. It has also been demonstrated that people who have personally experienced privacy invasions are more likely to utilise their privacy settings[59].

These available means for protecting users' online privacy have proven to be inadequate, and offer only restricted control over their data. The later rule-based approach presents a promising solution for location privacy management, yet needs more improvement to embrace all aspects of location disclosure on GeoSNs. More detailed discussion about the role of privacy awareness is introduced in Section 2.4, followed by a review of relevant privacy-enhancing methods in Section 2.6.

2.3 Location-Based Inference on GeoSNs

A significant interest can be witnessed in research studying the value and utility of location information on GeoSNs to understand users' behaviour. The means used for user location-based profiling can be systematic using algorithms, as reviewed below, or as simple human-based strategies and observations [68, 69].

2.3.0.1 Implicit Location Extraction

Studies have utilised publicly available information from GeoSNs to derive or predict users' location. Cheng et al. [70] estimated Twitter users' city-level locations by exploiting their tweet contents which was sufficient to predict more than half of the sample within 100 miles of their actual location. Similarly, Pontes et al. [71] examined how much personal information can be inferred from the publicly available information of Foursquare users, and determined the home cities of more than two-thirds of the sample, within 50 kilometres. Sadilek et al. [72] investigated novel approaches for inferring users' location at any given time by taking advantage of knowledge of the GPS positions of their friends on Twitter. Up to 84% of users' exact dynamic locations were derived even when setting their location data as private, whereas accuracy of 57% was accomplished by knowing information of only two friends. Location profiles of Twitter users were constructed by exploiting the following connections and tweets of these users in terms of home location prediction and inferring users' multiple locations and explaining the following relationships, where 62% of users' home locations were derived [73]. Another study conducted on the social web showed the ability to detect users' location based on utilising rich contextual information including place names, geo-tags, and local linguistic context mined web resources [74]. The geo-location of the users of some LBSNs was extracted with high accuracy, and the top 5 locations were also inferred [67] . In addition, Huo et al.[75] demonstrated that it is feasible to infer the users' hidden location that they have visited without sharing it by exploiting the check-in history of this user's on GeoSNs and her social ties as well .

Interestingly, it is possible to infer future movements of GeoSNs' users. Gao et al. [76] formulated predictive probability of the next check-in location by exploiting the social-historical ties of Foursquare users. They were able to predict with high accuracy possible new check-ins for places that users have not visited before by exploiting the correlation between their social network information and geographical distance in LBSNs [77]. Pietro [8] predicted users' future movements by exploiting users' previous transition history between venue category as well as their probability of returning to a previously visited venue category, which demonstrated an ability to estimate users' location even if with a limited history of a user.

2.3.0.2 Social-ties Inference

Other works focussed on investigating the potential inference of social relationships between users of GeoSNs. Crandall et al. [78] investigated how social ties between people can be derived from spatial and temporal co-occurrence by using publicly available data of geo-tagged pictures from Flickr. They found that relatively limited co-occurrence between users is sufficient for inferring high probability of social ties. Sadilek et al. [72] also formulated friendship predictions that derive social relationships by considering friendship formation patterns, content of messages of users and their location. They predicted 90% of friendships with accuracy beyond 80%. Additionally, Scellato et al. [79] investigated the spatial properties of social networks existing among users of three popular LBSNs and found that the likelihood of social connections decreases with distance. [12] developed a link prediction system for LBSNs by utilising users' check-in information and properties of places. 43% of all new links appeared between users with at least one check-in place in common, and especially for those who have a friend in common.

2.3.0.3 Human Behaviour Detection

Sharing location information on GeoSNs can be utilised to analyse and predict spatiotemporal user behaviour, including their interests, activities, mobility patterns and future movement. Vosecky [10] modelled users' interests shared on microblogs in relation with their corresponding disclosed locations, where users' geographical information from Twitter was extracted whether the location is directly tagged by them or mentioned in their tweets. Then, users' geographical regions of interests are derived that represent clusters of geographical activities.

Studying and extracting spatiotemporal movement and activity patterns of users on GeoSNs has attracted much research in recent years. Dearman et al. [11] exploited location reviews on Yelp in order to identify a collection of potential activities promoted by the reviewed location. They derived the activities supported by each location by processing the review text, and validated their findings through a questionnaire that showed a mean precision of up to 79.3%. Noulas et al. [45] studied user mobility patterns in Foursquare by considering popular place categories, the time interval between two consecutive check-ins, place transition from one place category to another, and spatial and temporal patterns. Cheng et al. [9] examined a large-scale data set of users and their check-ins to analyse human movement patterns in terms of spatiotemporal, social and textual information associated with this data. They were able to measure user displacement between consecutive check-ins, distance between users' check-ins and their centre of mass, as well as their returning probability to venues. They also studied factors affecting users' movement and found considerable relationship between users' mobility and geographic

and economic conditions. More recently, Preotiuc-Pietro et al. [8] investigated the behaviour of thousands of frequent Foursquare users. They analysed users' movements, including returning probability to places, check-in frequency, inter-event time, and place transition among each venue category. They were also able to group users based on their check-in behaviour, into categories such as generic, businessmen or workaholic, as well as predict their future movements. Similarly, in [80], users' check-ins were analysed for the purpose of revealing potential similarities in mobility behaviour between different users' groups, such as their favourite places or their place pattern. They were also able to calculate time and distance between check-ins as well as co-location with other users or friends.

In addition, location and activity recommendation systems for GeoSNs provide evidence of the feasibility of predicting users' future movement in regard to location or activity preferences based on gathering and utilising their historical location and contextual information. A travel route recommendation method that exploits geo-tagged pictures in geo-social photo sharing application, Flickr, was introduced in the work of Kurashima et al. [81]. They developed a probabilistic photographer behaviour model that estimates what landmarks users are likely to visit considering the user's current location and preferences, and predict users' personal interests. Another study, carried out by Kurashima et al[82], proposed a model for analysing the location log data of users in order to recommend places to visit. It estimates the users' topic interests and their spatial activity areas depending on the geographical features of the visited places derived from geo-tagged content in GeoSNs. Based on these estimations, this model also predicts the visiting behaviour in daily life, and can therefore recommend new places for the user to visit when they are in unfamiliar areas. In [83], as part of their work on how the temporal aspect can influence the effectiveness of location recommendations, they derived and clustered geographic activities by utilising users' history of geo-tagged photos in order to infer the Areas of Interest (AOI) for recommendation.

2.3.0.4 User Identification

Disclosing location information on GeoSNs can be used to identify users. Rossi and Musolesi [6] proposed and tested three approaches to identifying users' by exploiting their check-in information on LBSNs, particularly their spatiotemporal tracks, frequency of visits, and social ties in LBSNs. Evaluation showed that only a small amount of check-in information was necessary to identify users with high accuracy, where up to 80% of users were successfully identified in some datasets. Zhong et al.[84] were pioneers in exploiting the predictability aspect of location check-ins in order to develop location-to-profile framework that infers demographics of users. They derived enriched check-in semantics based on three main factors that are related to the user profile: spatiality, temporality, and location knowledge, such as customer review sites and social

networks. A series of experiments were carried out on the dataset that revealed that feasibility of deriving users' demographics from their check-in information, where gender and educational background attributes provided the best outcomes, followed by age, sexual orientation, marital status, blood type and zodiac sign. More recently, researchers have exploited GeoSNs to explore the personality aspect by examining the reciprocal relationship between users and spatiotemporal features. In the work of Chorley et al.[85], a study was conducted with the aim of understanding human behaviour in terms of examining the relationship between the type of locations visited by Foursquare users and their personality. A five-factor personality model was used as a way of studying users' personalities. As a result, there were correlations between the five-factor personality traits and the Foursquare check-in attitude.

The above studies show a significant potential for deriving personal information from GeoSNs and hence also imply the possible privacy threats posed to users of these applications. These studies considered extracting types of information from location datasets for a large group of users to determine general patterns and collective behaviour. Hence, in this work, we answer the question of how location data for a single user can be used to infer personal information and how it impacts this user's privacy, with the aim of understanding possible implied user profiles from location data stored in GeoSNs. This work does not focus on improving or testing information inference algorithms, but mainly uses the previous studies as a foundation for the possible personal information inferences that can be made from location disclosure on GeoSNs.

2.4 Location Privacy Perception and Disclosing Behaviour: The Privacy Awareness Gap

Individual's sense of privacy varies. A pioneering study of users' feelings towards their privacy, conducted by Westin in 1999, showed that 25% were privacy fundamentalists with a strong interest in protecting their privacy, 54% were privacy pragmatic, and try to balance between having the benefit and their personal privacy , and 22% were the privacy unconcerned. In 2003, carrying out the same study showed that the privacy pragmatics increased to 64% and privacy unconcerned fell to 10% only [86]. Online users tend to have concerns regarding the disclosure of their online personal information and what it is used for. Wang et al. observed a number of relevant studies conducted between 1998 and 2003, and showed that 70% - 89.5% of Internet users are actually concerned about the privacy of their personal information, and 89% - 90% of people are concerned about their data being shared for a different purpose than originally intended[87].

Studying the privacy attitude and behaviour of Web users has been a subject of interest for the

research field. These aspects have been subsequently examined in the areas of LBS, social networking and, more recently, geo-social location sharing platforms, in which privacy concerns, awareness and information-sharing habits are questioned. Related studies are classified first based on their evaluation approaches and methods used, as follows:

- **Evaluation approaches:**
  - Quantitative and qualitative user-based experiments to validate research goals using relevant subjects, which are used by the majority of research in this field
  - User data sets utilised to observe patterns and use behaviour to generate data without actually self-reporting [40, 16, 88].
- **Methodologies:**
  - Online surveys [20, 59, 26]
  - Hypothetical use and location-sharing scenarios [89, 90, 91, 29, 4]
  - Field trail and lab experiments including prototypes [15, 62, 65, 64]
  - Projection (visualisation) of users' data [92, 57]

In the domain of LBSs, users tend to be generally concerned about their privacy, yet their concern decreases as long as they gain benefits [89] or when using the service within a known and restricted environment [14]. Privacy concerns seem to be stronger when using online social networks where users can share more diverse personal information with a wider audience. For example, a survey on Americans' use of the Internet showed that 63% of profile owners have removed people from their networks, 44% stated that they deleted comments that others shared on their profile, and another 37% removed their names from photo-tags that reveal their identity [19]. Another Facebook study found that the amount of publicly displayed data decreases with time, where users restrict their visibly as a way protecting their data [40]. The social factor of location-sharing can profoundly impact users' privacy attitude. Tang, Lin and Hong [15] drew attention to how the social aspect of the social-driven location-aware applications affects users' privacy perception when sharing their location. They found that privacy concerns were a determining factor for social location disclosure, since users preferred sharing semantic location names and obscuring their location information. Another work showed that users prefer to set their location manually using a vague location as a way of managing their location exposure, which resulted in users being more comfortable with sharing their location [62].

Other factors that impact users' location-sharing decisions have also been investigated. Lederer et al.[90] carried out a web-based survey that tests users' social location sharing using hypothetical scenarios. They stated that users' privacy preferences in terms of location data granularity

are determined first by the requestor identity and then by the situation that these users are in. Similarly, Consolvo et al.[91] investigated people's reactions towards imaginary location inquiries from their social relations, and found that participants' responses to location requests depend on the three essential aspects: who the requestor is, why the participant's location is needed, and what the granularity of location information would be in order to better serve the requester's purpose.

Growing research interest has been witnessed over the past few years in studying users' attitudes and concerns regarding their location privacy in social location-sharing applications. Patil and Lai [64] showed that users were most concerned about their location information when using a social application that presents live feeds about users' information, and wanted to utilise the offered privacy setting to protect their privacy. Furthermore, Kelley et al.[65] examined users' behaviour regarding disclosing their location to mobile advertisement systems and found that they were highly concerned about their privacy, especially when sharing their location with corporate-oriented parties. Using visualisations of location tracks, Tang, Hong and Siewiorek [92] found that the majority of participants stated their concerns related to their physical privacy when showing them their visualised location history, and they selected the least information-exposing visualisation to share with others. Similarly, Brush, Krumm and Scott [57] demonstrated that users were also concerned about their location privacy, and selected obfuscation techniques that enabled them to hide their sensitive information.

With regards to public GeoSNs, there are relatively few research works that examine the privacy concerns of users. Lindqvist et al.[26] considered users' motivations for using Foursquare and questioned their privacy concerns using qualitative (18 users) and quantitative (219 users) surveys. Their analysis showed that most of the participants had some concerns about their privacy, and users who were more concerned about their privacy chose not to check into their private residence or to delay checking into places until after they leave, as a way of controlling their safety and privacy. A similar observation was noted by Jin et al. [16] using a dataset containing thousands of check-in details from Foursquare, where it was found that users were generally aware of the privacy of their place of residence and tended not to provide full home addresses or blocked access to their residential check-ins to other users.

The above studies reveal that location privacy concerns are evident, yet users can be influenced by incentives to share their location information. As a result, the Privacy Paradox can be observed when having the contradictory behaviour between the users' privacy perceptions (being concerned) and their actual information disclosure attitudes (sharing their information), which can occur due to their insufficient understanding of the link between their sharing actions and resulting location privacy implications [93]. For example, a survey-based study on Facebook using 318 participants showed that even concerned users reveal a considerable amount of their

information due to having misconceptions about how privacy threats can occur [20]. Another survey on 119 Facebook users found that the perceived benefits of using social network services can impact how users appreciate the privacy risks [59]. It also revealed that users accept friendships from people who they do not personally know, which allows strangers to access a large amount of their personal data. In a similar Facebook experiment, users were willing to share a large volume of their personal identifying information, including real names, birthdates and current address [41]. It seems that lacking awareness of how users' location information can be utilised contributes to their ability to appreciate their location privacy and impact their perceptions. Krumm [39] suggested that people would not adequately understand the potential detrimental consequences of penetrating location information unless stories of others being actually endangered were published.

People's awareness of the practices carried out on their data and related privacy implications can play an important role in shaping their privacy perceptions and hence behaviour. Users of social networks are not fully aware of the privacy threats that can occur as a result of the information published on their account on these networks [22]. They showed to share information willingly without clear comprehension of who can access their data and how it can be processed or manipulated [20, 27]. Furthermore, Rader [29] conducted a scenario-based online survey to investigate users' concerns regarding possible information collection and inferences made based on their online behaviour provision using 701 web-savvy users recruited from Amazon Mechanical Turk (MTurk). The study revealed that even the web-savvy users of online services are not fully aware of data collection made by third party services, and they were more concerned about unwanted access to their data when knowing the possible privacy implications of information sharing behaviour. The design of these GeoSNs focuses on providing an enjoyable and useful experience for users, yet lacks support of awareness of data collection, accessibility and utilisations [26, 4]. A lab study that investigated users' perception of possible unauthorised information disclosure including location showed that none of the participants had a profound comprehension of data sharing and its dimensions [23]. They were also unaware of potential sensitive information collection by mobile apps, and were most shocked about how often sensitive data is leaked and the destination of that leak. In a similar study, most of the participants stated that their location was disclosed to unexpected apps or more frequently than they anticipated [58]. Moreover, a user-based study was conducted where the participants were asked to perform pre-defined scenarios and tasks on two location-based services, and showed that the participants felt safer using the LBSN that actually collects the most of their personal data and pose greater privacy threats due its social and playful aspects [4]. Additionally, in GeoSNs, user data can be revealed indirectly by being shared by others on the networks or without the user's explicit consent, which is another dimension of compromising user privacy. A user survey that investigated privacy threats and awareness of shared media in GeoSNs showed that 52% of par-

ticipants found out about pictures of themselves by coincidence, and up to 34% of pictures from two Flickr datasets contain GPS data [88].

Consequently, a gap of knowledge can be noted in users' privacy perceptions about the potential privacy implications of disclosing location-related information on these GeoSNs. In order to raise awareness and increase the perception of personal location privacy, users need to comprehensively understand what they are really disclosing and who has access to it [93]. Gambs et. al.[2] suggested that the transparency of these GeoSNs needs to be improved by integrating a 'privacy lens' that present visually how user's profiles are shown to others . Thus, user awareness of the practices carried out on their data in terms of collection, utilisation and accessibility can educate them about the potential privacy consequences and hence enable them to provide informed consent about their information disclosure that matches their desired privacy preferences [94]. Kang et. al.[28] argued that the extent of privacy management action undertaken by users is related to their personal awareness and experience with how a system processes their data, which is limited, especially for non-technical users. Gross and Acquisti [41] found that users of social networks generally underestimate the privacy implications associated with their information disclosure, and hence are less likely to use the provided privacy settings. Similarly, Rader [29] showed that users are unable to make use of design and policy based privacy solutions if they are not aware of the violations that are actually occurring.

Limited work has been carried out on users' privacy attitudes and behaviour in the domain of public GeoSNs. What has been conducted was constrained only to questions users' motivations to use Foursquare, and their residential privacy in this application. The specific characteristics of data shared on GeoSNs and the nature of user interaction on these networks suggests the need for more specific studies of location privacy on these networks to evaluate privacy attitudes and behaviour which is addressed in this work. Factors that can influence privacy attitude and location-disclosure behaviour in this domain are also investigated in this work. Considerable research on examining the impact of privacy feedback and visualisation systems on users' privacy perception and attitude showed that many participants change their privacy permissions effectively due to the projection of relevant privacy information that raised their awareness of the undesired implications. These studies are discussed in detail in Section 2.6

2.5 Privacy Models and Frameworks

Privacy models provide principles and guidelines to be considered when designing a privacy-aware system. They present insights into how to design or assess a system that serves users' awareness of potential privacy implications based on their interaction with it, convey these insights clearly, and also suggest means of effective privacy management by users. These models

have common aspects, but can vary based on the application and privacy domains. A pioneering privacy model was introduced by Bellotti and Sellen [95], who proposed a design framework for privacy in ubiquitous computing as a way of addressing privacy problems resulting from a lack of user awareness of ongoing activities concerning their data. This framework suggests that the design drivers of a system should be to provide feedback and control about information collection, how it is processed, its accessibility, and purpose of use. The framework also identifies eleven design criteria that can help in evaluating design solutions. Adams and Sasse [96] identified three main privacy factors based on previous work as a way of finding boundaries under which a privacy breach can take place. The first is information sensitivity, which is concerned with the degree of sensitivity or privacy of data, and can vary among individuals. The second is the information receiver, which is concerned with who knows what about a user. The final factor is information usage, which deals with how the data are utilised and for what purposes. In addition, Langheinrich [97] proposed six general guidelines for privacy in ubiquitous systems based on design concepts that mainly include informing users of data collection, giving them to explicit control over of their participation and information about their data accessibility. Lederer et. al. [98] introduced five design pitfalls that need to be considered when designing interactive systems. They place emphasis on the transparency of potential and actual information flow in terms of collection, accessibility and utilisation. The rest involve allowing users to manage their privacy easily using top-level privacy settings and established social practice. Moreover, Friedman [94] developed a conceptual model for informed consent for information systems in the context of online interactions. This model describes the practice of informed consent depending on offering easily-interpreted feedback on benefits and risks of information-disclosing actions. It also assures users' right to be explicitly consented and to determine their participation under free will with minimal distraction from their main task.

In the domain of location privacy, Shokri et al. [99] provided a unified framework for recognising and clustering essential elements of location privacy. It basically describes how privacy threats can take place in relation to users, applications and LPPMs. The model revolves around three main factors: means available to adversaries in terms of accessibility and knowledge about the system, actions undertaken by them considering the spatial and temporal scopes, and the purpose of such action, whether to reveal user existence in a place, or track or identify them on an individual or group basis. Similarly, Zafeiropoulou et al. [100], 2012 proposed a framework that analyses essential properties of location data that involves classifying the information based on its complexity, identifiability, quality, consent and source.

In this work, relevant principles of these privacy models were considered, specifically in the design of the location privacy awareness solutions proposed in this work, and they were moulded to serve the domain of location privacy in GeoSNs. In addition, an analytical framework that discusses the factors contributing to the location privacy problem on GeoSNs is necessary to

understand the dimension of this problem in this particular area.

2.6 Privacy-Enhancing Technologies for Privacy Awareness

As discussed in Section 2.4, awareness of the implications of information-sharing, including location in GeoSNs, has a key role in empowering users and enabling them to manage their privacy effectively. In particular, it allows users to clearly perceive related privacy consequences and make informed decisions towards their data disclosure. Hence, information and privacy are difficult to manage when user interpretations of online privacy are constrained [101]. On one hand, online users seem to be only conscious about the task they are currently involved in, and can lack cognition about their previous interactions and sharing of information. On the other hand, they often fail to recognise the possible privacy threats associated with their sharing actions, since the service providers tend to hide them for the purpose of maintaining their business [55]. Thus, according to Vihavainen et al. [60], the service providers need to make their practices transparent in order to address user privacy concerns. This would include supporting user understanding of how these practices work and their consequences. Debatin [59] demonstrated the need to educate people about potential privacy risks in a way that actually modifies their behaviour. Privacy awareness approaches that enable informed consent of data sharing and use can be utilised to address location privacy in GeoSNs. These methods not only enable users to effectively manage their privacy, but also take into consideration the varying interpretations of privacy, where individuals can decide upon their sharing actions according to their personal privacy preferences [35]. The approaches used in this area can be classified under three main groups:

- *Privacy feedback and notification systems* that alert users about threats to their privacy related to information-sharing action carried out by them.
- *Design implications of such systems* that consider specifically improving the usability and effectiveness of privacy notice by utilising the design aspect.
- *Visualisation techniques for privacy* that focus on employing information projection and presentation to users as a way of achieving privacy awareness of their interactions with systems

Relevant work in these groups is reviewed in the next sub-section. Evaluation methods and experimental design can also be clustered to give a brief idea of how these studies were conducted, as follows:

- **Evaluation methods:** mainly in the form of quantitative and qualitative user-based studies using
  - *Lab and field experiments* using running prototypes (e.g. [30, 21])
  - *Simulations and mock-ups* (e.g. [102, 38])
- **Experimental design:**
  - *Within-subjects* in which all (or one group of) participants is exposed to all study conditions (levels of the independent variable) in order to capture their influence on the same subject [103](e.g. [21, 31, 104])
  - *Between-subjects* in which participants are divided into different groups where each group is randomly allocated to one of the study conditions in order to objectively gauge subjects' feedback [103](e.g. [30, 102, 58, 64])

2.6.1 Online Privacy Feedback and Notification Systems for Privacy Awareness

Feedback and notification tools are commonly used for warning users about security and privacy risks on the web. They mainly include privacy nudging, which is an approach for allowing users to make informed decisions without constraining their freedom by designing informative systems [105]. Studies are emerging that assess user awareness of privacy implications and the impact of such tools on users' attitudes while interacting with systems.

2.6.1.1 Feedback and Notification Systems in Online and Social Applications

Feedback and notification tools have been used widely on web applications with the aim of warning the user of any security or privacy risk they may encounter while using the service. Some studies have specifically investigated how such systems can influence users' awareness of the privacy implications, and hence their attitude. In particular, Malandrino [30] examined whether informing users' about hidden activities that are running while web browsing, which may threaten their privacy, can motivate them to take action to prevent them. They developed a browser-embedded privacy-improving tool that informs users about any privacy-threatening activity and allows them to manage their privacy. They quantitatively studied whether this tool can increase the transparency of implicit information leakage during web navigation and enable the control of privacy by recruiting two groups of 18 ICT (technology-related) and 18 non-ICT students for evaluation. All of the participants were willing to change their online attitude and the tool was effective in raising their awareness of the privacy risks, as well as stimulating them

to take steps to protect their privacy. In addition, Balebako [21] developed a privacy notification tool that provides feedback to users about what of their sensitive information has been illegitimately leaked through these applications, the frequency of leaks, and the recipient that the information has been disclosed to. It shows just-in-time notifications to instantly inform the users about a data leak and displays a frequency-based summary of the illegitimate disclosure using a grid layout. The researchers ran a lab study for their evaluation, where 19 participants were asked to play two games and then answer questions, first without the tool, and then with it, followed by interviews. Feedback showed that none of the participants initially lack understanding of the implications of data sharing or sensitive information collection by the app. Using the tool increases their awareness of their information leakage, and they showed their interest in using such a tool. Similarly, Almuhimedi et al.[31] studied how a permission manager for mobile applications and privacy nudges can increase users' awareness of data collection by these apps. The researchers carried out a field study where they implemented AppOps- the permission manager. The privacy nudge was designed to show which app has access to what data and how often. It also offers the options to change app access permissions or view more details. Firstly, 23 participants installed AppOps without accessing it, and the second phase included accessing AppOps where they can change and restrict the app permissions. Finally, they started receiving privacy nudges. As a result of the second phase, 95.6% checked their permissions and 65% restricted 272 app permissions. 70% of them restricted a further 122 app permissions in the last stage, where 78% restrictions were reactions towards a nudge. Moreover, Bargh [106] discussed how feedback systems can be employed to notify the data disseminator of any detected privacy breach by analysing feedback mechanisms and categorising them by completeness (complete and partial) and timing (real-time and delayed) factors.

Other studies have employed feedback mechanisms for increasing users' privacy awareness in the domain of social applications. Fire [32] implemented a Facebook application that enables the user to enhance their security and privacy, where it distinguishes which of the user's friends can pose privacy threats by examining their in-common activities, and hence determines whether their accessibility to the user's personal information needs to be restricted. This tool also warns users about their installed Facebook applications that can access their private information. Evaluation showed that 42% of 74 participants restricted their information access to 392 friends. Furthermore, Emanuel [102] developed a privacy warning prototype that provides real-time alerts for social network users about the risks of disclosing their information in order to enable them to make informed decisions about their information exposure. This was implemented as a plug-in in the users' browser, where it scans what personal information users entre in to a website, manually detects the type of information entered, and then warns the users about what other personal information can be inferred. The system was tested using two fictitious dating and professional social networks . 41 participants were asked to review their profiles before

submitting, where the privacy feedbacks were shown to half of the sample. They were also asked to set privacy accessibility conditions for their information. As a result, users provided with the feedback system reduced their disclosed information in the review stage (by 7.7%) and specified stricter accessibility levels, whereas users with no feedback system provided even more informing in that stage . Wang et al. [104] designed a privacy nudging tool that helps Facebook users consider their post's content before sharing it on their profiles using visual cues. It shows profile pictures of a random sample of who can view the post, a timer-based interface allows for post-editing, and indicating how others would perceive this post using sentiment analysis. A trial was carried out on 21 participants by installing the tool as chrome plug-in in two stages; the first was with the original Facebook interface, and the second with the tool. Although the quantitative analysis of system log do no show significant impact of the nudging on user behaviour , the exit survey showed that it actually had positive affect on some participants.

2.6.1.2 Feedback and Notification Systems for Location Privacy in Geo- Social Applications

There has been recent interest in employing feedback techniques for the purpose of studying location privacy. Fu [58] conducted a field study to investigate location disclosure on run-time. They first developed a novel approach using an app for detecting location access by location-enabled mobile applications. The participants were randomly assigned to either a Disclosure group (13) where they can know about the location access from their mobile app, or a No Disclosure (9) group with no access feedback. They found that an Android notification system for location disclosure is not effective for informing the users. Most of the participants stated that their location was disclosed to unexpected apps or more frequently than what they anticipated, hence they took steps to manage their app's location access. Jedrzejczyk [107] examined how real-time feedback is efficient in supporting the social translucence concept in terms of awareness, visibility and accountability when disclosing sensitive information in location-sharing applications. They implemented a mobile location sharing application that builds upon the social translucence concept. This shows a friends list and their location on a map, presents how friends can view their profile, and allows the user to define preferences. They carried out interviews (No=5) and a field trial (No=12). As a result, real-time feedback was shown to influence people to be more accountable action-wise, and that they were more willing to reject unnecessary location requests. Similarly, Patil and Lai [64] developed an interactive graphical representation of the users' actual workplace that offers live feeds about users' information and allows them to specify rules for sharing location, activities, and availability. 36 participants were asked to set information-sharing permissions using pre-defined accessibility groups. They divided their participants into three groups, based on information disclose to the system and feedback of

permission implications. They found that participants were most concerned about their location information, and they effectively utilised the group-based permission feature. In addition, Sadeh [38] developed a geo-social application where users can control their location requests from others by enabling them to specify their privacy preferences (rules) based on time, location and requestor(s). In addition, the application shows instant notifications including details about users' location request. Evaluations included using an initial lab experiment (No=19) with a mock-up application that introduces situations derived from their real life and communication with their social relations, and a field experiment (No=28) over several weeks using a deployed version. They found that methods that raise users' awareness about the way the application operates and provide comprehensive explanations tend to stimulate users to produce more accurate preferences and increase the user's trust in the application.

Tsai [37] implemented a mobile location sharing system which utilised Facebook. This application exploited WiFi access points to specify users' location, where users can specify time-based rules for their location exposure and can see who viewed their location. Participants installed and used it for a one month period and were split into two groups: one with feedback access, and one without. Evaluation showed that users who had feedback about who was querying their location were more willing to share their location, had less privacy concerns, and made themselves more available to others. Christin et al. [34] used picture-based privacy warnings to increase user awareness about potential privacy threats. The warning is shown to the users based on their selected privacy sharing preferences, including location in terms of the recipient and data granularity, in which they can choose to change their sharing settings or continue without change. The continue button is colour-coded to indicate the level of data granularity set by the user, where green represents coarse, orange moderate, and red fine granularity. A history view was also proposed to enable users to see who acquired what of their data, the granularity, and the time of the request. A user study (No=30) was carried out with participants' performing three instructed tasks. Consequently, participants understood the warnings and 70% of them wanted to change their privacy settings. More recently, Patil [36] studied how employing feedback and control mechanisms can impact users' location sharing decisions and how these vary from their pre-defined privacy preferences. They implemented an Android application that collects participants' locations, gave them a questionnaire of hypothetical location requests, and captured their attitudes. 35 participants were asked first to specify access permissions based on the day and time, and then were divided into two study conditions: feedback condition in which participants are notified about location disclosure made based on their pre-defined preference, and decision condition in which participants are enabled to respond to a location sharing request. The findings showed that 65% of the responses represent a mismatch between the pre-defined privacy preferences and the in-situ sharing attitude, which suggests that using general privacy preferences is not effective, since they have no consideration for contextual factors be-

ing in an atypical location at a certain time. More importantly, immediate feedback of location disclosure triggers users' feeling of being overexposed (51%), whereas this feeling is significantly less when the users were in control of choosing their location sharing regardless of their pre-specified preferences (10%).

The above studies generally assume that awareness of location information is confined to the visits the user makes to places and focuses on the accessibility aspect of user awareness, but does not consider a holistic view of possible personal information inferences that may be made in the network, as described in Section 2.3. They also may have been confined by the limited features of the proprietary applications used in testing that do not fully represent the GeoSNs' environment. This work attempts to fill this gap by utilising both lines of research in studying user awareness in public GeoSNs, using similar user-based experiments.

2.6.2 Design Implications of Usable Privacy Information Presentation and Notification

Usability of privacy notices and feedback tools is of relevance to this work. It focuses more on examining the design implications of the privacy feedback systems in order to ensure the effectiveness of such a system in conveying warning of any potential privacy risk to users. Some studies examined the behavioural impact of improving the interfaces of existing applications. Kelley et al.[17] tackled the problem of users' lack of understanding of privacy policies by proposing a new label-based presentation inspired by nutrition labelling. They developed a privacy label for presenting how their personal information is collected and utilised by organisations through an iterative design approach. A laboratory user study was conducted where participants were shown examples of natural language and label formatted privacy policies and asked to answer relevant questions. They found that the participants were able to answer the questions faster and more accurately when using the labelled version. In addition, Liccardi et al. [108] modified the permission interface of Google Play to show a sensitivity score to indicate the applications' rights to share personal information and to highlight the related permissions. Their validation experiment showed that the enhanced interface design actually motivates participants to select apps that have the lowest access permissions. In the social application field, Wang et al. [18] carried out an experimental study to investigate how improving the limited designs of the current privacy authentication dialogs on Facebook can impact behaviour in regard to granting access. They developed four designs that offer the user different levels of granularity for access permissions and presentation styles. Facebook users were recruited from MTurk to participate in an online experiment where the user is redirected to the semi-functional implementation of one of four designs or left with the current default dialog. Consequently, the users who interacted with the default Facebook authentication dialog recorded the highest installation rate for

the third-party application (84%) among users who interacted with one of the four proposed designs (72.5% on average).

Other studies focused more on how to improve privacy indicators to successfully attract a user's attention to a potential threat. Wang et al.[18] carried out an online experiment to investigate how the cues of security warnings influence users' behaviour towards mobile websites. A mobile website simulation was implemented for this experiment in four versions, where the absence and presence of security warnings and instant gratification cues varied in the registration page of each version. Evaluation ran on participants from MTurk. The findings showed the cues to have a significant impact on the users' attitude by enhancing the privacy perception of the potential personal information risk. Similarly, Shi et al. [109] formed three design principles of security indicators in web browsers based on visually attracting users' attention, offering suitable icons for information mapping, and indicating the security status for various website types. They implemented a prototype of new designs for Firefox which was used in a user study. As a result, the majority of the participants favoured the new design in terms of effectively drawing users' attention and hence warning them of the websites' frauds . Bravo et al. [110] developed five inhibitive attractors that aim to reduce users' chance of making potentially harmful actions by obliging the user to wait for a period of time or perform another action . These attractors only appear when the users hover over the harmful choice. Three user-based experiments were conducted using MTurk in third-party websites providing online games that show warning dialogs, including different attractors for highlighting important information. They revealed that participants exposed to the attractors showed less likelihood of ignoring the indicators for the potential threat and were more capable of making informed decisions. Moreover, Maurer et al.[111] introduced an in-context warning concept that immediately shows a security alert right next to data after it has been entered by users to attract their focus to the critical data they have provided. They implemented a Firefox plugin and validated it with a lab study followed by a focus group session. The results indicate that the plugin enabled the participants to recognise fraudulent websites and avoid them. Hughes-Roberts and Kani-Zabihi [112] explored the role of the user interface in informing users about their privacy during interaction with social networks. Three designs were developed: grouping personal information based on their disclosure consequences represented using a traffic light, showing a recommended action based on other users' actions, and offering a review screen to display a privacy score based on the sensitivity of the disclosed information. An experiment was conducted including three groups for each of the designs and a control group. The findings showed that participants in control group shared more information than other groups. Bravo et al. [113] also carried out a user study with advanced and novice users to investigate what can lead to ignoring security warning by presenting them with five random warnings and scenarios related to them . They found that the novice users tend not to care about the sensitivity of the entered information, which implies that it should be noted

in the warning. Moreover, novice users seem to consider the safety of an action after performing it, and take less steps to assure their safety than advanced users. The researchers also suggested using less text in the warnings, since participants tend to not read them thoroughly.

As demonstrated in the above studies, the design aspect of security and privacy systems contributes significantly towards facilitating information awareness and hence triggering a desirable user attitude. This work considers the design element in providing location privacy solutions in the context of GeoSNs that can effectively contribute to increasing users' awareness .

2.6.3 Visualisation

2.6.3.1 Visualisation for Online Users' Profile

Many studies have employed visualisation techniques as a way of presenting users' profiles due to their simplicity in conveying information [114]. Middleton et al. [115] developed a system for searching paper databases and online paper recommendation . The system offers profile visualisation for users that displays their interests and enables the collection of further information on their interests. User interest profiles are derived and updated daily and visualised to users as a time/interest graph. Heer and Boyd [116] presented a visualisation system that enables end-users to explore and analyse large-scale social networks by simplifying the discovery of their online communities and awareness of information exposure. It enables the accessing and searching of users' profiles on an online dating site via visualisation, and delivers the social network based on a common node-link layout, where nodes are the users and links are the friendships. Evaluation revealed that users were able to explore their online community while enjoying the experience, and they used the system's features effectively. Similarly, Tchuente et al. [117] introduced a method for presenting temporal graph visualisation of users' short and long-term interests that considers the evolving interests of users. A third-party application was implemented using the Facebook API which 85 users installed in their profiles. The derived interests were projected as a 3-D co-occurrences matrix where the level of temporal granularity of their interests can be specified in various periods. Furthermore, Plumbaum et al.[118] developed a user-centric system with a personal user interface for aggregating users' profiles from various web applications, and enabling users to manage and share their personal information in a privacy preserving environment. The personal user interface visualises users' personal profiles and allows them to track their personal information in different applications as well as control their data sharing. Church et al. [119] conducted a study investigating the impact of the visualisation type used in a LBS search interface on users' experience of information discovery . They found having both map- and text-based user interfaces is more effective in providing the user with the required information and enhancing their experience. More recently, Cuttone et al.[120]

implemented a personal informatics tool for Android smart phones that offers interactive visual projections of personal information on users' mobility and social interactions. They introduced a visual spiral timeline that displays the mobility patterns across various temporal periods. The social interactions are captured by monitoring the detected devices, where they are grouped in the form of bins depending on their timestamp, and each user is assigned a weight based on their meeting frequency. Initial results from users' activity log files denote that they are able to know more about their behavioural patterns . In addition, Vosecky et al. [10] modelled users' geographical interests shared on Twitter in relation with their corresponding disclosed locations by developing an interactive visualisation system. The system presents the users' geographical topics integrated with recommendations in many visualisations, including user-based by showing an individual's user interest profile and news recommendations, region based by showing topics and users associated with a certain region, and topic based by presenting the terms related with a particular topic.

Generally, visualising location data facilitates perception of the information presented. We used visualisation methods in this work as a means of representing a users' location-based profile, hence enhancing their awareness of their shared data.

2.6.3.2 Visualisation for Privacy Awareness

Visualisation of privacy warnings was found to be effective in increasing user awareness of privacy implications [101]. Studies employed these techniques as a way of providing feedback on users' information accessibility based on the user specified privacy preferences. Anwar and Fong [33] developed a visualisation tool that enables users to explore how their profiles are viewed from the user's perspective in their social connections in order to provide means for understating the privacy implications of the access controls. Participants were asked to perform policy analysis with and without the tool. As a result, the participants were able to perform policy assessment more accurately when using the visualisation tool . In addition, Rode et al. [121] explored visualisation techniques to present system activities and integrate configuration with action as a way of increasing users' understating of the consequences of their action with a system. They implemented a prototype that represents the client's shared workplace as an interactive pie-shaped interface where each slice corresponds to a user's space. An initial user study revealed that the participants were easily able to set permissions and found that the privacy level is intuitive. They were also able to understand the activities broadcasted on the interface. Similarly, Wang et al.[122] implemented an interactive visualisation system that aids users in specifying sharing preferences for their personality traits that are extracted from Twitter. This tool allows users to understand their derived personality traits by showing three types of traits as labelled bars, where each is linked to further sub-traits and coloured distinctively. The filled

bar's length indicates the score for the extracted trait. Evaluation revealed that participants found the visualisation tool useful in configuring their settings.

Several studies have focused on raising users' awareness of collection of their online data by web services. The concept of online interactive privacy was introduced by Kani-Zabihi and Helmhout [123]. This is concerned with any tool or user interface that supports privacy awareness and understanding of online privacy threats by showing the personal information flow. A user study was conducted using a prototype in the form of a mock-up council where participants were asked to perform pre-defined tasks. As a result, online interactive features increase users' privacy awareness and motivate them to learn more about how their personal information is used. Angulo et al. [124] developed a tool that visualises online data disclosure with the aim of supporting usable data transparency. It is basically a user interface that shows an overview of users' disclosed data to various online services and enables them to access their data collected on the service side. The prototype represents a user as a profile picture in the centre of the interface that is connected with the attributes of their disclosed data at the top, and with service providers that data are resealed to at the bottom. Scenario-based usability testing and a workshop session were conducted, which revealed improvement to users' awareness of their data disclosure to web services.

Focussing on GeoSocial environments, Brush, Krumm and Scott [57] studied users' attitudes towards their location privacy when sharing and tracking long-term location information. They collected location traces of 32 participants using GPSs over a period of two months, and then showed them personal visualised maps of their location tracks using five different obfuscation methods. They argued that some of the participants choices of obfuscation methods were not compatible with their privacy concerns, and that this might be due to participants' lack of understanding of the implications of obfuscation. More recently, Tang, Hong and Siewiorek [92] developed three types of isomorphic visualisations: text-, map-, and time-based, that also consider spatial and temporal properties of sharing historical location, such as the physical location and duration, and included four place labels categories (geographical or semantic, general or specific). They applied these three visualisations to their participants' collected GPS data and then showed it to them. The majority of the participants mentioned concerns related to their physical privacy when showing them their visualised location history. Consequently, they all preferred text-based visualisation, since they believed it to reveal the least location information and to be the least appealing to others because it requires more effort to understand.

Visualisation methods have shown to effectively serve privacy-oriented purposes. Yet, utilising and studying the impact of visualisation on location privacy awareness, particularly in GeoSNs, is yet to be investigated. Thus, this work employs visualisation techniques within the proposed location privacy awareness solutions that aim to present the privacy implications of location

disclosure in GeoSN in an attractive and simplified manner, and study their influence on users' degree of awareness.

2.7 Discussion

This chapter discussed location privacy in GeoSNs in terms of its implications and the methods used for protection. It also explored users' privacy concerns and awareness. In addition, feedback and visualisations approaches for privacy awareness are reviewed. Limited work has been conducted to study and address the problem of location privacy in public GeoSNs using privacy awareness methods of location accessibility by other users and content (personal information) that can be explicitly or implicitly revealed based on location disclosure. Table 2.1 presents a comparative summary of related work on location privacy in the area of geo-social applications. GeoSNs can be compared with respect to a number of criteria including:

- Collection methods of location data, whether automatically acquired by the application or manual shared by the user, continuous collection of location or periodic based the application need.

- The use context of location; whether the application is used to communicate with other users based on location information as in LBSNs (location is mandatory), or location information is used as an identifier of users' activities as in LESNs (location is optional).

- Density of the shared location information; whether users are encouraged to share their location using incentives and whether sharing location information in main driver of the application.

- Accuracy of the acquired location information; whether it is pre-defined within the application or there is a need for utilising other resources for extracting it.

- User consent of their location information disclosure; whether it is explicit for each sharing action or implicit within the use of the application.

More detailed discussion on these criteria and how they can impact location privacy in both LBSNs and LENSs is provided in Section 3.2.

A number of gaps were identified in the literature and addressed in this work:

Table 2.1: A comparative summary of work conducted on location privacy in the domain of social-driven location sharing applications.

Related work in geo-social location privacy	Source used for evaluation	Approach	Aim
Tang, Hong and Siewiorek [92]	GPS data	Visualisation of users' location tracks	Studying the impact of visualising different levels of granularity of users' location track on their privacy concern and sharing behaviour
Brush, Krumm and Scott [57]			
Lindqvist et al. [26]	Public GeoSNs	User survey	Studying what motivates users to use Foursquare application
Jin et al. [16]		Analysis of check-in datasets	Studying users' check-in behaviour into residential places from a privacy perspective
Jedrzejczyk [107]	Geo-social location sharing applications specifically developed for their research	Providing location privacy feedback	Studying the effect of providing feedback on who of users' friends viewed their location
Patil and Lai [64]			Studying the impact of providing feedback on who of users' friends viewed their location based on the privacy preferences they specified
Sadeh [38]			
Tsai [37]			
Kelley et al.[65]			
Christin et al. [34]			
Patil [36]			

1) **Implications of location disclosure in GeoSNs on personal privacy.**
Studies have investigated the value and utility of location information on GeoSNs to understand users' behaviour. These studies demonstrate the potential for location-based profiling in terms of the variety and amount of information that can be derived from users' location information. However, they focus on deriving information from a large group of users to discover mobility trends. In this work, we focus on the privacy implications of location sharing for each individual user in terms of what personal information can be inferred and how it affects the user's privacy, with the aim of understanding possible implied user profiles from location data stored in GeoSNs.

Thus, the location privacy implications for individual users need to be considered and explored in terms of feasibility and the range of threats that can be posed. It is also a

necessary to identify and discuss aspects that contribute to the location privacy threat, particularly on GeoSNs, and analyse how privacy risks can occur in relation to location information disclosure on GeoSNs. This gap is linked and addressed by *RQ1* presented in Chapter 1.

2) **Privacy attitude and behaviour in GeoSN domain**
Most research on privacy attitudes and behaviour regarding location information has considered users' reaction towards those who request their location at a given time. In addition, very little effort has been made to understand users' privacy concerns and behaviour on public GeoSNs, hence there is a lack of insight in this area. The specific characteristics of data shared on GeoSNs and the nature of user interaction on these networks suggests the need for more specific studies of location privacy on these networks to evaluate privacy attitudes (users' concerns and awareness of their location data accessibility and inferred information) and behaviour (actions towards data sharing and management) on GeoSNs, using similar research methods (user-based experiments). The support of privacy awareness by GeoSNs deserves to be closely analysed and assessed since it can significantly influence users' privacy attitudes and behaviour. This gap is mapped to *RQ2* and *RQ3*.

The next two gaps are the main limitations found in the literature and highlight the novel contributions of this thesis:

3) **Location Privacy Modelling in the Context of GeoSNs**
Privacy models generally provide a useful foundation for designing effective privacy-aware systems. However, to the best of our knowledge, there is no work that examines location disclosure aspects and it relation to potential privacy threats in the domain of GeoSNs. Such modeling is needed for:

- Proposing location privacy solutions by modelling contents of user location-based profiles derived from their GeoSNs by investigating the type of data dimension involved, and how they can be mapped to possible privacy implications (*RQ4*).
- Identifying main factors that influence users' location-sharing attitudes in the context of these networks using similar approaches (user-based studies with hypothetical location-sharing scenarios) to ultimately model users' perceptions towards location privacy threats (*RQ5*).

4) **Privacy-awareness methods for supporting location privacy in GeoSNs**
Privacy feedback and visualisation approaches have shown to be effective in increasing users' awareness of potential threats to their online privacy. Few studies have examined

the impact of providing location awareness (feedback) on users' sharing behaviour. There are two main limitations in these studies, as can be noted in Table 2.1:

- First, they generally assume that awareness of location information is confined only to the visits a user makes to places, and therefore focus on the accessibility aspect of user awareness. However, these studies neglect providing a holistic awareness of the content that can be disclosed to other uses based on their location sharing actions. In particular, they do not consider what possible personal information inferences that may be made based on others having access to their location data.
- Secondly, their test environment was limited, as proprietary applications with limited features of interaction were used in their evaluation, which does not provide sufficient representation the public (commercial) GeoSNs environment. In GeoSNs, as oppose to these proprietary geo-social applications, the user:
  - Can share with wider audience
  - Can share many type of information, not only location
  - Is provided with incentives to share location
  - Has prior experience with the application
  - Can use it for long-term period

  Hence, these interaction differences can impact these studies outcomes and generalisability on public GeoSNs.

Thus, these limitations need to be addressed by proposing location privacy-awareness methods that consider both aspects of awareness, including the data visibility to other uses and extracted content of users' location profiles, and taking into account their design usability in studying user awareness in public GeoSNs. These methods should also be evaluated using related experimental methods including simulations and prototypes in user-based experiments (*RQ4* and *RQ6*).

Chapter 3

Location Privacy Analysis in GeoSNs

3.1 Introduction

In this chapter, we study the location privacy of users in the domain of GeoSNs. The aim of this study is to investigate the potential privacy implications of location-based information provision and collection. Firstly, factors contributing to the location privacy problem in GeoSNs are examined in terms of data collection, its visibility and accessibility to users of the application, as well as the possible exploitation of these data and the level of security of such services, in order to provide a comprehensive understanding of related privacy issues. Secondly, an analytical study will be carried out using a representative data set, to explore the location data content and the range of possible inferences that can be made from them. Foursquare was chosen as a representative LBSN application for this study due to its popularity. Usage patterns in the dataset are used to guide a classification of users of the application and in the analysis of the data.

3.2 Factors Contributing to the Location Privacy Problem on GeoSNs

In order to understand the potential privacy implications resulting from using location services on GeoSNs, four main aspects of location privacy can be identified. These are related to data collection by the application, its visibility and accessibility, its possible utilisation by potential users, and the level of security offered to the user by the application. This discussion focuses on the type of privacy-related questions that can be asked and the confidence level in the information that can be derived; both factors can affect the degree of privacy concern to users. The study considers both LBSNs (Foursquare[1]) and LESNs (Twitter[2]), the difference in the way location data are acquired and the issues these imply.

[1]Foursquare and its check-in application- Swarm `https://foursquare.com/about`
[2]Twitter `https://twitter.com/about`

3.2.1 Location Data Collection

GeoSNs in general collect any information that has been transmitted through their services. For example, Twitter collects any information which has been published through their services , as stated in their terms of use [3] .

"You understand that through your use of the Services you consent to the collection and use ... [Any information that you provide to Twitter], including the transfer of this information to the United States and/or other countries for storage, processing and use by Twitter."

Foursquare does the same as stated in their terms of use [4] where they declare *"We receive and store any information you enter on our Service or provide to us in any other way."*

Here the type of data, its density and quality, as well as the methods of collection and storage are considered.

3.2.1.1 Method of Collection

This factor is concerned with how the location information of users is collected and under what circumstances. Both LBSNs and LESNs depend on the user's device to acquire the user's current location using GPS, wireless access points (WAP) or cellular networks. When using LBSNs, location data are collected automatically since location is mandatory to providing the service. In Foursquare specifically, a user's location is implicitly acquired on a continuous basis, even when the service is not being used (background tracking). User's check-ins into specific places are verified against their estimated current location and recorded explicitly. In LESNs, user's location data are collected only when location-based features are enabled and used. Some features require continuous collection of location data, for example, when tailoring trends to the user's location in Twitter. The mode of data collection, whether continuous or periodic; automatic or manual, will impact the volume of data collected and its accuracy, and hence also the degree of confidence in inferences made from the data.

In addition, location-based information about users can be collected in two ways, depending on how GeoSNs are used, which can impact the extent of inference and hence pose potential privacy risks, as follows:

1) **Virtual Presence (Passive Location Disclosure)**
   GeoSNs users can perform location-oriented tasks that tag a location without requiring the user to be present in the given location, which can be referred to as virtual presence.

[3] Twitter terms of use, `https://twitter.com/tos` [Accessed: Nov. 2014]
[4] Foursquare terms of use, `https://foursquare.com/legal/terms` [Accessed: Nov. 2014]

For instance, the user can share a comment about a place, rate a place or save it for a potential future visit. In this mode, the feasible privacy implication would be extraction and profiling of the user's interests, preferences and activity in association with location.

2) **Physical Presence (Active Location Disclosure)**
The GeoSN can be used to perform tasks that reveal the user's physical presence in a location where users explicitly tag where they are or the service collects the location of the user implicitly. For example, the user can check into a place in Foursquare and share a geo-tagged tweet in Twitter. Disclosing the user's current location along with other contextual information such as comments and who they are with allows for more inference potentials, such as mobility patterns and prediction of future movements.

3.2.1.2 Types of Data

The completeness and accuracy of location information are primary factors that determine the possible inferences made based on this information and the possible privacy threats to users. Three types of data can be associated with location data collected in GeoSNs: spatial, non-spatial and temporal.

- *Spatial semantics:*
  These refer to any type of information that can be used to identify the places a user has visited. In both LBSNs and LESNs, the user's location is identified as a point in space with a latitude and longitude. In LBSNs, users identify their locations explicitly, allowing for a rich definition of place identity, including place name, type classification and street address. On the other hand, location in LESNs is determined automatically by reverse geo-coding the registered latitude and longitude coordinates, and thus carries a degree of inaccuracy and ambiguity.

  Increasingly, LESNs are able to use resources from LBSNs for defining locations. For instance, Instagram allows users to geotag their pictures using the Foursquare API [5]. After purchasing the application by Facebook, picture geotagging is linked with Facebook Places [6]. Twitter also uses Google API for linking users' selected place names with a location on a map. Hence, in both cases it can be assumed that detailed and precise user location data can be stored by the applications. Knowledge of the spatial semantics reveals details about locations visited by users, including sensitive places such as police stations

[5] Instagram Location Endpoints, `http://instagram.com/developer/endpoints/locations/#` [Accessed: Nov. 2014]

[6] Instagram Location Endpoints, `instagram.com/developer/endpoints/locations/` [Accessed: Jul. 2016]

or mental healthcare centres. Such data can also be used to derive users' place-related interests.

- *Temporal semantics:*
  These represent the time of a user's visit to a place and the duration of their visit. In LBSN, the time of visit is registered by the user as they check-in to a place. The user's physical presence in the place may be validated by comparing their actual GPS coordinates with those of the place they check into. In LESN, a time stamp is encoded with the resource used, for example, a tweet location. However, in this case it is difficult to ascertain whether the user is intentionally visiting the place or is just passing by. In both cases, further processing of the user tracks is needed to estimate the duration of the user's visit. Temporal semantics allow for tracking of users over place and time, which leads to deriving their absence as well as presence in particular places.

- *Non-spatial semantics:*
  Non-spatial semantics are other types of data about both users and places that may be associated with location information, and can be categorised as follows:

  - Contextual information: This includes explicit user data, as, for example, defined in a user's personal profile on an application. It also involves place-related data, such as reviews that express users' experience with place, tags, pictures, as well as ratings where they can explicitly indicate whether they like or dislike a place. With the user's permission, applications will identify users and share their personal information. Rich place-related semantics may also be mined from resources on the web [125]. This information can be exploited to infer users' interests and activities performed in a place.
  - Social ties: Users of GeoSNs are mainly motivated by the social aspect, where they can add friends and follow others. The social connections and what they share can be exploited along with the user's spatiotemporal data to infringe or estimate the user's personal information. It can be utilised to derive where users meet (co-locations), common interests and activities among the user and their friends, as well as the nature of their relationship based on where or how often they meet (e.g. a work colleague or a close friend).

3.2.1.3 Data Volume

The amount of location data collected is another important factor to be considered and is dependent of the user attitude and behaviour when using the application. The pattern of data

logging and the frequency of usage will determine the density of the data collected over time and will thus influence the type of information that may be inferred from the data. Having location data that is dense enough to reflect the users' real-life mobility pattern and consequently enables more reliable yet more privacy-threatening inferences. For example, regular visits to specific places can determine routine mobility patterns. It also can uncover where a user live or work. Mobility patterns can also be exploited for predicting users' future movements or potential transitions from one place to another. However, having less dense locations information can also lead to revealing sensitive location information of users where, for instance, incidental visits to other places can signify special events or activities.

3.2.1.4 Data Extraction

Personal information can be extracted from users‘ location data into two levels based on how it is acquired. The first level is a direct derivation of information from shared location data that can include place coordinates, name and type, as well as users‘ reviews of locations. The techniques used at this level are fairly simple, and can be as straightforward as human-based observations. In the second level, personal information of users is implicitly inferred based on the first level data. This level of data can require more sophisticated inference techniques and algorithms. Further private information can be revealed in the second level, such as mobility patterns and predictions.

3.2.1.5 User Consent

User consent regards the process of asking users for their permission before their information is disclosed. This element affects individuals' awareness of their information retrieval and utilisations as well as relevant privacy implications. User consent can be explicit, where individuals are required to give their permission before their information is stored or published. However, the general notion of consent in web applications is that users are asked to grant permission only once, when they begin using the service, but not every time their information will be shared. Further use of the service and related data sharing is considered consensual, which in this case the consent is implicit. Implicit consent reduces users' awareness of potential privacy risks since they are not given the opportunity to further consider their information disclosure. In addition, the privacy settings of some GeoSNs services are silently updated without notifying the users, and are consented to by default, where users must opt-out in order to deny a certain sharing. For example, Foursquare updated their service to include 'background location', where they track the user's movement even when they are not interacting with the application, and this is consented to by default.

3.2.1.6 Data Accuracy

The accuracy of location data refers to what extent the given information represents the real-life experience of the user. It is mainly related to location and time information, as discussed in the following:

- *Location Accuracy:* Location accuracy considers how closely the collected location represents the actual location of a user. It plays an important role in determining the accuracy of the personal information inferred. Generally, LBSNs provide predefined venues where they have a database of geo-locations and their spatial semantics, including coordinates, name, type and address. Users of these services check-in into nearby venues based on the coordinates acquired from their smart devices. Thus, LBSNs offer a more precise location of users since determining their location is not only based on their GPS coordinates, that can have issues with accuracy, but also other spatial semantics that can reveal exactly what place the user is in. Nevertheless, users can check into places that they are not currently in for other purposes, such as deception or self-representation. As for LESNs, the GPS coordinate is mainly used for acquiring the users' location, which affects the location accuracy. The coordinates can be reverse geo-coded to retrieve more location semantics that can help pin-point the visited place, yet still, the results of this process can have limited accuracy. Some LESNs allow their users to manually enter their location, which can result in vague place information.
- *Time Accuracy:* Whether the recorded time associated with a location reflects the actual time the user is in the given location influences temporal accuracy of the derived personal information, and hence leads to privacy concerns. The time accuracy of LBSNs can be higher than LESNs since their users tend to check into locations at the time of their visit, which is the motive of these services. Whereas in LESNs, users can geo-tag their content even when they are no longer at the location, which reduces the possibility of relating the extracted personal information with the real temporal preferences or the user. Users of LBSNs may also check into venues after leaving them for diverse purposes, including concerned about their privacy [26].

3.2.2 Location Information Accessibility

Location information accessibility represents how much of the user's data are available and visible to others, including the user, other users and third parties. In terms of users' accessibility to their collected location and location-related data, GeoSNs provides only limited means for accessing these kinds of information. In Foursquare, users' previous check-in information is

available in the form of check-in history, where users can view their visited venues, dates of visits and tips they made. These raw data provide only a limited view of the information content in the data. Twitter users can request to download their tweet history, but location information is not included in this data. Although users generally can access what they have shared on the application to some extent, they cannot see what other information has been collected while using the service, such as information about the device used to access the service. For example, Foursquare tracks its users' location and shares it with their friends even if they are not using the application (background tracking), yet users have no means of access this location tracking information. As for information visibility, most of the users' information published on GeoSNs is available to their friends and can be visible to other users also.

Generally, users of GeoSNs have limited control over the visibility and accessibility of their information, since the privacy settings provided to them do not manage all aspects of their information accessibility. In Foursquare and Swarm, almost all of the user's information is publicly available by default and can be viewed by other users. This includes profile information, tips, likes, friends list, photos, badges and mayorships. Check-ins are only shown to users' friends, yet they by default appear publicly in the list of people who are at a venue. Check-ins also appear to other users if a user is the mayor of that place (the user with the most check-ins to a place over a certain period of time). Similarly, in Twitter, users' profiles and their tweets are public by default and can be accessed by others. This means that location information attached to tweets is publicly available also, unless users mark their profile as private, where only followers can view their data. All of the publicly available user information is accessible by third parties, including the GeoSN API's users. Third parties can also access a user's personal and publicly unavailable information. In the case of Foursquare, third parties can obtain check-in data in an anonymous form, but they also indicate that they will share user's personal information with their business partners and whenever is necessary in certain situations, such as enforcement of law. Twitter, on the other hand, states that any content the user submits or displays through the service is available to their third parties without anonymity.

The privacy risks of disclosing location information in GeoSNs actually occur when users' information can be accessed by others, and where an adversary can utilise it for undesirable purposes. An adversary can be an insider from the service provider and their third parties, or other users of the service, including the users' social ties on the application. Users might not personally know all of their friends who can see almost all of their data and exploit it maliciously [59]. Figure 3.1 shows the levels of data accessibility to the parties involved in GeoSNs.

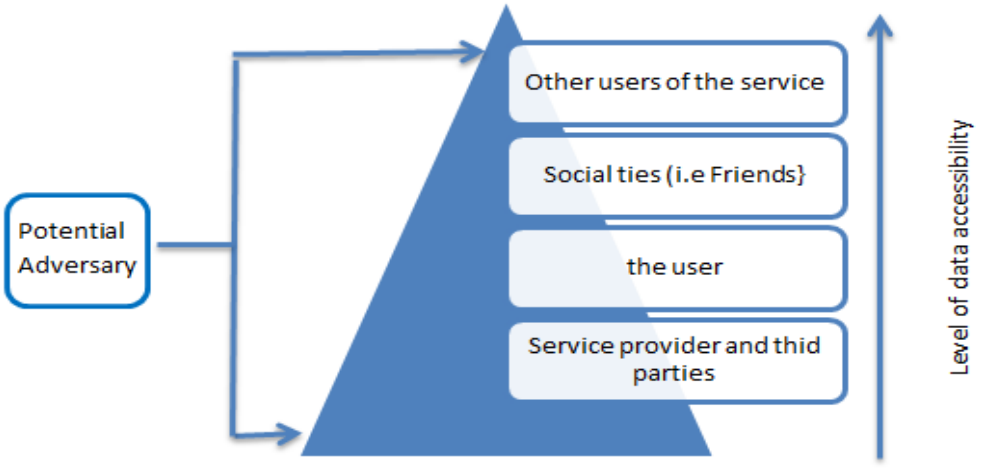


Figure 3.1: Possible levels of accessibility for the parties involved in GeoSNs.

3.2.3 Location Data Exploitation

Location information exploitation refers to how the application or third parties can utilise the data and for what purposes. This dimension involves the actual exploitation of a user's location and location-related data that can lead to various levels of privacy threat. It seems that GeoSNs have unlimited rights to utilise their users' data in any way, and for any purpose, as stated in their terms of use. For example, Foursquare gives itself absolute privileges in using and manipulating user information, as stated in their terms of use [7]

"By submitting User Submissions on the Site or otherwise through the Service, you hereby do and shall grant Foursquare a worldwide, non-exclusive, royalty-free, fully paid, sublicensable and transferable license to use, copy, edit, modify, reproduce, distribute, prepare derivative works of, display, perform, and otherwise fully exploit the User Submissions in connection with the Site, the Service and Foursquare's (and its successors and assigns') business, including without limitation for promoting and redistributing part or all of the Site (and derivative works thereof) or the Service in any media formats and through any media channels (including, without limitation, third party websites and feeds)."

Similarly, Twitter has the right to utilise its users' data, including location information, in various ways, as stated in their terms of use [8]

"By submitting, posting or displaying Content on or through the Services, you grant us a worldwide, non-exclusive, royalty-free license (with the right to sublicense) to use, copy, reproduce, process, adapt, modify, publish, transmit, display and distribute such Content in any and all media or distribution methods."

[7] Foursquare terms of use, `https://foursquare.com/legal/terms` [Accessed: Nov. 2014]
[8] Twitter terms of use, `https://twitter.com/tos` [Accessed: Nov. 2014]

It is therefore clear that there are no restrictions on GeoSNs as to how the data may be used or shared by the application or other parties. In addition, the reasons given for the potential exploitation of users' data are vague (e.g., to improve the services) or even not stated. Hence, by agreeing to the terms and conditions, users are effectively giving away unconditional rights to the use of their data by the application.

3.2.4 Location Data Security

Location data security refers to the level of data protection provided by an application for securing the user's data against loss or unauthorised access. The fact that data are stored on servers opens the doors for potential undeclared access and use, and hence it is almost impossible to guarantee the security of user data. Foursquare declares that the security of users' information is not guaranteed, and any *"unauthorized entry or use, hardware or software failure, and other factors, may compromise the security of user information at any time"*. Without any commitment to responsibility for data security, the application provider is declaring the possible high risk of data abuse by any adversary or even by the application provider themselves. Twitter states that *"Twitter complies with the U.S.-E.U. and U.S.-Swiss Safe Harbor Privacy Principles of notice, choice, onward transfer, security, data integrity, access, and enforcement"*, but gives no additional explanation or examples of situations or access methods that these laws apply to.

3.3 Empirical Investigation

This section describes an investigation of the data collected by GeoSNs. Real datasets have been collected and analysed. The effect of the usage patterns, frequency, and density of the data collected are analysed to understand the implicit semantics/data content between users, places, and other users.

3.3.1 Dataset

The Foursquare dataset used in this analysis is provided by Jin et al. [16] from their work on examining users' residential privacy by analysing their check-ins in a temporal manner. The dataset includes raw location-based data collected using Foursquare APIs. It contains venue information and public check-ins for anonymised users around the wide area of Pittsburgh, USA from 24th February, 2012 to 22nd July, 2012. Places on Foursquare are associated with pre-defined and structured place categories, e.g., Home, Office, Restaurant, etc. The data set contains 60,853 local venues, 45,289 users and 1,276,988 public check-ins of these users.

There are four files used in this data set:

- Checkin.txt containing userID, venueID and timestamp for all users' check ins.
- UserFriends.txt containing all friends IDs of a user.
- Venue.txt containing id, name, latitude, longitude, category information of a venue.
- VenueCategory.txt containing the category information of a venue including the parent category.

3.3.2 Approach and Tools Used

To study the possible impact of location data density on users' privacy, users of the dataset were first classified into groups based on their check-in frequency. A filter was initially imposed to disregard sparse user activity. Hence, users with less than five check-ins per month were removed from the dataset. The remaining users were categorised into three groups based on their daily check-in frequency; moderate, frequent and hyper-active user groups, as shown in Table 3.1. Figure 3.2 demonstrates the classification approach used. One representative user who has the nearest average check-ins per day to the average check-ins for the whole group is selected from each group. Table 3.2 shows some statistics of the selected users. The R statistical package was used for analyses and presentation of results. Mainly, the SQLDF package was used to present the results of the analysis [9].

3.3.3 Results

Analysis of the data set questioned the sort of implicit user-related information that can be considered to be private that may be extracted using the location data collected. User's spatial location history can be extracted in the form of visits to venues and the exact times of such visits. The places visited are identified and described in detail. For example, *user7105* visited 'Kohl's', a department store, located at latitude 40.5111 and longitude -79.9934 at 9 a.m. on Monday 27/2/2012. The basic information on venue check-ins can be analysed further and combined with other semantic information from the user profile to extract further information that can compromise user's privacy. Analysis will investigate the relationship between users and places visited, their mobility patterns and the relationships between users and other users as follows.

[9] R Project, `http://www.r-project.org` [Accessed: August, 2016]

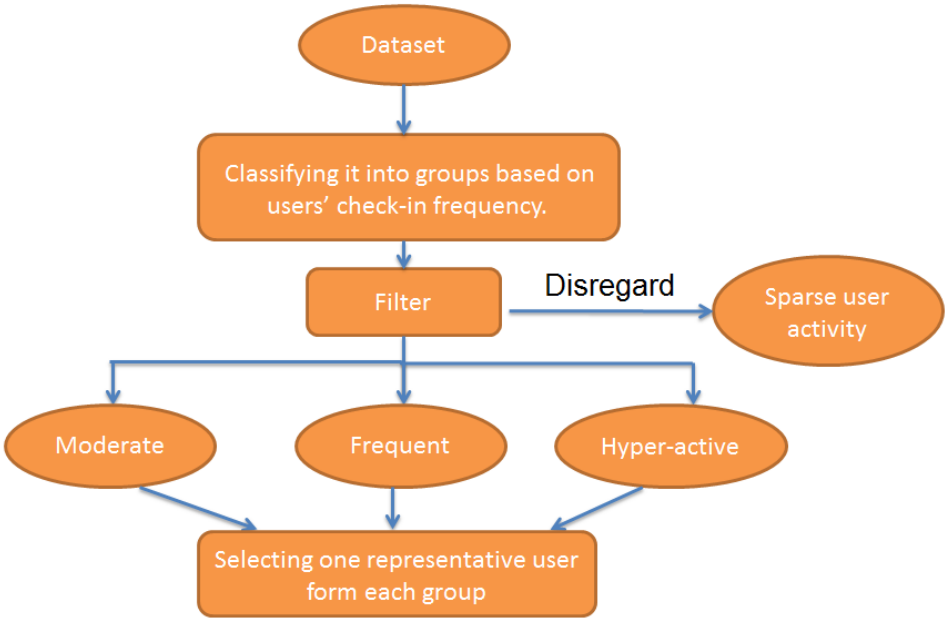


Figure 3.2: Classifying approach used for the experiment..

Table 3.1: Statistics of user groups in the Foursquare dataset..

Group Name	Check-ins Range in Total	Users Count	Check-ins Range per Day	Average Check-ins per Day
Moderate	Between 50 and 300	4902	0.3 to 2	1.15
Frequent	Between 301 and 750	880	2 to 5	3.5
Hyper-active	Between 751 and 1303	24	5 to 8.6	6.8

- *Degree of association between user and place:* Relationship with individual place instances as well as with general place types or categories will be studied. Elements of interest will include visit frequency, and possible commuting habits in terms of the association between the visit frequency of places and their location.

- *Spatiotemporal movement patterns:* Visiting patterns to individual places or to groups of places can identify regular movement patterns. In addition, a change in visit patterns can also be a significant pointer of user activity.

- *Degree of association with other users:* Relationship between users can be derived by studying their movement patterns and analysing their co-occurrence in place and time.

3.3.3.1 The Moderate User

The analysis results of *user9119* selected from the moderate group are as follows.

Table 3.2: Profiles of selected users..

Factor	Selected Users		
	User9119	*User7105*	*User2651*
Number of total check-ins	144	511	1019
Average check-ins per day	0.96	3.4	6.8
Number of visited venues	21	99	101
Number of visited venues' categories	17	47	57
Number of visited venues' main categories	10	11	17
Number of friends	20	10	19

1) **Degree of Association Between User and Place:**
Two venues frequently visited by *user9119* are 'Penn Garrison', whose category is 'Home,' and 'USX Tower' whose category is 'Office', representing 44% and 36%, respectively of the total check-ins. Home and Office are highly sensitive places, yet they represent 80% of this user's check-ins. Other place types visited with significantly less frequency include, 'Nightlife Spot': 0.5%, 'Travel & Transport': 0.27%, and 'Shop & Service': 0.27%. *User9119* is also interested in 'Hockey', 'Garden Center' and 'Museum' place types. As could be predicted, the location of venues visited indicates that most of them are close to 'Home' and 'Office', whereas this user commutes further away to less frequently visited venues such as 'Hockey Arena'. Figure 3.3 shows this user's check-in frequency for different categories of venues classified by the time of day. As can be seen from the figure, this user's association with sensitive places such as home and place of work can be identified. In addition, a strong association with other place categories is also evident.

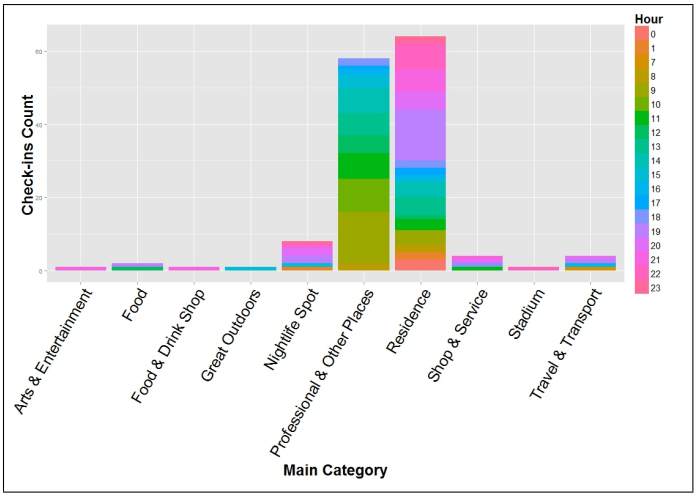


Figure 3.3: The moderate user's check-in count, classified by the category of venues for different hours of the day.

2) **Spatiotemporal Movement Patterns:**
About 40% of this user's total check-ins occurs at 9 am, mostly in the 'Office', and at 7pm,

mostly at 'Home'. More than two-thirds of the check-ins are between 10am and 2pm and between 6pm and 11pm, which indicates that this user commutes more frequently during these hours. From the weekly patterns of movement, it can be seen that 71% of the venues were visited after 6pm. Mondays and Thursdays are when this user is most active, representing 41% of the check-ins. *User9119* tends to go to 'Nightlife Spots' more frequently during working days, whereas visits to other specific place types occur only at weekends, including, 'Salon or Barbershop', 'Coffee Shop' and 'Garden Centre'. This user typically starts commuting earlier on working days and visits more places than on weekends. Observing the check-ins by month shows that the months of May and June are the most active in terms of check-in frequency, comprising 60% of total check-ins, as well as diversity of category of venues visited (99% of the total visited categories of venues occurred in those months, including the emergence of new categories such as 'Museum', 'Airport' and 'Hotel'). The user was least active in April. Figure 3.4 demonstrates this user's check-ins count in different categories of venues, classified by day and grouped by month. Some changes to this user's habits can be noticed as well, which can suggest a change of personal circumstances. For example, the user did not visit any Nightlife Spots in March and April and has not checked-in to any place on Sundays throughout June and July, including 'Home' and 'Office'. In addition, the user did not check-in to any place for a period of a week between the 21st and 28th of April. *User9119's* last check-in before this week was on the 20th of April at 'Home'. This may indicate a possible period of time-off work in that week.

3) **Degree of Association with Other Users:**
Co-location is used here to denote that users have visited the same venue at the same time. This can be used as a measure of interest in a place and relationships between users. *User9119* was co-located in 6 unique venue categories with two (out of twenty) friends. He shared three co-occurrences with two friends; once with *friend1236* at 'American Restaurant' and twice with *friend15229* at 'Office', which may indicate that *friend15229* is a work colleague. In fact, this user shared 95 co-occurrences with 52 other users, 90% of whom were in the 'Office' suggesting the probability of those users being work colleagues.

3.3.3.2 The Frequent User

Analysis of the results of *user7105* from the frequent user group follows.

1) **Degree of Association between User and Place:**
Similar to the moderate user, *user7105's* most checked-in venue category is 'Home',

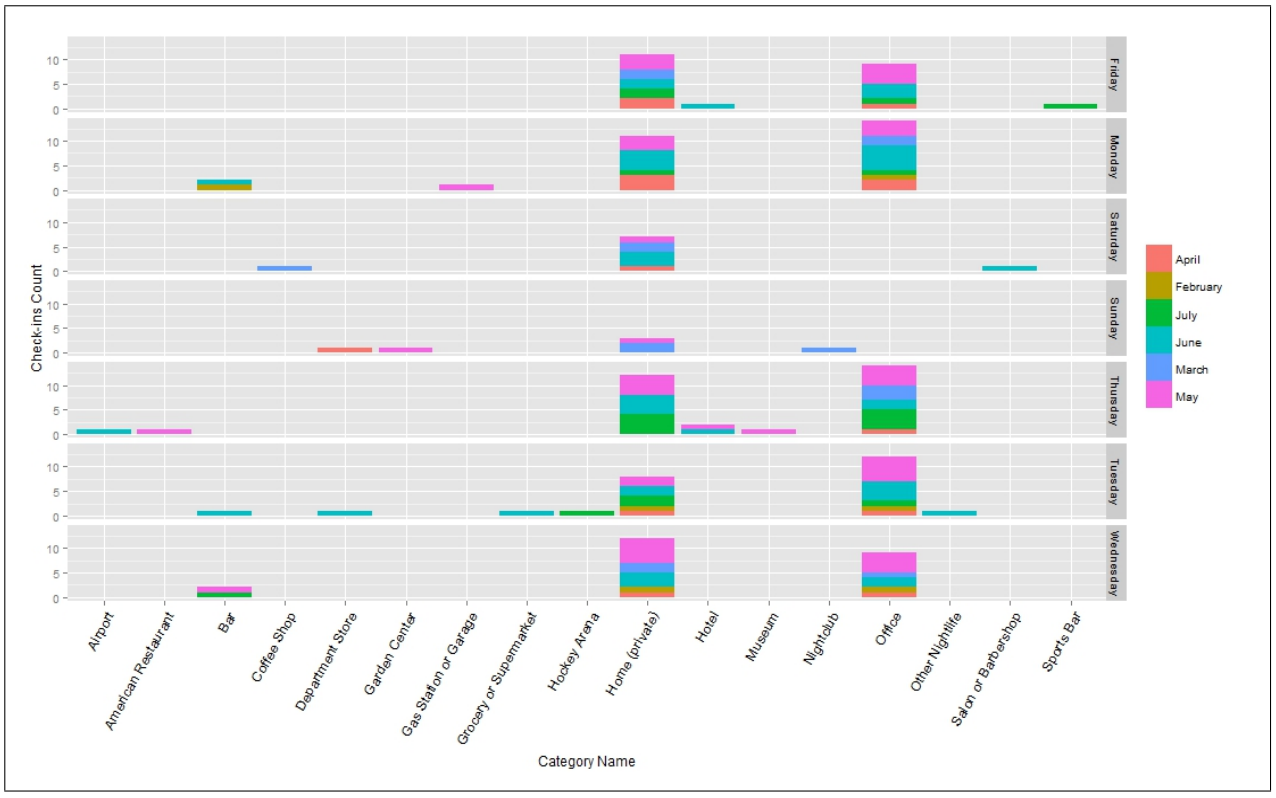


Figure 3.4: The moderate user's check-in count in different categories of venue, classified by day and grouped by month.

whose location is identified in detail. However, the second most visited venue is a specific restaurant, whose category is 'American Restaurant', representing 25% of the total check-ins and 28% of category check-ins. This visit pattern may indicate that this is the user's work place. The third most visited venue category for this user is 'Bar' (4%), which is a subcategory of 'Nightlife Spot', representing about 7% of check-ins. The third most visited main category is 'Shop & Service' corresponding to 10% of check-ins, where specifically 40% of those are to 'Gas Station or Garage' and 25% are to 'Drugstore or Pharmacy'. *User7105* is occasionally interested in visiting places described as 'Great Outdoors', 'Professional & Other Places' and 'Arts & Entertainment'.

The majority of the most frequently visited venues are within close distance to 'Home' and to the 'American Restaurant', whereas *user7105* commutes further away to other less frequently visited places, such as, 'Medical Center'.

2) **Spatiotemporal Movement Patterns:**
About 20% of the check-ins occurs from 10am to 12pm, half of which are at 'Home'. In addition, *user7105* tends to move the most between 3pm and 5pm, representing 23% of his total check-ins to 46% of the visited venue categories. More than half of the check-ins

are at 'Atria's', which may indicate that the user starts his work shift in this place at that time. This hypothesis can be ascertained by examining his subsequent check-ins, where 18% of the check-ins occur between 12am and 3am at 'Home', possibly when the user returns from work. There is a high correlation in terms of place transition between 'Home' and the 'American Restaurant'. When examining weekly mobility, *user7105* is more active on Tuesdays followed by Saturdays, corresponding to 19% and 16% respectively of total check-ins. Noticeably, the majority of Friday and Tuesday check-ins occur at 12am, whereas on Monday and Saturday, the bulk take place at 4pm. Furthermore, this user has visited more diverse venues on Tuesdays, followed by Thursdays and Wednesdays, representing 53%, 43% and 38%, respectively of the total visited categories. During the working week, this user tends to visit a 'Bar' (5%), especially on Tuesdays, and 'Gas Station or Garage' (4%). This is reasonable considering his working shifts. While on weekends, the 'Grocery or Supermarket' and 'Drugstore or Pharmacy' venues are among the top four visited categories corresponding to 4% and 5%, respectively, of weekend check-ins. *User7105's* check-in patterns were regular throughout the entire period. However, visits of this user are more frequent and diversified in the month of March. Noticeably, about 28% of the check-ins between 12am and 3am occurred in March, indicating a possible change of lifestyle. Figure 3.5 presents this user's check-in count in different categories of venue, classified by day and grouped by month.

3) **Degree of Association with Other Users:**
User7105 had co-locations in 36 unique venues from 19 different categories with 7 friends. In particular, 26 co-locations are shared with *freind38466* at 14 venues categories including 'Coffee Shop', 'Bar', 'Fast Food Restaurant' and 'Other Nightlife'. Co-locations shared with the rest of the friends include 'Bar', 'Mexican Restaurant', 'Hospital' and 'Government Building'. Moreover, *user7105* has 16 spatiotemporal co-occurrences at 14 unique venues from 6 different categories with two friends, where 14 co-occurrences with *friend38466* at 6 different categories include mostly 'Bar', 'American Restaurant', and 'Sandwich Place', which can denote a close friendship between them. The other two co-occurrences are with *friend15995* at 'American Restaurant' on May 13th and June 17th, 2012. The place and time of this user's co-occurrences with friends is shown in Figure 3.6. Similarly, this user has 89 co-occurrences with other users, who are not stated as friends, at 29 unique venues, where 38% of these co-occurrences are at 'American Restaurant' and 24% at 'Plaza'.

3.3.3.3 The Hyper-Active User

The results of the analysis of *user2651* selected from the hyper-active user group are as follows.

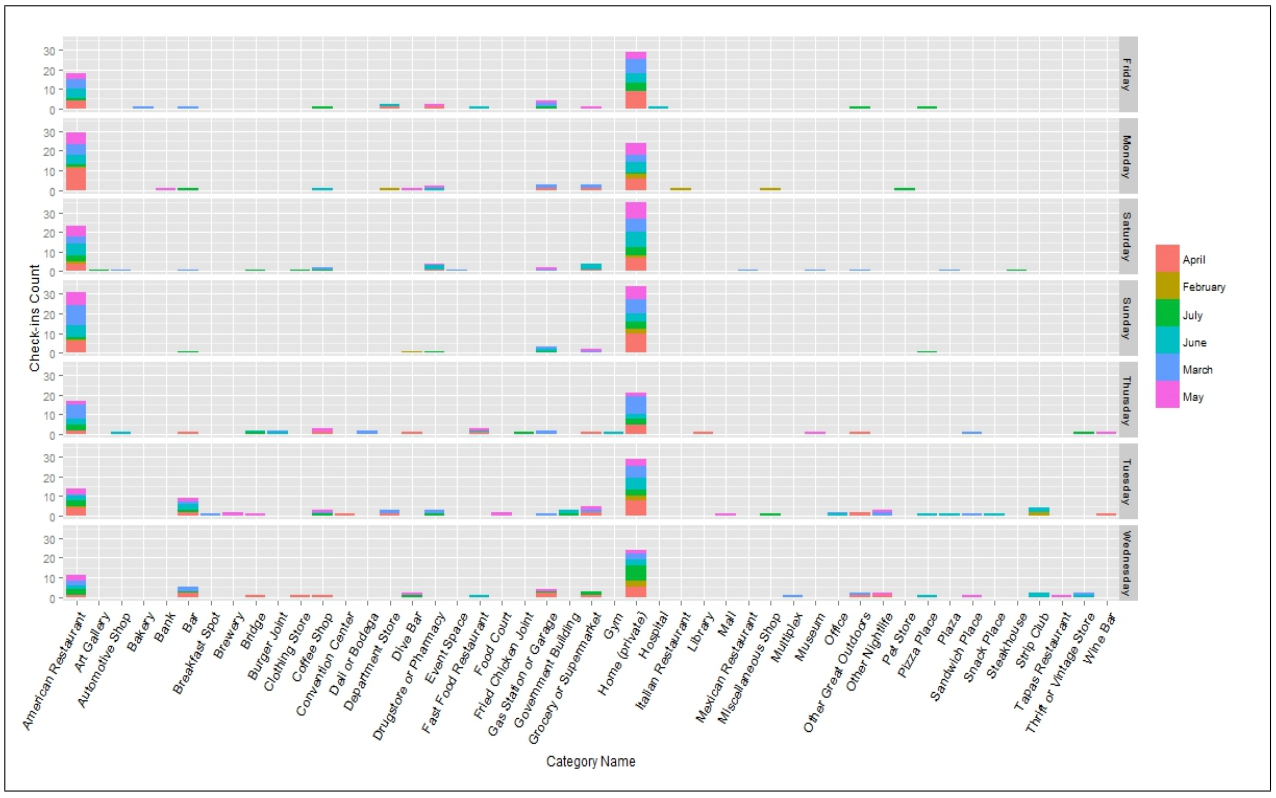


Figure 3.5: The frequent user's check-in count in different categories of venue, classified by day and grouped by month.

1) **Degree of Association Between User and Place:**

The first most visited venue by this user is a 'Nightlife Spot' corresponding to 15% of total check-ins. Two 'Home' venues were recorded, 'My Back Yard' and 'La Couch', representing 23% of check-ins. Both home venues have the same location coordinates, implying that they are actually the same place. 'Automotive Shop', 'Pool' and 'Italian Restaurant', representing 9%, 8% and 5%, respectively, of this user's total check-ins indicate the user's interests and activities; swimming and Italian food in this case. A particular instance with a vague category of 'Building' was among the top 10 most visited venues. Further investigation of this venue using the given place name revealed that an international summit for creative people is held at this place [10], which may indicate that *user2651* is an active participant of such event.

When considering the main category of visited venues, this user visits 'Shop & Service', 'Nightlife Spot', 'Arts & Entertainment' and 'Food' on a regular basis, representing 17%, 14%, 11% and 10% respectively of this user's check-ins. *User2651* also usually visits 'Gas Station or Garage': 4%, and 'Church': 3%. The location of the visited venues can

[10] World Domination Summit, `http://worlddominationsummit.com/faq/#primary-content` [Accessed: Nov. 2014]

be clustered into two main areas on a map as illustrated in Figure 3.7. One area includes 'Home' as well as other frequently visited venues such as 'Nightlife Spots' and 'Gym or Fitness Centre'. The other area includes less frequently visited venues such as 'Hospital'.

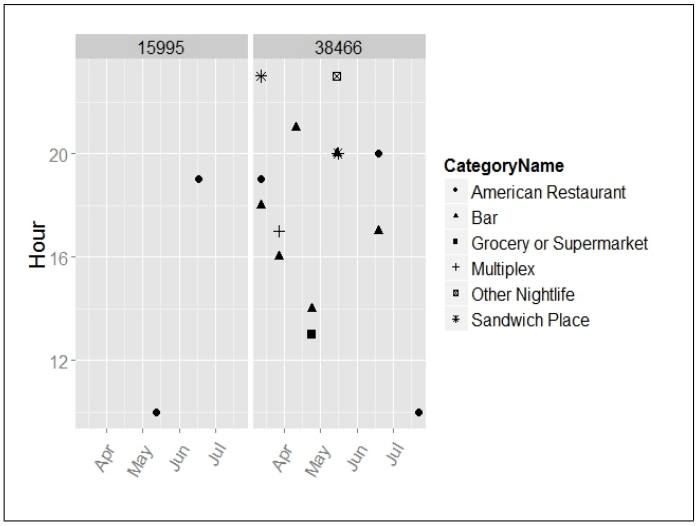


Figure 3.6: Spatiotemporal tracks of the frequent user co-occurrences with friends.

2) **Spatiotemporal Movement Patterns:**
Overall, 53% of residential check-ins occur between 9am and 12pm. A significant number of check-ins (10%) occur at 2pm, of which almost two-thirds occur in an 'Automotive Shop'. Check-in frequency reaches another peak between 11pm and 12am (18%), of which more than half are in 'Nightlife Spot'. Noticeably, this user tends to be more active at night, where about 70% of the check-ins are registered after 6pm. In his case, weekends have similar check-in frequency as the working week, but Sundays register as the most active day in terms of check-in frequency. Moreover, *user2651* checks in considerably less frequently at the 'Automotive Shop' and the 'Pool' on Wednesdays and Fridays, but checks-in to the 'Automotive Shop' and 'Nightlife Spot' on weekends. This may indicate that he works shifts on weekends.

User2651 has regular check-in patterns over the whole period. However, in the months of June and July, check-ins into 'Hotel' and 'Pool' significantly increased, representing 75% and 60%, respectively, of these venues' total check-ins. Figure 3.8 demonstrates this user's check-in count in different categories of venues, classified by day and grouped by month.

3) **Degree of Association with Other Users:**
As with other users, *user2651* was co-located with 23 users at 12 distinct venues. Half of these co-occurrences took place in 'Bar', 'Automotive Shop' and 'Grocery or Supermarket'. *User2651* is co-located in 27 unique venues from 19 categories with 9 friends, 13

of which are with *friend12432*, and 9 with *friend12046*. Most of the co-locations are in 'Nightlife Spots', 'Gas Station or Garage', 'Pool', 'Flower Shop' and 'Bar'.

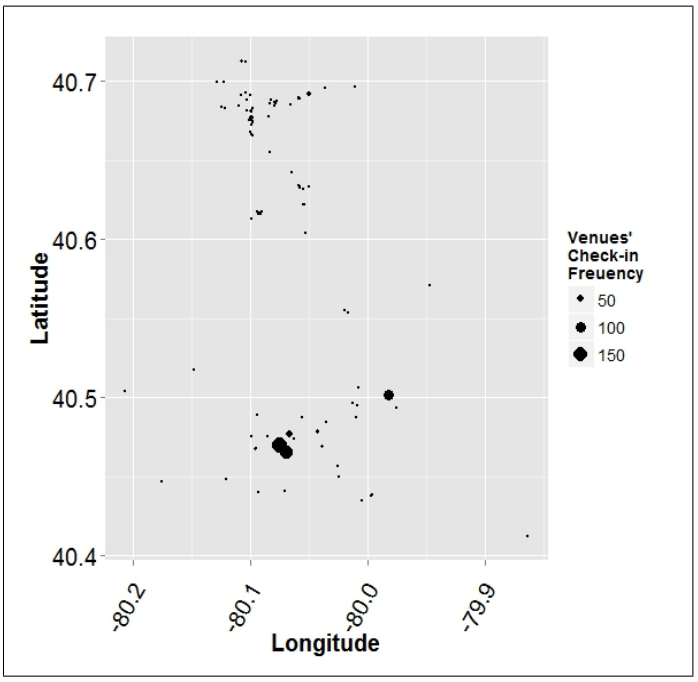


Figure 3.7: Coordinates of venues visited by the hyper-active user, considering the frequency of visit.

3.4 Discussion

The proliferation of location-based GeoSNs and their large-scale uptake by users suggest the urgency and importance of studying the privacy implications of personal information collected by these networks. Identifying user profiles is a goal of many businesses that is now commonly accepted as being for the purpose of improving the quality of service. However, GeoSNs do not explicitly present similar business goals, and thus their motivations for collecting and sharing personal location information are not clear. Also, the issue is further complicated by the fact that the data collected may be shared or accessed by other users and applications. The results of this study highlight the possible implications to user privacy and the need to develop means for increasing user awareness of these issues, and possibly also giving the user the ability to manage access to their data.

The data analysis experiment conducted here shows the amount and types of personal information that can be inferred using location data. Users' spatiotemporal mobility tracks can be analysed to identify where they are, where they are likely to be, and sometimes, more significantly, where they are not. Tracking user location data may also give indications of their preferred

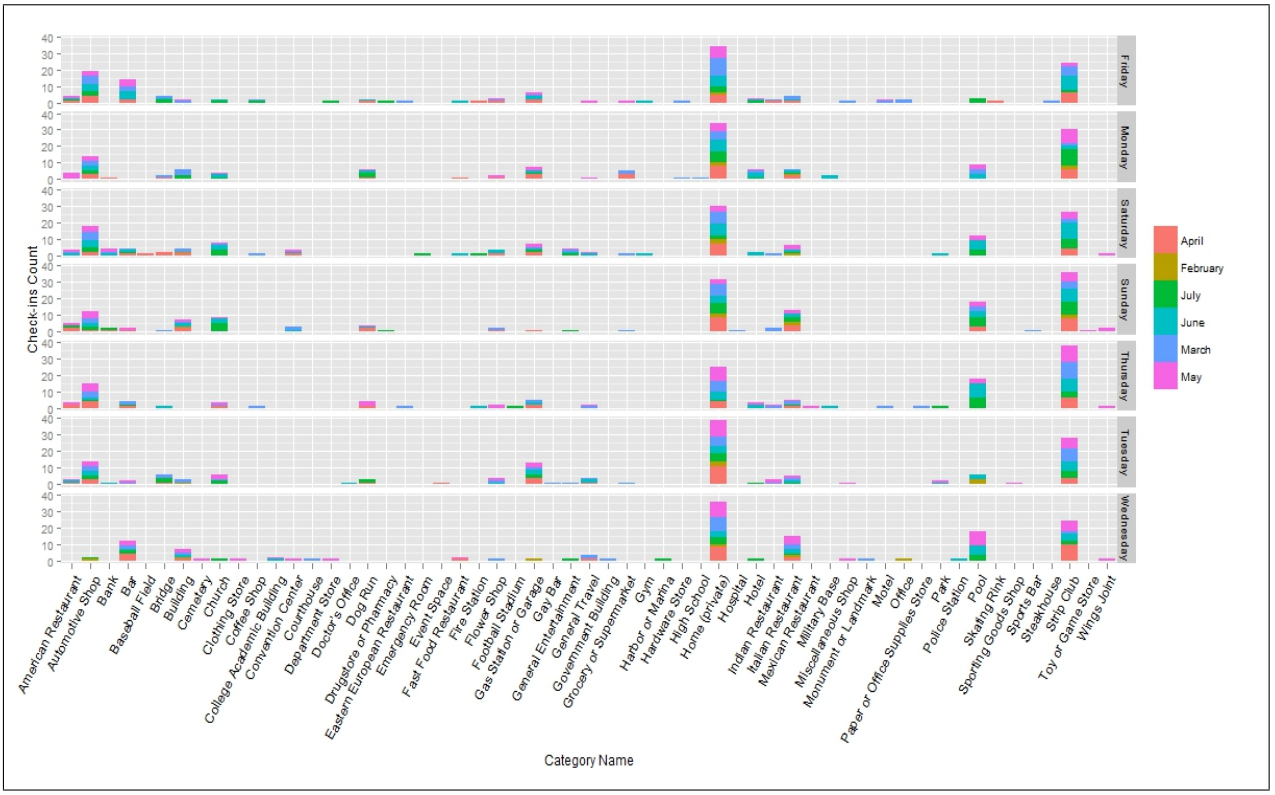


Figure 3.8: Count of check-ins for the hyper-active user in different categories of venues, classified by day and grouped by month.

activities, places, habits and friendship community. In addition to these inferences that have been identified in the literature (see Section 2.3, this study also highlighted the potential reveal of users' commuting habits, nature of relationship with friends or other users of the service, any change in users' mobility routines, work places, and absence from home. Please note that this analysis focuses on deriving users' location-based profile to understand how it can impact their privacy, whereas other related studies aimed to investigated to general mobility trend from large number of users in certain areas. The driver of this study is to explore what personal information can be inferred from sharing location data in relation to the spatial, temporal and social aspects. There is no a particular set of inferences that were specifically targeted, but just to see the type and density of the information that can be derived. Sensitive locations are subjective to users. Although users can have different opinions about the sensitivity of the work place for example, most of them agree about the sensitivity of other places such as home or medical centres (see Section 6.2.1.1 and 6.3.1.1). Thus, a privacy-aware system can always warn users about disclosing location that are considered sensitive by the majority of users, and then adapts the sensitivity of location based on each user's sense of privacy.

As can be expected, the more frequently the applications are used, the denser the spatiotemporal history of user data collected and the greater certainty in the information extracted from this

data. At the same time, sparsity of data resulting from less use of the application can impact the density and the certainty of the derived information. This sparsity can occur in the spatial aspect where a user visits limited number of places, in the temporal aspect where a user visits places in certain times only, or in both aspects. Consequently, it might be not possible in such situations to infer users' patterns or extract wider range of their interests and activities which in turn enhances their privacy. however, having sparse data do no means that there are no privacy implications of disclosing location. For example, a user can check-in to her home location which is considered a sensitive place for most users. Thus, users of GeoSNs need to be informed about their data collection and potential privacy implications of their interaction with the application and allow them to decide upon what of their sharing actions can cause risk based on their own privacy preferences. Whilst the statistical analysis carried out in this study highlights some of the basic interesting inferences that can be made, more sophisticated location-based inference methods can be developed to infer, for example, the probability of future movements, methods of transport and places visited. The now common practice of linking user accounts in several GeoSNs increases the availability of data and compounds the privacy risks to users, who sign up to different, possibly contradicting, terms of use and policies of different applications. For example, developers now use the Twitter API to collect user check-ins in Foursquare.

It is not clear, from the dataset source, to what extent the users in the dataset used have utilised of the provided privacy settings. However, these settings still offer fairly limited control on the data accessibility by other users as discussed in Section 3.2.2. The user's data are generally accessible by at least friends on the application whom some of them might be personally known by the user and yet are able to infer all possible information about this user. Some of users' data are set by the application to be publicly available to other users without users' control even if they used the strictest privacy settings. For instance, a user' presence in a certain place are revealed to other users that are also in that place. In addition, if a user is the most frequented visitor to a particular place, or wrote a review about their visit to a place, then this user presence is also reveal to all other users. The only aspect that can affected by making use of these settings is the density of the extracted information and the possibility of revealing pattern of a user mobility to other users of the application. Nevertheless, users in general do not use the provided privacy settings or have difficulty and misconception about how they can be utilised (see section 7.6.1) mainly because they are unaware of the potential risk to their privacy that would motivate them to make use of the available settings. Thus, such limited utilisation of the privacy settings would even increase the availability of users' data to the public.

Although this analysis (and most of this work evaluations) is conducted on LBSNs, LESNs are also considered in all of the work stages including understanding the location privacy problem and proposing privacy awareness designs. LBSNs provide more location-oriented services that include those offered by LESNs and hence choosing LBSNs for evaluations can yield general-

isable outcomes that cover both types of GeoSNs. The only issue is that the retrieved location semantics (e.g. place name and type) in LESN can be less accurate since reverse geo-coding is used for acquiring them as oppose to the LBSNs where locations and their semantics are pre-defined within the application as discussed in Section 3.2.1.6.

3.5 Conclusion

This chapter has identified four main aspects of location privacy in terms of data collection, accessibility, utilisation and security in order to understand the potential privacy implications resulting from the use of location services on GeoSNs. The privacy implications of location-based information provision and collection in GeoSNs were also investigated. The study is supported by analysis of a representative dataset from Foursquare. The results show that it is highly feasible to infer rich personal information about users and their mobility. In particular, some of the possible inferences demonstrated are:

- Users' spatiotemporal mobility tracks.
- Visiting frequencies and possible degree of association with specific places or place types.
- Users' spatiotemporal movement patterns.
- Users' absence and presence in particular places.
- Users' commuting habits.
- Co-location patterns with other users and friends.

The study also demonstrates the need for users to improve their visibility and accessibility of their collected and derived information by GeoSNs in order to allow them to better assess the privacy implications of their location sharing activities on these networks.

Chapter 4

Privacy Awareness, Concerns and Attitude in GeoSNs

4.1 Introduction

As demonstrated in Chapter 3, there is a substantial feasibility of inferring users' personal information that may pose a threat to their privacy on these networks. Other studies also utilised publicly available information from GeoSNs in order to derive or predict users' location [70, 71, 72, 77], infer relationships between users of GeoSNs [78, 72, 12], study and extract spatiotemporal movements and activity patterns of users on GeoSNs [45, 9, 8], and more recently, reveal personal details of users including their gender, educational background, age and sexual orientation [6, 84], as presented in Chapter 2. However, few studies have explored users' concerns specifically in regards to their residential privacy when sharing their location in GeoSNs [26, 16]. Hence, there is a need for a more elaborate study on users' privacy attitude and behaviour on GeoSNs that takes into consideration the nature of user interactions and data disclosed.

This chapter assesses the need to improve the visibility of user profile information on GeoSNs. An online survey was undertaken to gauge users' understanding and reaction to privacy implications that result from disclosing their location data and their need to control access to their data on GeoSNs. Three main aspects are addressed in this study, including users' awareness of practices carried out on their shared data by these services, their reaction towards possible personal information inferences, and their preferences in terms of controlling access to their personal information on these applications. The outcomes can assess the need for further work on improving the visibility of the information extracted, to allow users to better understand the implications of their location sharing activities.

4.2 Experimental Design

This user study is designed as an online questionnaire in order to be able to capture responses from a large and diverse sample of users for the purpose of yielding representative feedback. This survey was carried out using Google Forms, an online survey service, to facilitate dissemination of the survey as well as collect participants' responses, since it is a fixable and powerful online survey tool that has many options for deploying the questions and capturing users' inputs. The targeted participants of this study are any users of online social networks who use location features such as adding location to their posts and pictures or checking into places. This study targets any user of social networks from any background (no restrictions on their demographics such as age and gender); this is to prevent bias in the study results and to achieve generalisable outcomes.

Before disseminating the survey, a pilot study was conducted to ensure the clarity and coherence of the survey question as well as to check for any possible difficulties in answering the questions or the flow of the study. Volunteers from inside and outside the Cardiff School of Computer Science and Informatics with no specific background took part in this study, where they completed the survey and were then interviewed about it. They provided valuable feedback into the wordings and layout of the questions used. The survey was improved according to the pilot study input.

Finally, the survey was disseminated widely to staff and students within Cardiff University as well as other UK universities. It was also advertised to overseas connections on social networks through the author's account. A token incentive of $\pounds 10$ Amazon voucher was offered to ten randomly chosen participants who completed the survey.

The questionnaire consists of four main sections. The first section collects background information on the participants including their demographics and their use of GeoSNs, particularly Facebook, Twitter Flickr, Instagram, Foursquare, and Yelp since they are the most popular. For example, Facebook has more than two billion monthly active users [1]) and Foursquare has over 60 million users across the world [2]). This section gathers application use and frequency of GeoSNs, as well as other social network features. The second section examines users' knowledge of the terms of use and privacy policies of such networks, followed by a section studying their perception of possible inferences of personal information. The final section aims to capture users' attitudes to privacy on social networks as well as their attitude to controlling their personal information.

[1]Statista, Number of Facebook users worldwide 2008-2017https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/[Accessed: 25-Mar-2018]

[2]Foursquare, About Us https://foursquare.com/about[Accessed: 27-Jul-2016]

The participants are presented with an introduction for the survey explaining its purpose. The extended sections include a brief description of the task to ensure the integrity of results. The questionnaire is divided into many pages, with each containing a few lines on the nature of the question in order to simplify it and reduce the possibility of users being influenced by other related questions. This survey is designed to be completed by participants in no longer than ten minutes. The survey is provided in Appendix A.

The questionnaire data was analysed using R statistical and programming language, since it is powerful and flexible in analysing and manipulating data as well as generating effective graphs. The results are grouped and presented in their corresponding section. Participants' background information collected form the first section, such as age, gender and frequency of use, are used as factors to investigate whether they influence on user awareness and attitude explored in the rest of the survey's sections. Chi-square test of independence was used to examine the significance of these factors' impact on participants' inputs.

4.3 Analysis of Results

186 users of online social networks completed the survey. In what follows, the results of the different sections of the questionnaire are analysed.

4.3.1 Background and Profiles of Location Information Use on Social Networks

Most of participants are young adult users, as shown in Figure1. 60% are in the age group of 15-24, representing the top age group of the participants, followed by the age groups of 25-34, and 35-44, representing 27% and 8%, respectively. There is no gender bias, as the number of male and female users is almost equal. In terms of the frequency of using social networking applications, the vast majority use these services frequently (several times a day), representing 77% of users, followed by moderate use (several times a week), and occasional use (once a week or less), corresponding to 13% and 9%, respectively.

Regarding location services use, 72% of the participants use the location services in the Geo-SNs. Moreover, it seems that the participants prefer to use a few number of location services as illustrated in Figure 4.1(a), whereas approximately 60% use only one location service, followed by two services, and three services, representing 20% and 12%, respectively. As for the geo-social services used, the top location service is adding locations to posts and pictures on Facebook corresponding to 47% of the total number of location services used followed by

adding location to tweets on Twitter, photo mapping pictures on Instagram, and checking in on Foursquare representing 17% ,16% and 10%, respectively, as illustrated in Figure 4.1(b).

(a) (b)

Figure 4.1: (a) Percentage of the number of location services used by participants. (b) Percentage of the type of location services used by participants.

In addition, when observing the frequency of use of the location services by participants, it is clear that almost all geo-social application are used sometimes with considerable percentage of 'always' use which is no less than 21% for each of the location services presented as demonstrated in Figure 4.2. Nevertheless, Foursquare is the only application where it users use it 'always' more than 'sometime's, which generates even richer location-based profiles for users, given that Foursquare provides place semantics as well as social information of users. Furthermore, 25% of the participants have linked their social network accounts, whereas the rest have not.

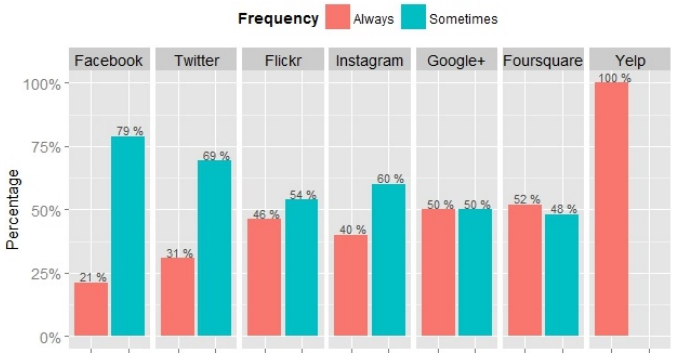


Figure 4.2: Users' awareness of general terms and policies of GeoSNs (Term1-Term4) grouped by frequency of use.

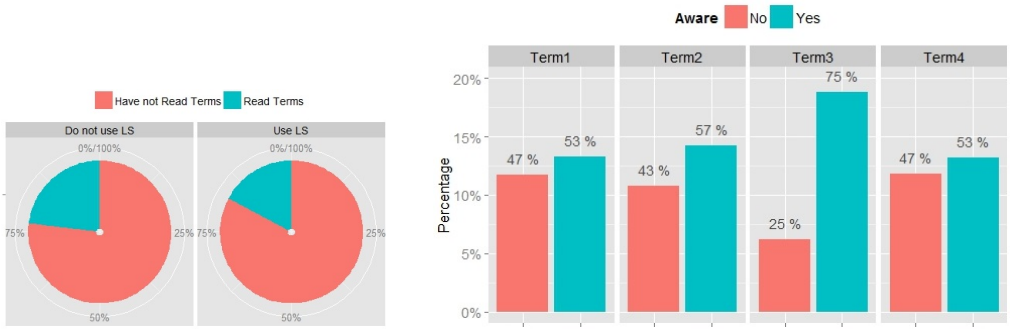
4.3.2 Knowledge of Terms of Use and Privacy Policies for Social Networking Applications

Here, the awareness of the terms of use and privacy policies are examined and analysed against users' profiles. A general question was used to begin with to understand the extent of users' knowledge of privacy policies. The majority of the users (81%) have not read the terms of use or privacy policies of the social networking applications they use. This percentage increases to 83% when considering only participants who use location services in social networks as demonstrated in Figure 4.3(a), which suggests that they are less aware of privacy risks associated with information sharing on these services. Users were presented with the following typical statements representing the terms of use relating to location information, and were asked to indicate whether they are aware of the information in the statements. Note that the following statements are representative of the terms of use of all the GeoSNs in question. The results are shown in Figure 4.3(b).

- *Term 1: The application collects and stores your precise location (as a place name and/or a GPS point), even if you mark your location as private, for a possibly indefinite amount to time.*
- *Term 2: The application can use your location information in any way possible including sharing it with other applications or partners for various purposes (commercial or non-commercial).*
- *Term 3: If you share your location information, your friends and any other users are able to access and use it in any way possible.*
- *Term 4: The application can collect other personal information, such as your personal profile information and browsing history from other web applications.*

More than half (59.5%) of the users acknowledged awareness of all of the statements and of those 23% have read the terms and policies, yet a considerable portion of them are unaware of them (up to 47%). Most users (75%) are aware of Statement 3, relating to the sharing of information with friends, but are generally unaware of statements 1 and 4, relating to how their location and other information may be collected and stored by the applications.

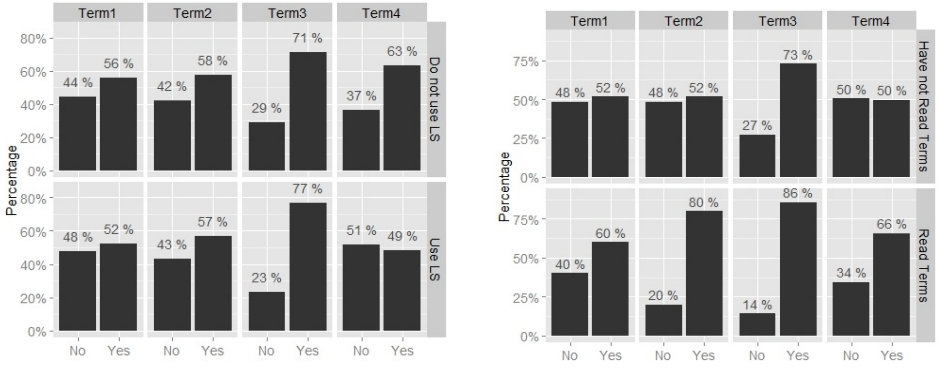
When exploring how users' profiles or factors can affect their awareness of the privacy-provoking terms of use or privacy policies of social networks, it seems that users' awareness level varies to some extent based on these aspects. In terms of using location services, the participants who use



(a) (b)

Figure 4.3: (a) Participants' knowledge about the application terms and policies grouped by location services use. (b) Participants' awareness of the general terms and policies of GeoSNs (Term1-Term4).

geo-social services are generally less aware of these potential terms than the participants who do not (by 3%). In particular, users of location services are less knowledgeable, specifically of Term 4, of which the majority are unaware (51%), and Term 1, about the collection of location and other personal information (48%), as shown in Figure 4.4(a). As expected, participants who reported that they read the terms of use are more aware of the terms presented (73%) than those who do not (56%). About half of the participants who do not read the terms are particularly unaware of terms related to the collection and use of personal location information (Terms 1, 2 and 4), as presented in Figure 4.4(b).

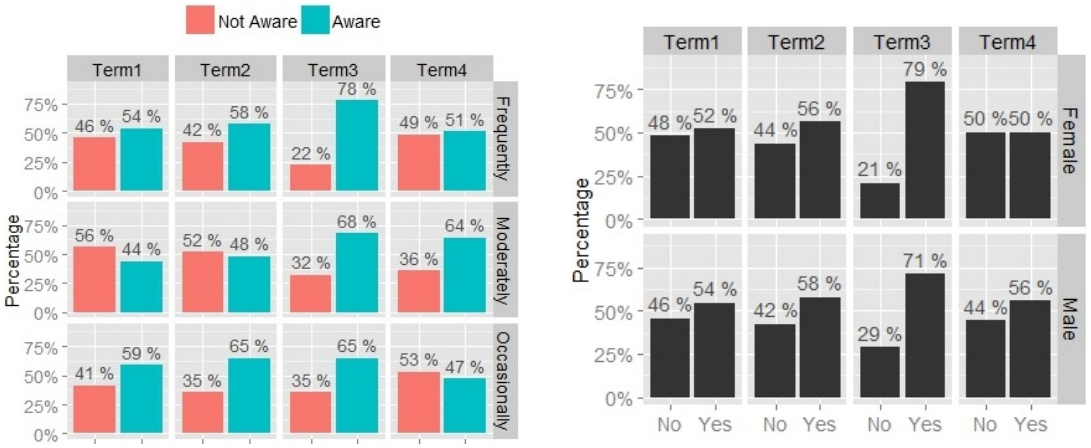


(a) (b)

Figure 4.4: Participants' awareness of terms and policies grouped by each term by considering (a)location service use, and (b)reading them.

Overall, frequent users of social networks have the highest awareness rate followed by occasional and moderate users representing 60%, 59% and 56%. Nevertheless, it is interesting to note that frequent users of such applications are generally unaware of such statements (up to 49%) and the least frequent users are more knowledgeable of Terms 1 and 2 about the collection and use of location information as demonstrated in Figure 4.5(a). In addition, gender

does not seem to have an impacting factor on participants' knowledge of terms (awareness in male:60% and female:59%), but when observing responses to the individual terms, male participants are least aware of Term1: collection of location data (50%), whereas females are least aware of Term4: collection of other personal information (46%), as shown in Figure 4.5(b).



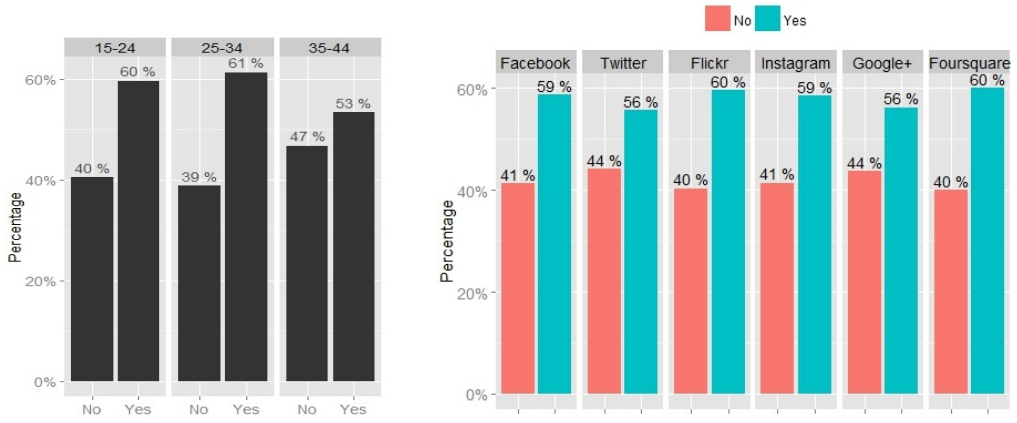
(a) (b)

Figure 4.5: Participants' awareness of terms and policies grouped by each term by considering (a)the gender, and (b) frequency of use.

Younger users aged between 15 and 34 tend to be more knowledgeable of these policies (60%) (the age groups that has less than five participants was discarded), as illustrated in Figure 4.6(a). When considering the location services used within the GeoSNs, it seems that Flickr and Foursquare users are more aware of the terms and policies than users of other applications, representing 60% of their users, whereas Twitter and Google+ users have the highest level of unawareness representing 44% of their users total number, as presented in Figure 4.6(b). However, the difference in awareness level between different location services users is not significant. The factor of whether the participants' link their social application accounts has minor impact on their terms awareness where those who linked their accounts showed 7% increase in knowledge of policies.

4.3.3 Perceptions of Possible Privacy Implications

In this section, users' attitude towards the possible inference of personal information resulting from location sharing on GeoSNs is examined. In particular, the questions aim to gauge users' awareness of plausible inferences about their private places, activities at different times, their connections to other users, and possible knowledge of this information by the application. Participants were presented with 14 statements, shown below. They were then asked to indicate, for each statement, whether they are aware that the statement is possible and to score their reaction to the possibility of this statement as either 'OK', 'Uncomfortable' or 'Very Worried'. The



(a) (b)

Figure 4.6: Participants' general awareness of terms and policies by considering (a)the representative age groups, and (b)location services used.

first twelve statements refer to information that can be derived from the user's location history, which is commonly accessible to 'friends' in most application, if the account is not visible to the public. The last two statements reflect commonly used terms of use, such as giving the application the right to share the user's data with other users and third parties. The dimension of the data queried is indicated against the statement number (Spatial (Sp), Temporal (T), Social (Sc)).

These 14 set of statements were chosen to cover different aspects of information inferences can be made in relation to location disclosure on GeoSNs and to capture users' attitude towards them. Generally, these information extractions can be affected by how much users share of their location data including the type of places they visit and their frequency of visit. Most of them can be known using small amount of users' data such as S1, S3, S6 and S7, while the remaining need more dense data to be derive such as S4 and S5. If the privacy settings are set properly, the first 12 statements are true and extractable at least to friends, while some of them would also be visible to public such as S6 and S10. Nevertheless, these 12 inferences are also accessible to public if the provided setting are not utilised. The last two statements (S13 and S14) are applicable whether the privacy settings are used or not since they are concerned with information visibility by the application which has all of users' data. Moreover, the presented statements cover the potential information inferences that can be made using both LBSNs and LESNs because all of them can be derived basically by knowing the location and time of visit (timestamp) which are disclosed in both types of GeoSNs. The only difference is that they would be generally more plausible in LBSNs since all of users' shared data revolves around location. Therefore, these set of information can be extracted from using any of the GeoSNs that participant were asked about their use of earlier in Section 4.3.1.The results are summarised in Figure 4.7.

- *S1-Sp: I can guess where your home is.*
- *S2-Sp: I can guess where your work place is.*
- *S3-SpT: I know which places you visit and at what times.*
- *S4-SpScT: I can tell where you normally go and what you do on your weekends.*
- *S5-SpScT: I can tell you where you go for lunch or what you do after work.*
- *S6-Sp: I know your favourite store (your favourite restaurant, your favourite coffee shop, etc.)*
- *S7-SpSc: I can guess what you do when you are in a specific place.*
- *S8-SpT: I can guess when you are AWAY from home.*
- *S9-SpScT: I can guess when you are OFF work.*
- *S10-Sc: I know who your friends are.*
- *S11-SpScT: I know when and where you meet up with your friends.*
- *S12-Sc: I can guess which of your friends you see most.*
- *S13-SpT: Other people can know where you are at any point in time.*
- *S14-SpScT: Other people can know what you are doing at any point in time.*

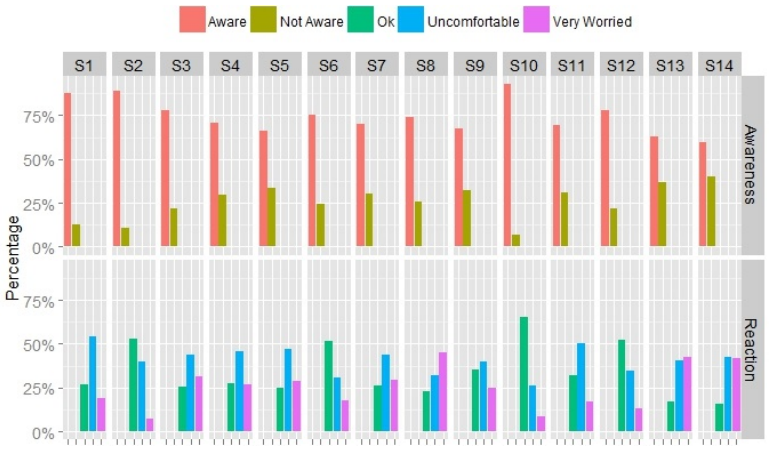


Figure 4.7: Participants' awareness and reaction about potential information inferences grouped by inference statement.

In terms of awareness, it is noticed that most participants are aware of these potential information inferences (74%). In particular, users seem to be most aware of statements S10, S2 and

S1 regarding knowing their friends and estimating work and home locations which correspond to 93%, 89% and 88% of each inference awareness respectively. On the other hand, , the participants are least aware of S14, S13 and S5 that relate to other users' knowledge of personal mobility patterns and activities, representing 40%, 37% and 34% of unawareness level of each inference respectively.

Despite a reasonable level of awareness about the plausibility of these statement, users seemed to be relatively concerned about their privacy. 66% of users' reactions were either 'Uncomfortable' (41%) or 'Very Worried' (25%). Over half of the responses to S2 (awareness of workplace - 53%) and S10 (awareness of friends - 65%) were not concerned. On the other hand, participants were most concerned with S13 and S14, with the 'Very Worried' category scoring 83% and 84%, respectively. S1 and S11, relating to the location of home and meetings with friends were rated most 'Uncomfortable', corresponding to 53% and 51%, respectively. Statement S8, suggesting the knowledge of user's absence from home and S13, indicating the possible knowledge of this information by other people presented a significant source of worry to users, with 45% and 42%, respectively indicating that they are 'Very Worried' about these statements.

4.3.3.1 Influencing Factors on Inference Awareness

This section examines what aspects of users or their social network usage effects their awareness of the previously mentioned personal information inferences. Non-users of location services are more aware of these inferences than users of location services by 6% in general, where this difference in awareness increases up to 20% when clustering awareness for each statement, as demonstrated in Figure 4.8. Participants who read terms and policies are 9% more aware than users with no knowledge about them, as illustrated in Figure 4.9(a). Indeed, reading terms and policies was found to significantly impact users' awareness of potential inferences (Pearson Chi-Square = 16.637, $p < .00001$). Furthermore, gender has less impact on inference awareness, where male users tend be 5% more aware of the potential information extraction than females, as shown in Figure 4.9(b). These two general observations regarding users' knowledge of applications' terms and their gender are still applicable even when clustering the awareness for each of the inference statements.

When considering the frequency of use aspect, there is no significant different in awareness among frequency-based user groups (3% increase in awareness for the 'moderate' users). Some variation in awareness level appears when classifying these three user groups based on the individual inference statement, where moderate users seem to rate the highest knowledge of S5 and S4 related to deriving the normal mobility pattern, representing 17% and 11% more awareness, respectively, than the other groups, as presented in Figure 4.10. 'Frequent' users are the most aware of S2 and S10 regarding knowing workplace and friends, representing, 9% and

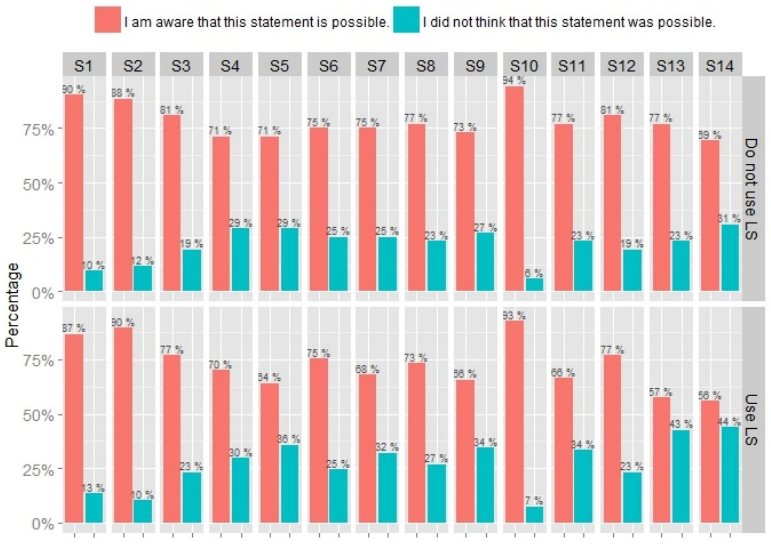


Figure 4.8: Participants' awareness of potential inferences by considering location services use grouped by inference statement.

(a) (b)

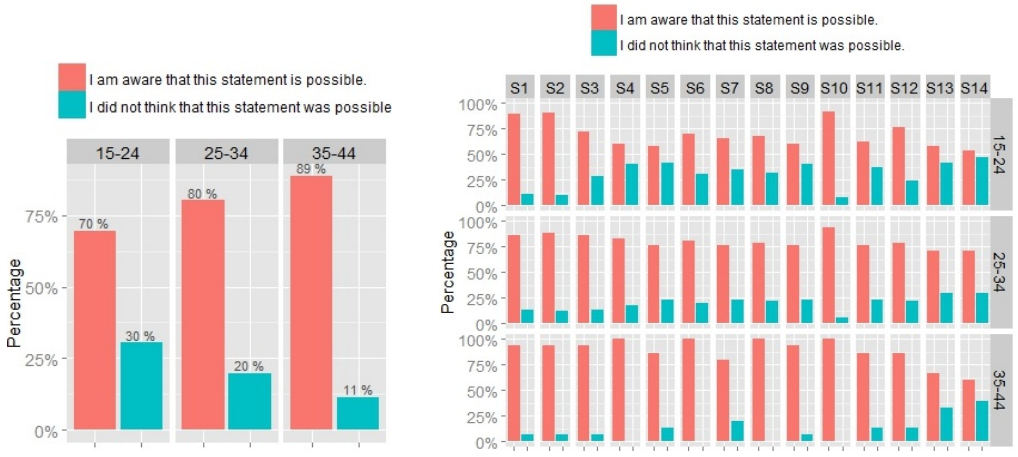
Figure 4.9: Participants' awareness of potential inferences by considering (a)whether users read terms and policies, and (b) gender.

7% more awareness, respectively, than the other groups. Whereas occasional users are the most knowledgeable about S8: inferring when a user is absent from home, representing 8% more awareness then the other groups.

Examining the impact of the representative age group shows that there is a positive correlation between the age of the participants and their level of awareness (Pearson Chi-Square = 46.50, $p < .00001$). In other words, the awareness of potential inferences considerably increases in parallel with the age group, as observed in Figure 4.11(a), where the oldest active age group (35 to 44 years) have the highest awareness rate, scoring 89%, while the youngest (15 to 24 years) has the least knowledge, scoring 70%. The general level of awareness remains the same when exploring the awareness of age-based groups for each inference, yet it slightly varies when considering individual inferences. Users aged from 15 to 34 are most knowledgeable

Figure 4.10: Participants' awareness of potential inferences considering frequency of Social networking use grouped by each inference statement.

about S10, followed by S2 and S1. Whereas all 35 to44 year old users are aware of S4, S6, S8 and S10, about mobility patterns, favourite places, absence from home and knowing friends, as demonstrated in Figure 4.11(b).

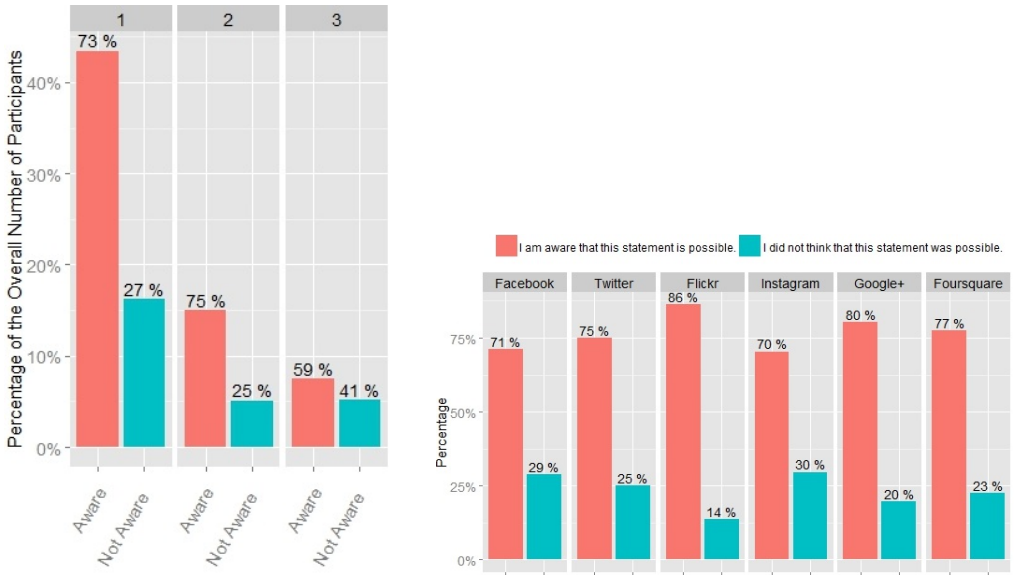


(a) (b)

Figure 4.11: (a) Participants' awareness of potential inferences considering age group, and (b)grouped by each inference statement.

Observing users' awareness by clustering it based on the number of location-sharing applications used reveals that users of one or two services are more aware of the introduced information derivation (about two-thirds of users) than users of three services (59%), who are exposed to

more personal information inferences due to having more platforms to share on, as illustrated in Figure 4.12(a). The awareness of users' of more than three applications is discarded, since the user sample is not representative (participants are less than 10% of total users). Moreover, users of Flickr have the highest awareness rate, followed by Google+, corresponding to 86% and 80%, respectively, while Instagram users have the top unawareness rate, followed by Facebook users, representing 30% and 29% of their users, respectively, as presented in Figure 4.12(b).



(a) (b)

Figure 4.12: Participants' awareness of potential inferences considering (a)the number of location services used, and (b)GeoSN used.

4.3.3.2 Influencing Factors on Inference Concerns

In this section, the factors' influence on participants' reaction toward the presented information inferences is investigated in relation to their demographic and social networking application usage. The use of location services in these applications seems to have an impact on users' privacy concerns. Non-users of location services showed more concern than actual users of these services, representing 62% and 75%, respectively, as demonstrated in Figure 4.13(a). This can suggest that privacy concerns are what are preventing them from sharing their location information. The majority of non-users are 'Very Worried' about the potential personal information inferences (39%), whereas most of these services users are 'Uncomfortable' about the issue (43%). More than half of non-users of location services are specifically 'Very Worried' about S14, concerning other users' knowledge of a user's activity performed at a certain time, followed by S13: other people knowing the place and time of a user occurrence with others, S8: being absent from home, and S3: spatiotemporal visiting patterns. While location service users

are ‘Very Worried’ about S8 (42%). More than 50% of location services users are particularly ‘Uncomfortable’ about S1 and S11, regarding deriving home location and spatiotemporal co-occurrence with friends. Whereas more than half of non-users of location services are ‘Uncomfortable’ about S1 and S2, related to knowing home and work places. Figure 4.14 shows participants’ reaction towards potential inferences considering location services use grouped by each of the inference statements.

(a) (b)

Figure 4.13: Participants’ reaction towards potential inferences considering (a)location services use, and (b)reading of applications’ terms and policies.

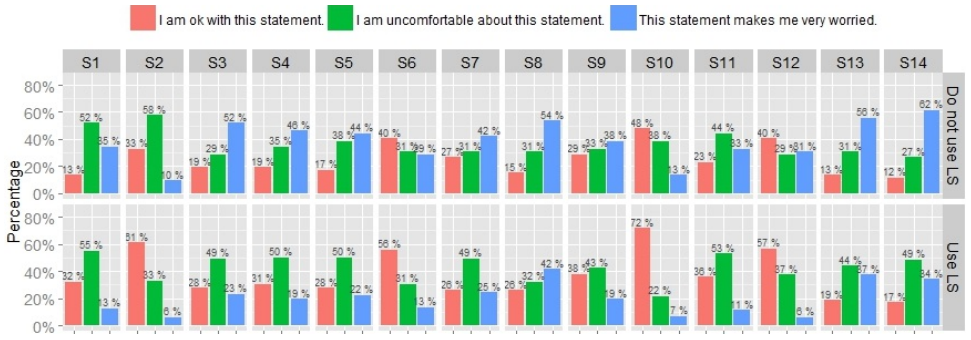


Figure 4.14: Participants’ reaction towards potential inferences considering location services use grouped by inferences.

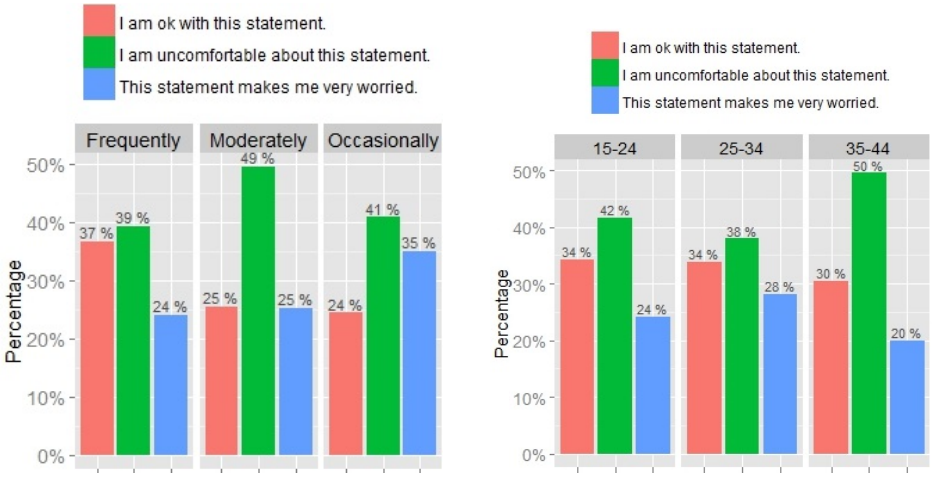
Furthermore, participants’ who have read the terms and policies of the applications seem to be more worried than those who have not, by 9%,which could be due to their knowledge of potential information collection and utilisation, as demonstrated in Figure 4.13(b). However, when observing reactions toward individual inferences, users with no previous knowledge of these terms score the highest reaction of ‘Very Worried’ towards S14 and S13: other people knowing their activity as well as the place and time of an occurrence, and S8: being absent from home. Whereas terms-knowledgeable users are ‘Very Worried’ the most about S8, as shown in Figure 4.15. Moreover, the highest rated ‘Uncomfortable’ inference for terms-knowledgeable users

is S7: estimating the activity performed at a place, while it is related to knowing home place for terms-illiterate users, representing 63% and 55%, respectively. Indeed, reading terms and policies was found to significantly impact users' concern levels (Pearson Chi-Square = 16.867, $p < .00001$). Interestingly, grouping reactions based on individual inferences reveals that male participants are particularly 'Very Worried' about 29% of the inferences, while females are 'Very Worried' about only 7%.

Figure 4.15: Participants' reaction towards potential inferences considering reading applications' terms grouped by inference statements.

Frequency of use of social networking can be seen to impact users' privacy concerns, where there is a negative correlation between the frequency of use and worry regarding the inferences. As noticed in Figure 4.16(a), the level of concern increases with the reduction of use frequency, where 76% of 'occasional' users are concerned, compared to only 63% of 'frequent' users. It is noted that the frequency of use of these applications has a significant impact on the level of concern (Pearson Chi-Square = 35.636, $p < .00001$). Specifically, 'frequent' users are most worried about S8: being away from home (45%), and most uncomfortable with S1:knowing home location (55%). Moderate users are worried the most about S8 and S13:other people knowing the place and time of a user occurrence, representing 44% each, whereas they score the highest 'Uncomfortable' score with S5: knowing the routine mobility patterns (72%). Similarly, 'occasional' users are 'Very Worried' the most regarding S14: other users knowing the activity performed at a place, while knowing the workplace triggers the highest 'Uncomfortable' score. Figure 4.17 presents participants' reaction towards potential inferences, considering the frequency of use of social networks grouped by each of the inference statements.

Considering the age factor, younger users in the age group of 15 to 34 years are 4% less concerned than older users of the age group 35-44, as shown in Figure 4.16(b). This can be due to the fact that young users are less aware of potential inferences, as discussed in Section 4.3.3.1. 15 to 34 years old users are worried the most about S8: inferring absence from home, repres-



(a) (b)

Figure 4.16: Participants' reaction towards potential inferences considering (a)the frequency of use of social networks, and (b)the age group.

Figure 4.17: Participants' reaction towards potential inferences considering the frequency of use of social networks grouped by inference statements.

enting on average 49%. S1: knowing home location, has the highest 'Uncomfortable' score for the 15-24 age group (60%), whereas the 25-34 age group are most 'Uncomfortable' about S2 and S4, besides S1 regarding knowing workplace and normal mobility patterns (45% each). As for the 35-44 age group, they are the most 'Very Worried' about S14 (47%), and the most 'Uncomfortable' regarding S12:deriving the most seen friends (73%). Figure 4.18 illustrates participants' reaction towards potential inferences by considering the age group grouped by inference statements. In addition, users of Facebook and Instagram registered the highest degree of concern among all users of GeoSNs scoring 63% and 62% respectively, whereas Flickr users are the least concerned (49%), as shown in Figure 4.19.

Figure 4.18: Participants' reaction towards potential inferences considering the age group grouped by inference statements.

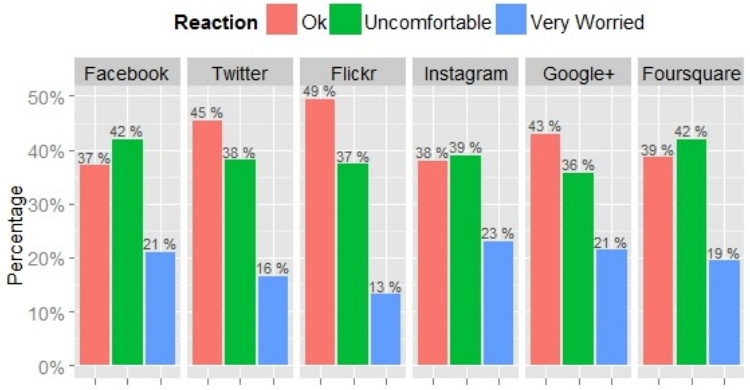


Figure 4.19: Participants' reaction towards potential inferences grouped by the GeoSNs used.

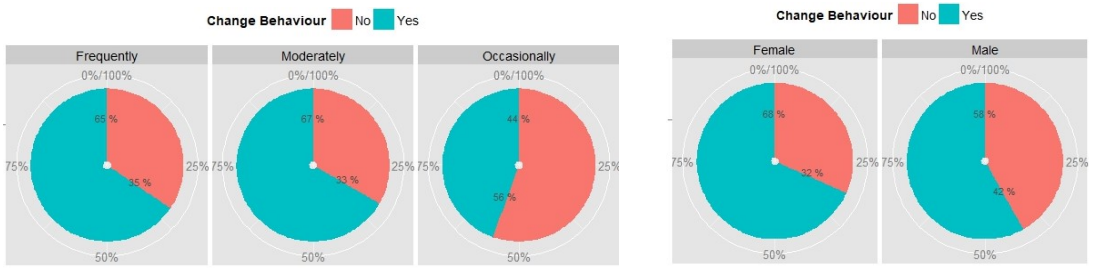
4.3.4 Attitude to Privacy on Social Networks

The aim of this section of the questionnaire is to understand the users' reaction with regards to sharing their location information on these applications, given the knowledge of potential implications on privacy from the previous section. Therefore, only the participants who disclose their location are considered in this examination.

4.3.4.1 Willingness to Change Location Sharing Behaviour

The majority of the users stated that they would change the way they share their location information on the GeoSNs representing 64% of them. This fact denotes that raising users' awareness of the potential personal information extraction based on location disclosure to the end user have a significant impact on the way they perceive the location-based privacy threats and hence their share their location information. Frequent and moderate users seem to be the most mo-

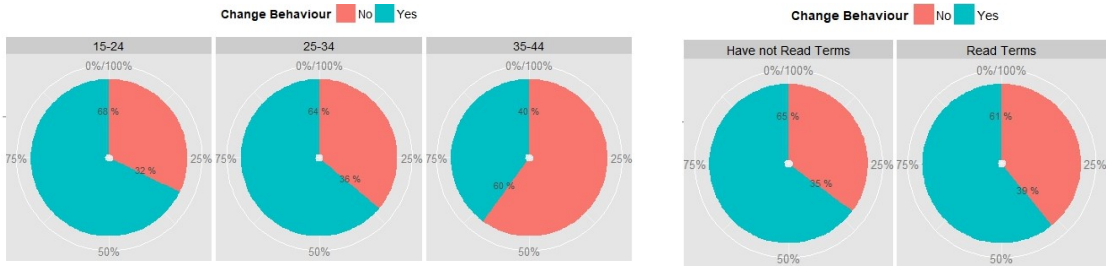
tivated to change their sharing behaviour (10% more than occasional users), as illustrated in Figure 4.20(a). Female users showed grater tendency to change their attitude regarding location disclosure (by 10%) than males as shown in Figure 4.20(b).



(a) (b)

Figure 4.20: Participants' reaction to location privacy risk, grouped by (a)the frequency of use of social networks, and (b) gender.

Interestingly, Younger users (15-34) are more willing to change their usage behaviour (by 26%) than older age groups (35-44) who were mostly less motivated for such a change, as illustrated in Figure 4.21(a). However, whether users are knowledgeable of the terms and policies of social networking applications has a minor influence of their behaviour where policy-literate users have a slight increase in their tendency to modify how they share their location information (of 4%) over illiterate users, as presented in Figure 4.21(b).



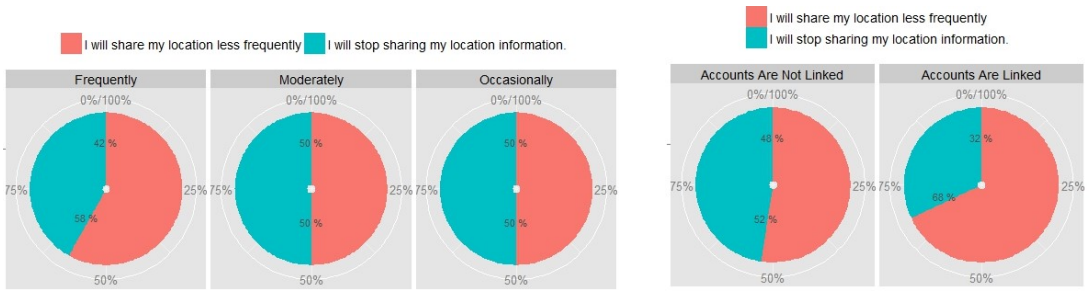
(a) (b)

Figure 4.21: Participants' reaction towards location privacy risk, grouped by (a)age group, and (b)literacy of application terms and policies.

4.3.4.2 How to Change Location Sharing Behaviour

This section investigates how the participants would change the way they their share their location information on the GeoSNs. More than half the users would stop sharing their location information, while 43% of them would prefer to share it less frequently. Clustering users' responses based on how frequently they use the social networks shows a relatively negative correlation between willingness to stop location sharing and the frequency of use, as seen in Figure

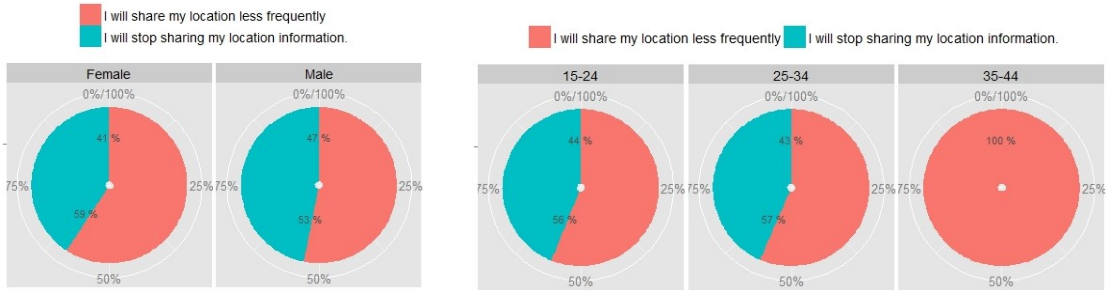
4.22(a). In other words, the less use of social networks, the greater tendency to discontinue location information disclosure. Frequent users are less willing to stop location sharing, and would rather reduce their location-sharing frequency (58%), than infrequent users. Similarly, Users who have linked their accounts on social networks tend to be less willing to stop sharing their location (32%), and would prefer to share it less frequently instead, than users who have not, as illustrated in Figure 4.22(b).



(a) (b)

Figure 4.22: Participants' preference for changing location sharing behaviour, grouped by (a)the frequency of using social networks, and (b)whether their account are linked.

Although female users are more motivated to change their sharing behaviour, as discussed in Section 4.3.4.1, males showed more likelihood of discontinuing their location sharing activities (by 6%) than females, as presented in Figure 4.23(a). Age also has a noticeable impact as well where older users (35-44 years) showed 100% willingness to reduce their location-sharing frequency (with no interest in stopping such sharing), compared to only 57% of younger users (15-34 years)s, as presented in Figure 4.23(b). Whether users read the social networking applications' terms and policies has no effect on how they would change their sharing behaviour as the number of users who favour each of the reactions is identical.



(a) (b)

Figure 4.23: Participants' preference for changing location sharing behaviour, grouped by (a)gender, and (b)age group .

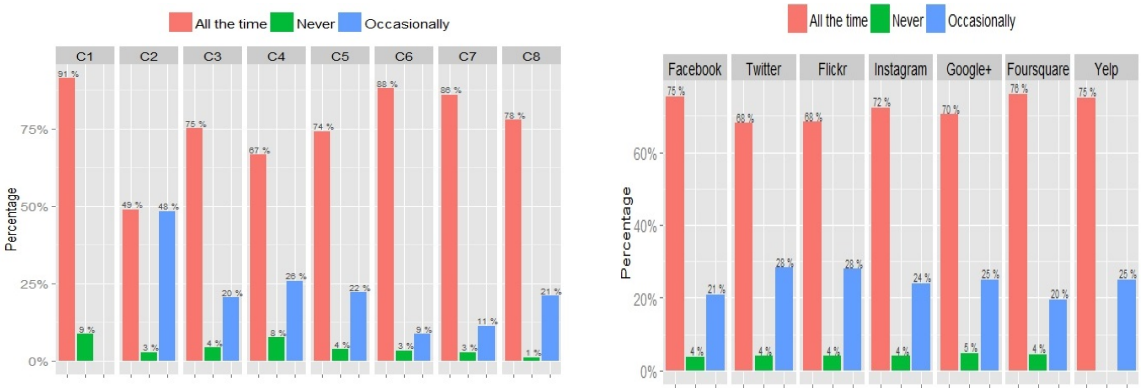
4.3.5 Managing Personal Information

In this section, users' views on managing and controlling access to their location information are explored. This includes several aspects related to what information is stored, how it is shared or viewed by the application and by others, and whether users need to manage access to their information. The following statements were presented to the participants who were asked to rate how often they would use them: 'All the time', 'Occasionally' or 'Never'.

- *C1: I would like to be able to turn off location sharing for specific durations of time.*
- *C2: I would like to turn off location sharing when I visit specific types of places.*
- *C3: I would like to decide how much of my location information history is stored and used by the application; for example use only my check-in history for the last 7 days.*
- *C4: I would like to see the predicted personal information that the application stores about me based on my location information.*
- *C5: I would like to decide how people see my current location; for example, exact place name, or a rough indication of where I am.*
- *C6: I would like to decide who can download my location information data.*
- *C7: I would like to know, and control, which information can be shared with other Web applications.*
- *C8: I would like to make my location information private; seen only by myself and by the people I choose.*

The results are given in Figure 4.24(a) and show a significant desire to use these controls for location privacy. 76% of participants would like to apply those controls 'All the time', 20% are happy to apply them 'Occasionally', and only 4% of users would not consider these controls. In general, C2, C6, C7 and C8 were most favoured controls to use, scoring over 97% each of users' responses. Controls C1, C6 and C7 were the most chosen controls to be applied all the time, representing 91%, 88% and 86% of users' responses respectively. It is worth noting that users of different location services have similar acceptance rates for these controls. Foursquare and Facebook users have the highest preference for applying the controls 'All the time', corresponding to 76% and 75% respectively as shown in Figure 4.24(b).

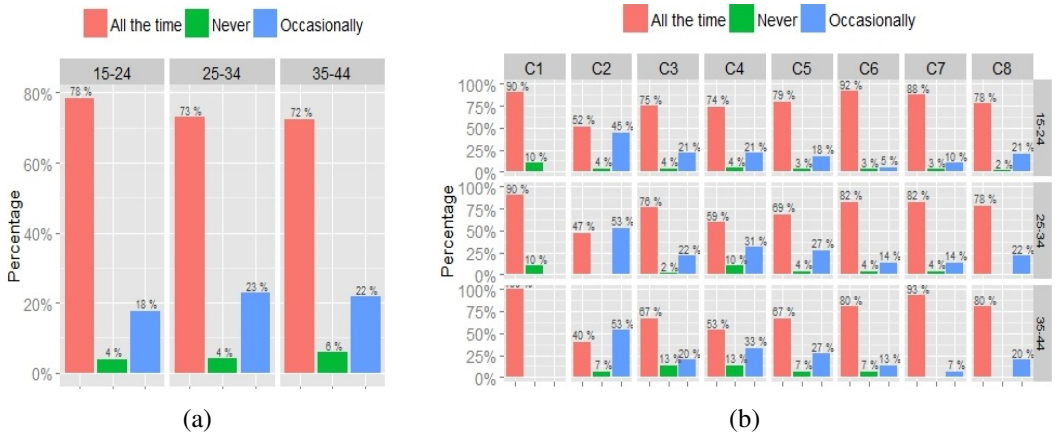
When taking into account users' age, a negative correlation appears to exist between users' tendency to use these privacy controls all the time and their age group as demonstrated in Figure 4.25(a). The youngest active age group of 15-24 years old has the highest desire for 'All the



(a) (b)

Figure 4.24: (a) Users' desire to use location privacy controls grouped by statement of controls C1-C8. (b) Data in (a) grouped by the location services used on GeoSNs..

time' application of controls representing 78% of this group's responses which corresponds to a 6% increase from the oldest age group of 35-44 years. This association is shown clearly when clustering responses based on the individual controls as illustrated in Figure 4.25(b), particularly for C4: personal information prediction by the application, where the youngest user group showed an average increase of 21% over the oldest users, as well as C2: turning of location sharing for specific location types, C5: how the user location is seen by others, and C6: managing who can download location data, which showed an average increase of 12%, descending by age group.

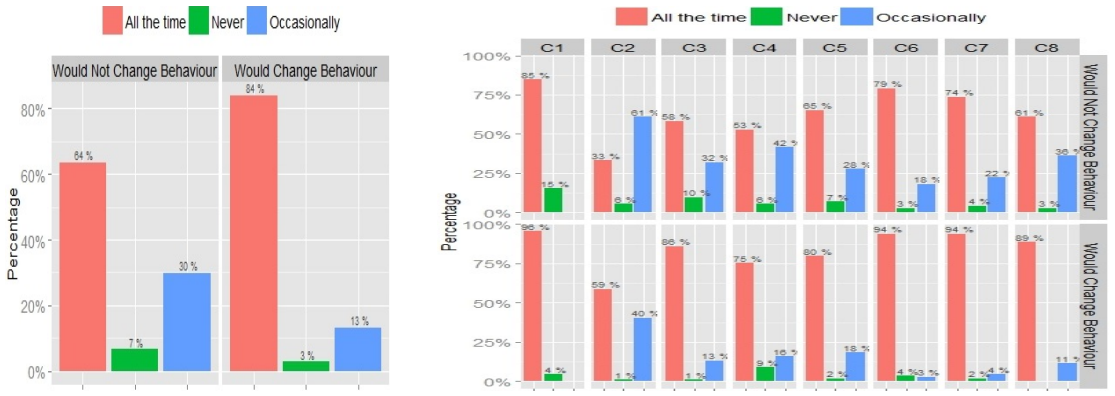


(a) (b)

Figure 4.25: (a) Users' desire to use location privacy controls considering the age group. (b) Data in (a) grouped by the controls.

As expected, users willingness to change their location-sharing behaviour has a considerable impact on their tendency to use the privacy controls (Pearson Chi-Square = 81.170, $p < .00001$). Users who are tempted to change their location sharing behaviour have relatively higher motivation to use these controls (by 4% for general use, and 20% to be applied 'All the time') than

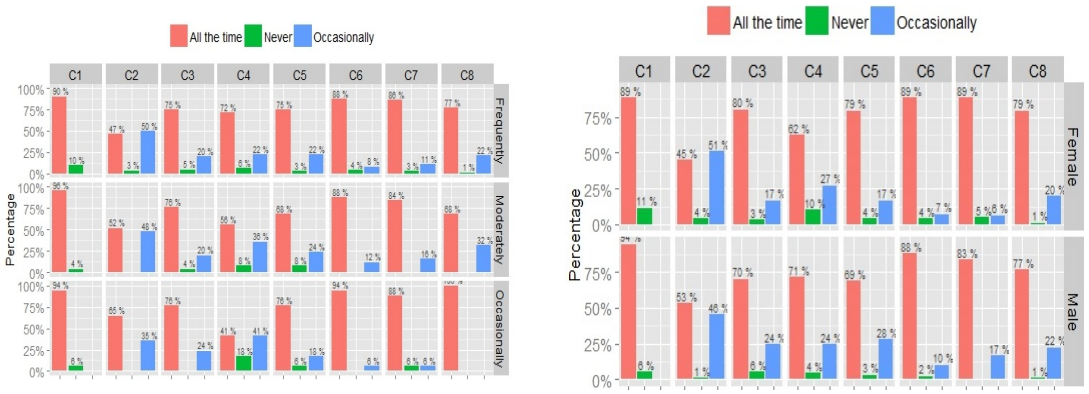
users who are reluctant to change, as can be seen in Figure 4.26(a). The most favoured privacy controls of users who want to change their behaviour are C6 and C8, related to managing the visibility and accessibility of location data, representing 97% of users' reactions. Whereas users who were not willing to modify their location sharing preferred C8, representing 100%, followed by C2 and C3 regarding controlling location information history storage and use by the application, representing 99% of these users' attitudes, as presented in Figure 4.26(b).



(a) (b)

Figure 4.26: (a) Users' desire to use location privacy controls considering their willingness to change their sharing behaviour. (b) Data in (a) grouped by the controls..

Other factors have no effect on users' desire to use privacy controls in general, yet they influence their preferences towards using individual controls. Frequent users are more tempted than moderate and occasional users to use C4 and C5 regarding personal information prediction and how the user's location is seen by others, as demonstrated Figure 4.27(a). On the other hand, moderate and occasional users are the most willing to use C2: turning of location sharing for specific location types, C6: managing who can download location data, and C8: making location information private, representing 100% of these users. As for gender, the most desired control for male users is C7: knowing and controlling which information are shared with other applications (100%), followed by C2: turning of location sharing for specific location types, and C8: making location information private (99% each). Whereas female users prefer C8 the most, followed by C3: deciding how much location information history is stored and used by the application, representing 99% and 97%, respectively, as presented in Figure 4.27(b). However, there is no considerable difference in general attitudes between users with or without knowledge of the application terms and policies.



(a) (b)

Figure 4.27: Users' desire to use location privacy controls grouped by the controls by considering (a)frequency of using social networks, and (b)their gender.

4.4 Discussion

This survey provides valuable insights that covers many essential aspects of location privacy on the social web from the end user's perspective. In particular, it investigates users' level of awareness about their location information accessibility, collection and utilisation, whether it is knowledge of the official terms of use and privacy policies, or another indirect means of inference. Users' reaction towards potential personal information inferences, and how likely users are to use possible location privacy controls are also captured. This survey examines how the users' attitude is impacted by factors related to demographics or use of social network services. The user sample that participated in this survey is a reasonable representation of the general users of social networks, especially users of location services, where more than half of users are aged between 15 and 24, and more than two-thirds use social networks several times daily, as well as equal participation by male and female users, with representation for the most common location services on the GeoSNs.

The majority of users, especially those who use location services, have not read the terms and policies of the applications they use, which is generally considered the only way they can educate themselves about how their location-based information is stored and exploited. This means if these terms are updated to include more application rights that imply more privacy-threatening use of users' data, the users would not be aware of them and continue use the application regardless. The unawareness level for the presented terms is significant, particularly for location services users, rising up to 48%. Overall, users are less knowledgeable about the terms and policies that contain more privacy provoking statements, such as collection of users' exact location and other personal information from other web applications. Moreover, frequent users who are generally more prone to the privacy threat since they have more publicly available data, are

considerably unaware of potential risks to their privacy. However, 15 to 24 year old users tend to be more knowledgeable of these terms. As expected, users who read the terms and policies are more aware of them. It seems that there is no gender bias for this level of awareness.

In terms of users' awareness and reaction towards potential personal information inferences, it seems that general knowledge is reasonable. However, the level of concern about privacy is significant (up to 84%, where 38% are 'Very worried'). This unexpected behaviour, where users are highly concerned in spite of the fact that their awareness of these inferences is considerable, might be due to the means and context of presentation or the trade-off between privacy issues and the benefits that these services provide. Moreover, it can be argued that users have no comprehensive understanding of the potential information inferences, since the top two inferences with the lowest awareness and highest concern rates are "Other people can know where you are at any point in time", and "Other people can know what you are doing at any point in time". Whereas other inferences implying the same meaning, such as "I know which places you visit and at what times", and "I can guess what you do when you are in a specific place" triggered less concern. This suggests that if the same inference is presented in a more privacy-provoking style, such as using "other people" instead of "I", can have a different impact on how users perceive the inference, and their level of concern.

Users who have not used location services on GeoSNs are more aware of these inferences and are, as expected, more concerned, which can explain why they have not utilised them. In addition, awareness and level of concern increases with users' age. Nevertheless, younger users who represent the majority of social networks' users tend to be even more worried about certain inferences than older users. Furthermore, users who read the applications' terms are more aware and concerned, as would be expected from users who care enough to read these terms.

When capturing the overall attitude of users after presenting to them the plausible inferences, the majority of users stated that they would change the way they share their location information, where more than half are willing to stop sharing completely, and the rest want to share less frequently. This indicates that they are triggered by these inferences as well as the terms and policies, and hence highly concerned about their location privacy. Users who disclose their location more often are the most motivated to change their sharing behaviour, which might be due to their realisation that they share more of their information online and need to change their behaviour to protect their privacy. Younger users aged between 15 and 34 have a higher tendency to change their sharing behaviour than older users. It seems also that female users are more motivated to modify how they disclose their location information than male users.

The proposed location privacy controls are highly desired by the users (96%), where two-thirds of them want to apply them all the time, particularly users of Foursquare and Facebook which are two of the most popular GeoSNs. As expected, the users who are tempted to change their

location sharing behaviour have higher motivation to use these controls than the users who are reluctant to change.

4.5 Conclusion

This chapter provides valuable insights that convey many aspects of location privacy on the Social Web from the perspective of the end user. The main and (possibly only) means of communicating how the collected user information may be used and exploited by the application is described in the application‘s terms of use. It is clear from the results of the questionnaire that the majority of users, especially those who use location services, do not read the terms of use and policy documents. The findings also indicate that users are aware of the potential information, and possible derivatives thereof, stored by the application. However, it appears that they are also concerned about the privacy implications. The reason behind this apparently contradictory finding may be because such awareness and concern is evident when users are actively questioned about these issues, but is somewhat screened from the users‘ minds during their continuous use of the application. The study also suggests that users may not fully understand the privacy implications, where their level of concern was much more pronounced when faced with statements that indicate that other people may be aware of their location information in comparison to statements indicating that the application holds such information.

The study reveals that there is a strong need for the users to be continuously aware of their data and how it is stored, and to have the ability to control access to and visibility of their location data sets. Further research into methods that enhance communication of information by the applications is necessary, and there is also a need to allow users to better understand and manage their personal profiles on such networks.

Chapter 5

Feedback Design for Location Privacy in GeoSNs

5.1 Introduction

As was evident in previous chapters, most users of GeoSNs are concerned about their location privacy and have limited awareness of potential privacy implications in terms of the underlying collection and exploitation of their personal information. At the same time, GeoSNs lack support for privacy awareness methods in terms of offering effective privacy management tools as well as presenting a detailed view of personal information collection and potential utilisations. Generally, users are only conscious about their current location discloser and tend to forget or have difficulty recalling all of their location history that is built based on their previous location sharing on GeoSNs and exposed to other users. For example, when a user is about to share a location, this location or its related properties such as frequency of visit can pose privacy implications, yet this user would continue sharing it because she would most probably fail to understand and see the related privacy consequences of her sharing action and at the same time there is no means available that notify her about such potential risks.

Generally, privacy feedback and control systems have proven to impact users' privacy behaviour in terms of increasing their awareness and changing their data-sharing attitude [30]. Moreover, offering real-time privacy warnings enables immediate awareness of and reaction to the implications and management over their sharing behaviour [36, 106]. In addition, employing information presentation techniques seems to be effective in enhancing users perceptions of the privacy implications, and hence motivate them to take actions to protect their privacy [33, 21].

In this chapter, feedback design for location privacy awareness is proposed and evaluated. A model is first introduced for a user goe-profile in relation to privacy on GeoSNs by examining the dimensions of location data sharing on GeoSNs and how user awareness is situated in this information space by analysing users level of awareness while using GeoSNs. Then, This feedback design is developed based on this model's outcomes which basically provides real-time

privacy notification for users related to a specific location-sharing action they are about to carry out including the visibility of their data to others as well as data that may be shared and extracted based on such an action. Then, an evaluation is conducted as an online user study using realistic location-sharing scenarios to test the impact of location privacy awareness in the form of a feedback tool on users' privacy attitudes and behaviour on GeoSNs.

Few studies have addressed location awareness and its impact on users' behaviour [36, 37, 38], and those that have mainly question users' attitude with respect to sharing their current location information. These studies also neglect to reveal the privacy implications in terms of what personal information can be extracted based on location disclosure using location-sharing applications with limited interaction features for their evaluations, which are insufficient for reflecting public GeoSNs environment. This chapter 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 behaviour on GeoSNs.

5.2 User Geo-Profile Dimensions

Users intentionally declare to GeoSNs 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). This section examines 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 GeoSNs 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 Facebook Places APIs [1], and Twitter uses the Google API for linking users' selected place name with a location on a map.

[1] Instagram Location Endpoints, `instagram.com/developer/endpoints/locations/` [Accessed: Jul. 2016]

Given the place identifiers on a ST track, other useful place properties can be extracted, for example, the type of 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. [126]

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

- What concepts are the user interested in?
- Where would the user be associated with 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's geo-profile would be capable of supporting the following queries.

- When did the user visit a place? How often? How much time did she/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)?
- What is the relationship nature between this user and a particular friend (e.g. housemate)?
- What is the user's association with a particular place (e.g. workplace)?

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?

5.2.1 User Awareness of a Geo-Profile

Unlike other information stored on a personal profile, for example, an email address or a phone number, user location information on a GeoSN is dynamically updated as the user makes use of the application services. These location tracks are not always visible to the user. When interacting with the software system, the user's attention is normally task-oriented [127], i.e., she needs only recall information that is relevant to the task at hand. On GeoSNs, a user either consciously shares a particular location, or the service automatically captures her location and uses it to tag her resources (photos, tags, etc.). In the first case, the user is aware of one single point on the place-time plane, indicating her presence in the place at a particular point in time and may be aware of the visibility scope she has previously set on her profile, and thus can recall who she is sharing this location with. In the second case, the user is not even aware of the location capturing activity and may not need the information for task execution.

A spotlight metaphor (as used to explain visual attention [128]) can be used to describe user location awareness, where a user has only a limited view of their geo-profile space. The question arises then as to how much of the profile needs to be revealed to the user to enable them to make informed decisions on the privacy of their own data. An even more fundamental question is whether the user data need to be recorded and stored to provide the services. For example,

if the service relies on identifying the types of places a user is interested in, does it need to continuously store their presence in these types of places or can it stop tracking if sufficient information has been inferred.

A practical investigation was carried out first to examine how users of public GeoSNs can view and retrieve their information submitted on these services by exploring the tools provided to them for accessing and visualising their data. The aim was to assess the effort and the extent of support detected for enhancing users' privacy awareness of their data disclosure and related privacy consequences, whether by the application itself (including Google and Foursquare as example), or third parties (including Foursquare TimeMachine, 4sqmap.com, and Creepy). A full review of information presentation and visualisation tools available for users can be found in Appendix B.1. The means provided to the users of GeoSNs for viewing their shared data is limited and do not focus on educating them about potential privacy implications. These tools basically allow users to see only what they have explicitly shared, which is mainly a stream of time-stamped content such as location tracks. They lack the capacity to show the users what data have been collected implicitly and what their data are utilised for, including inferring personal information such as mobility patterns and co-location with friends. They also do not specifically inform the users of who can access their data, whether from inside or outside the service. Hence, there is a need to support privacy awareness about these later practices in order to enable the users to take charge of their data and make informed decisions about what they share.

In addition, a systemic task analysis was conducted to investigate location-based tasks in GeoSNs in order to address the gap in the knowledge of users' awareness and the related privacy risks. In particular, it aims to explore the user's mental model in terms of what they are conscious of when undertaking a location-related activity regarding their information acquisition and projection. The approach used here is a combinations of hierarchical and cognitive task analysis [132, 131]. Foursquare, including Swarm, its check-in application, is selected as a representative case study for the task analysis as it one of the most popular GeoSNs. The analysis was carried out in February 2015. The full tasks' analysis can be found in Appendix B.2. As a result, Foursquare collects rich user data, yet it lacks support for privacy-awareness in regards to showing users privacy-related information and how their information is actually processed and handled by the application. Generally, users' awareness seems to be confined to their current interaction with the applications, including a) the information they are required to enter in order to complete the tasks, and b) the information displayed to them as a feedback from the application. In particular, information required to perform any task is fairly basic and discrete. This mainly includes:

1) A place name, type, or concept of interest.

2) The users' place reviews (tips).

3) Tagging friends.

Feedback information projected to users during or after completing a task is limited to what they have just explicitly submitted to the network, and confirmation of task completion including:

1) A place profile (details and address).
2) Other people's reviews(tips) on a place.
3) Where the friends are in correspondence to the users' location, and their ranks in terms of their check-in activity.

However, there is another hidden layer within the application that users are unaware of and may pose a threat to their privacy. It deals with aggregating their data over time and is capable of deriving more personal information about them based on their previously listed interactions, including:

1) Revealing the user current presence or absence from places, and sensitive places
2) Extracting frequency of visit to places.
3) Extracting interests and activities.
4) Inferring spatial-social-temporal mobility patterns.
5) Deriving co-locations with friends and other users and detecting relationships .
6) Predicting transitions between places and future movement of the user.

Consequently, there is a substantial difference between what the users intend to share and the extent of personal information that can be implicitly inferred. Thus, there are privacy-awareness gaps that users are unable to appreciate, which can be identified in the following three main areas:

- **Data collection:**
  When performing tasks on Foursquare/Swarm, it is not explicitly shown to the user what data of theirs is being collected or how it is stored and linked to their historical profile, which can include their extracted personal information as listed above.
- **Data utilisation:**
  The purpose of data collection, beside what is required to carry out a task, and how it is exploited for personal information inference, is hidden from the users. Their data can be used for advertisement/ marking purposes or shared with other organisations.

- **Data accessibility:**
  The are no means of informing the users of who exactly can view or has already accessed their data. Users are conscious that their general group of friends on the application can see their submissions, but the extent of their data that can be viewed is not made clear to them. In addition, their data can be accessed by the API users and third party companies, where such accessibility is not made explicit to users.

5.3 Feedback Design for Location Awareness

To enable user content awareness in GeoSNs, privacy-enhancing feedback and control tools need to be designed and incorporated within the services. The development of such tools must 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 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. A default estimation mechanism can be used to determine the threat level, such as the one described below, but this can be customised 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. Usability in HCI is generally concerned with the effectiveness, efficient and satisfaction within a particular domain of use [134]. Several research works have considered usability issue and proposed design elements and principles for designing privacy-feedback systems. One of the most important design aspect is to informing the users what it is collected of their information and how it can be used for, and allow them to explore how their information is processed and stored [95, 94, 97, 98, 96]. Meaningfulness of the presented notice is also key factor in

deign usability where the feedback should be expressed in understandable form for the user to ensure it effectiveness [95, 94]. In addition, providing controls for that are suitable for the context of use is related to how usable the feedback is to the users in supporting privacy management [98, 94, 97, 23]. Timing of the presentation of the feedback plays also an essential role in triggering the needed reaction by the user which ultimately enhances the effectiveness and efficiency of the feedback design [23, 95, 106]. Based on considering these design aspects, content provided in the context of location-awareness feedback tools should comply with the following properties.

- Relevant: Avoid information overload by presenting enough information that is relevant to the task at hand. The user should be able to recognise the relevance of the presented information to the location sharing activity they are undertaking.
- Transparent: The user should be able to query the system's reasoning on the data presented and the privacy threat indicated. The system should provide access to a complete view of the user's geo-profile.
- Timely: Feedback should be presented at the point it is needed to make a decision about privacy, and be presented in a way that avoids unduly diverting the user from the task at hand.
- Actionable: Appropriate control actions should be available to the user to respond to the feedback provided.
- Comprehendible: Presentation of content should enable accurate interpretation of all elements of content provided.

5.3.1 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 [15]. 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, an initial model of the levels of privacy threats with respect to the user geo-profile is proposed. The aim at this stage is to introduce logical mapping to threat levels, and use this to measure the influence of the provision of a privacy threat level on users' awareness and sharing behaviour, which will be discussed in the experiment findings, Section 5.5. A more elaborate study that focuses on objectively modelling

levels of threat to location privacy through investigating users' location privacy perceptions is presented in the second stage, in Chapter 6.

The model assumes that two variables determine the threat to a user's privacy, namely, the amount and content of disclosed information, and the visibility scope of the information. Three levels of visibility scope can generally be considered: a) private (no access to other people), b) friends (access only to user's friends), and c) public (access to others whether inside or outside the social network). The extent of exposure of the user's data can be measured along the following data dimensions: spatial, social, spatial-social, or spatial-social-temporal. Guided by the results of the user study presented in Chapter 4, three levels of privacy threat are proposed to represent users' perception.

- Green: safe to disclose the information,
- Amber: caution; disclosing the information can result in moderate privacy implications,
- Red: danger; disclosing the information can result in risky privacy implications.

The results of that user study are utilised to derive an abstract model of privacy threat levels. Table 5.1 is a summary of the level of privacy concern against the data dimensions revealed by that study. In Table 5.2 the values in Table 5.1 are classified under three categories, as follows. In the case of the public scope of visibility, an 'Amber' classification was used with a threshold of $\geq 30\%$ for the 'Uncomfortable' attribute value and a 'Red' classification was used with a threshold of $\geq 30\%$ for the 'Very Worried' attribute value. In the case of friends scope of visibility, a 'Green' classification was used with a threshold of $\geq 40\%$ for the 'Ok' attribute value and an 'Amber' classification was also used with a threshold of $\geq 40\%$ for the 'Uncomfortable' attribute value. Nevertheless, the disclosed data in the spatial-social-temporal dimension can reveal more about users' personal information, which not only includes their mobility patterns, but also their future movement predictions and private places. Users can fail to recognise these related implications. Thus, considering the potential value of the information disclosed in spatial-social-temporal dimension, a 'Red' instead of 'Amber' classification was assigned to the friends scope.

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 to visibility to the *Friends* group, access to data on the individual dimensions; spatial and social axes, are assumed to be granted (green), whereas the threat level increases with the likelihood of disclosure of implicit data along composite dimensions.

Table 5.1: The average privacy concern level categorised by the data dimension of the presented inferences.

Dimension	Privacy Concern Level		
	OK	Uncomfortable	Very Worried
Spatial	*43.9%*	*41.6%*	14.5%
Social	*58.6%*	*30.6%*	10.8%
Spatial-Social	26.3%	*44%*	*29.6%*
Spatial-Social-Temporal	25%	*42.6%*	*32.3%*

Table 5.2: 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	red
Spatial-Social-Temporal	green	*red*	red

Note that using the dimensions in Table 5.2 only provides an abstract model of the level of threats associated with these dimensions. Finer specification of the data elements, and relationships between data elements, shared or inferred, along these dimensions is possible and would give more insight to the threat levels inherent within the information space of a geo-profile.

5.3.2 A Design Proposal for Location Privacy Feedback Tool

As an application of the above design ideas, a possible feedback tool for location awareness is considered. 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's geo-profile.

Such a feedback tool can be integrated within GeoSNs either by the service provider or as a third-party application. In the first case, the feedback tool would be a built-in feature that is provided by the GeoSNs where users can be notified about their location privacy while using the application as a part of service provided. In the second case, this tool would be implemented independently from the GeoSNs as an application that runs on top of the GeoSNs. Such an application would need to use the GeoSNs APIs with users' permission to access their accounts in order to retrieved and construct their geo-profiles required for providing the location privacy notifications.

One possible design is shown in Figure 5.1. 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 beyond the scope of this work. What follows is the description of a privacy tool which revolves around three main factors: what to present to the user, how, and when. This tool design follows the design properties discussed above in Section 5.3 in addition to these finer design principles:

- Learnability of the tool in which it can be used easily by the user without any overloading. [135, 98, 95]
- Familiarity with the design elements used in the tool (e.g. icons and colour-schemes) [135, 112]
- Minimising distraction from the users' main task [95, 94]
- Using simplified and succinct language for information presentation [94, 17, 113]

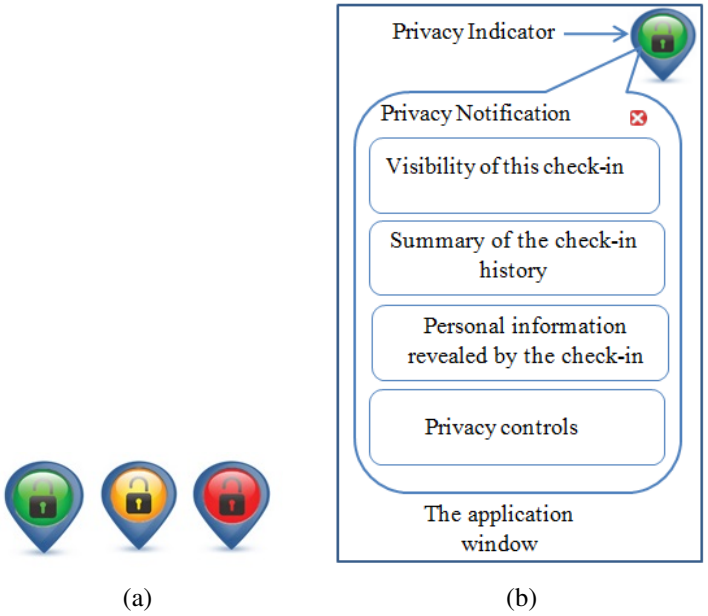


Figure 5.1: (A potential design of the privacy-enhancing feedback and control tool showing the (a) icon design for the privacy indicator, (b) content of the privacy notification tool.

5.3.2.1 What to Present: Information Content

The information presented to the users represents the essence of this tool, and actually shows the user the privacy risks triggered by performing a location-oriented task, mainly in terms

of what personal information can be dynamically derived by doing so. A security warning should show the sensitivity of the information entered by a user but not present general warning statements [113]. Basically, the tool dynamically retrieves the information associated with the current check-in from the recently updated user profile. The retrieved privacy-related personal information is presented in three main elements of the tool: Feedback, Privacy Implications, and Visibility that correspond to three main privacy aspects that a user should be aware of [95]. These elements are interconnected to provide complete information about the current privacy situation in order to allow users to make informed decisions about their location sharing actions as recommended by [113]. The tool's elements are defined as the followings:

1) **Feedback element:** This provides a summary of the user profile that shows frequency-based information related to the current check-in which represents the data collection aspect. The reasons for showing the feedback to the users is to enable them to recap their history related to the current check-in and to allow them to understand what triggers the presented privacy implications in the tool. Essentially, it presents three main pieces of user information as follows:

   - Check-in frequency to a place
   - Check-in association with temporal information
   - Check-in association with friends

2) **Privacy Implications element:** This presents a view of the geo-profile that lists possible constructed personal information based on this check-in as a way of presenting the associated privacy threats which represents the data utilisation privacy aspect. The related risks are derived based on the user's check-in history along with the current check-in task in hand.

3) **Visibility element:** This shows the visibility permissions granted in terms of who can view this check-in which represents the data accessibility privacy aspect. Although the user check-in is essentially recorded and viewed by the application and its third parties, the user is not necessarily conscious of them. Therefore, to avoid confusing the users with this default, the visibility is presented from the user's perspective, which involves friends (social connections), other users of the service, and other social networking applications if the user choose to share on them.

5.3.2.2 How to Present: Information Presentation

How the information is presented to the user is responsible for attracting their attention, properly conveying the required information, and enabling them to utilise the tool optimally. Based

on examining the findings of relevant work (e.g., [110, 34, 112, 124, 111, 17]) as well as considering the features of GeoSNs, our proposed privacy notification consists of two main parts as follows:

1) **The Privacy Indicator:**
The privacy indicator is embedded in the GeoSN and offers a simple and direct indication of the user's current privacy level in the form of a location pin with a "lock" icon, as shown in Figure 5.1(a),which is chosen based on its familiarity and link to privacy and security issues. The level of privacy is indicated by a three-colour scheme using traffic light colours. This scheme is chosen for its familiarity and association with safety and danger in order to simplify perception of the privacy information. It reflects the level of threat estimated by the system, which is based on the threat levelling model presented in Section 5.3.1.

2) **The Privacy Notifications:**
This part presents the privacy information related to a location-oriented action in detail, including the three main elements discussed previously as well as the privacy control options illustrated in Figure 5.1(b). The notification is shown when the indicator is clicked, allowing the user to explore their content to understand the basis of the threat indicated. In the privacy notifications window, the visibility element is represented by an eye icon. Parties who can view this check-in are also represented using icons. Then, the feedback is presented in the form of a natural language sentence including the three pieces of information in order to simplify the contents. Lastly, the Privacy Implications are shown as a list of labelled tuples. Previous studies have demonstrated presenting privacy information in the form of labelled short text in an organised manner improves comprehension of the presented information, as well as making it easier and faster to find [17, 113]. Each of the labelled tuples presents single or complex privacy implication including two types of information extracted based on the user's location profile, along with the information provided in the current check-in, as follows:

   - *Static Types:* This shows type of the personal information that can be inferred such as "Private place". It provides general warnings that can be shown to other users with similar profiles.
   - *Dynamic Contents:* This shows exactly what information is inferred such as "Home", which is specific to a particular user.

   The labelled tuple is formatted as *Static Types ("Dynamic Contents")*. In this way, the user can easily perceive the general privacy consequences of this check-in as well as the specific inference in this user's case. As for the privacy controls offered for the users,

a tick-box labelled 'Remove' is shown in front of each presented information inference in order to allow the user to remove this information from their profile. The visibility aspect is managed by choosing who can see this check-in, whether friends or just the user. Based on the provided information and controls, the user can decide to abort the check-in completely, apply the chosen controls and check-in, or check-in without any changes.

5.3.2.3 When to Present: Feedback Timing

The timing of displaying privacy alerts is also a key factor in triggering users' awareness of privacy issues. This also contributes to optimal use by the users. Essentially, the privacy indicator is shown when commencing the first step of performing a location-oriented task in the GeoSNs to provide initial reflection of the current privacy level of the user. Any interaction with the application carried out by the user while performing the task is dynamically captured and processed by the privacy notification tool and reflected by the indicator if the privacy level has changed. Therefore, the user is offered partial and real-time notifications of the urgency of the privacy warnings, that enable the use to make instant reactions, as advised in [106].

The privacy indicator to the potential threat level as shown in Figure 5.1(a) is displayed to users whenever they are about to share their location data to give them an initial estimation of their privacy status if they decide to share. However, the actual notification element showing the details of the privacy feedback (see Figure 5.1(b)) is only shown to the users if they click on the privacy indicator, so as not to disturb their experience using the application or enforce the notifications in order to allow them to freely choose to continue their task in hand or explore in details their current privacy status if they feel the need to do so. This approach addresses the difference in mental model of both advance users in security and novice users when considering the privacy consequence of an action as noted in [113]. If users feel that even the privacy indicator might interrupt their use of the application, this tool can be customised to be shown based on information inference or visibility to other users. For example, users can personalise the feedback tool so it just displays privacy indication if the type of information exposed are new and have not been revealed before or if their information is seen by public users.

5.4 Experiment

This experiment is designed to evaluate effect of enhancing users' location awareness using the privacy feedback tool on their perception of privacy and potential behaviour on GeoSNs. In particular, the experiment will aim to explore the impact on the user's perception of privacy due

to providing privacy feedback including a) the presentation of geo-profile content, and b) the use of a privacy threat level indicator. The experiment will also study the impact on a users' behaviour when sharing his location data online due to their perception of privacy, resulting from the introduction of the privacy feedback with and without offering privacy controls. Secondary objectives of the study are to evaluate the proposed design ideas concerning peripheral privacy awareness mechanisms and their impact on the utilisation of the applications.

5.4.1 Method

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 [26, 16, 71, 45]. Using a public GeoSN for evaluation provides more accurate insights in to the general user's attitude to privacy and sharing behaviour than using restricted location-sharing applications (e.g.[4, 36, 37, 38]). Foursquare offers place discovery and recommendation services based on users' location and previous visits to places (check-ins). User's friends have access to her/his place profile, and the user is also able to grant access to other users who visit the same places in her/his profile. A user can also opt out from background location tracking and from behaviourally targeted ads.

The experiment takes the form of an online user study that utilises realistic scenarios of using the Foursquare checking-in application. The scenarios were designed for checking-in to 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. Scenarios were used due to the need to capture users' privacy attitudes and behaviour when presented with potential privacy risks rather than capturing whether they would check-in to the particular places introduced in the scenarios in a real-life situation. Hypothetical requests and scenarios are exploited in order to gain generalisable outcomes in considerable location-sharing studies (e.g. [38, 36, 92]). On the spatial axis, patterns of presence as well as absence from places were used. 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.

Scenarios were designed under two conditions. First, the scenarios are presented with feedback only, and then with actionable controls over the information disclosed. A within-subjects design was chosen since we were interested in capturing the impact of privacy awareness with and without controls, as well as utilising its advantage of reducing error variance associated with individual differences (e.g. [21, 31, 107]).

Perception of privacy is dependent on the user's ability to comprehend the information being

Table 5.3: Summary of the check-in scenarios used.

Scenario No.	Privacy level	Visibility	Dimension	Inferred information
1	Green	Friends	Spacial	None
2	Amber	Public	Spatial	Interests
3		Friends	Spatial-Social	Interests; Friendship relation
4	Red	Public	Spatial-Social-Temporal	Pattern of visit; Place type; Nature of relationship with a friend
5		Public	Spatial-Social-Temporal	Private place (Home); Nature of relationship with a friend
6		Friends	Spatial-Social-Temporal	Pattern of activity; Absence from (Home); Predicted next check-in

disclosed. This study is not intended to measure comprehension, and thus it is 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 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.

5.4.2 Scenarios

Six scenarios were developed reflecting the three privacy threat levels: one Green, two Yellow and three Red, as shown in Table 5.3. Each privacy level is assigned a sufficient number of scenarios to convey the feasible privacy risks that can be associated with this level. The scenarios as well as the screen-shots of the related check-in with the privacy feedback tool are provided in Appendix C.2. Based on the proposed threat level model in Section 5.3.1, we generated a finer specifications for the threat level mapped to the features of Swarm application. These specifications were used to generate the check-in scenarios used in this study. However, we note that the sense of privacy is personal and hence these proposed levelling will be evaluated by the participants in this experiments.

- *Green (Safe):* The check-in is shared only with friends, and information is limited to the current check-in, including place and time, tip or friend tag; no further inference is possible.

- *Amber (Moderate):*

  - The check-in is shared with public (on other social networks), and information is limited to the current check-in, including place and time, tip or friend tag; no further inference is possible.
  - The check-in is shared only with friends, and the users' association with the places, concepts and social ties can be extracted when the frequency of the disclosed spatial-social data is above the threshold.

- *Red (Risky):*
  - The check-in is shared with public, and the users' association with the places, concepts and social ties can be extracted when the frequency of the disclosed spatial-social data is above the threshold.
  - The check-in can be public, and the disclosed spatial-social data constructs temporal patterns which can also be used to derive more private information such as future movement prediction.

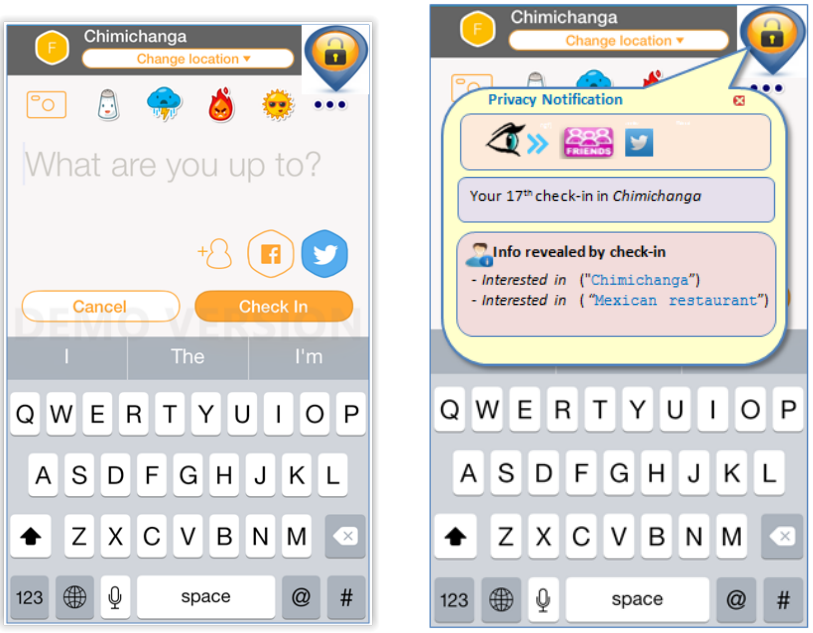
5.4.3 Procedure

The user study is an online survey with four main sections. **Section 1:** the participants' demographics and their experience using social networks are collected. Their location privacy concerns, awareness and behaviour when using them are captured. This is to allow a comparison of these variables to be made after the experiment.

Section 2: the feedback-only scenarios are presented. Six check-in scenarios are presented to the user in random order. 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[2] check-in screen that presents the check-in scenario details and a location privacy icon displayed on the top left corner, as demonstrated in Figure 5.2(a). 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 5.2(b). Users were asked to imagine themselves in such a situation and provide their inputs accordingly. 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 threat level associated with information and, finally, whether they would modify their check-in in any way if given the option to do so.

Section 3: 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

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



(a) (b)

Figure 5.2: Screen-shots of a check-in scenario showing (a) the check-in task details, and (b) privacy awareness pop-up window when clicked on the location privacy icon.

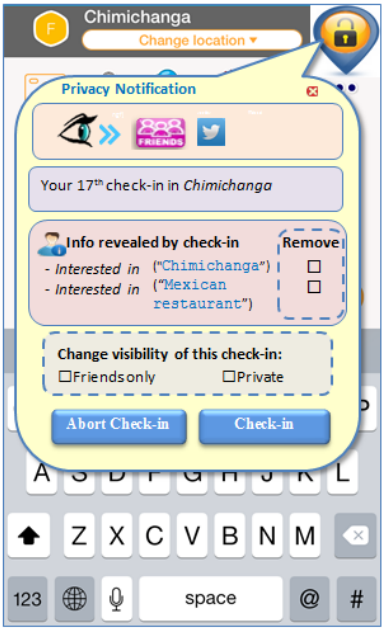


Figure 5.3: A sample of the location privacy awareness notice shown in the feedback and control scenarios.

information elements presented from their geo-profile, to change the visibility of the check-in or opt to abort this check-in all together, as shown in Figure 5.3. Each scenario is followed by a set of questions to measure the effect of the control options on users' location sharing attitude.

Section 4: 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 are examined. Moreover, questions are used to gauge their reaction towards the proposed location awareness tool and its usability. The question types provided in the study include a 5-option Likert Scale of strongly agree/strongly disagree, multiple choice, and open-ended questions for participants to explain their views or elaborate on a subject. The study survey is provided in Appendix C.3.

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

5.4.4 Recruitment and Participants

The experiment was conducted in June 2015. Participants were recruited using Amazon Mechanical Turk (MTurk) and were selected from those who use the Foursquare/Swarm application and check-in frequently (not less than three times a week on average) using a qualification test. 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 were also required to have a 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 $1.50 each. Based on the answers to the demographics and social networking experience question, most of the participants were young people (mean= 30.29, SD= 6.45) and male (57%) were more than female. Most of them were from North America (59.2%)and Asia (34.6%), whereas the rest were from South America (3.5%), Europe (2.4%), and Africa (0.3%).

5.4.5 Tools and Analyses used

Analysis of the survey data and presentation of the results were achieved using R statistical programming language. Statistical tests that are suitable for categorical (nominal) and ordinal data and within-subject (repeated measures) design were carried out using SPSS. Friedman, McNemar-Bowker and Cochran's Q were employed depending on the type and number of the

variables as well as responses to determine the significant differences between correlated user groups. Spearman's rank correlation test is also used to examine the dependence's strength and nature between the ranking of two variables. In addition, Cronbach's Alpha was employed to measure the internal consistency of the scales used in the final section (post-study).

5.5 Findings

An overview of the participants' social networking experience and pre-study privacy concerns is presented first, followed by the analysis results of the check-in scenarios. Finally, a post-study reflection on privacy perception and evaluation of the location awareness tool is given.

5.5.1 Pre-Study Phase: Privacy Concerns, Awareness and Attitude

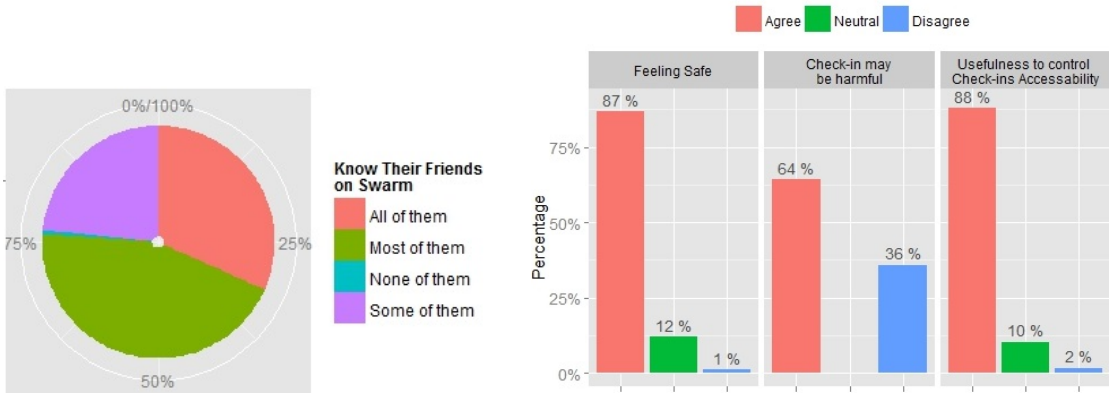
A pre-study evaluation of the participants' privacy concerns with the application was conducted to understand the relationship between their level of experience with the application, their location sharing behaviour and their privacy concerns. Most of the participants were moderate users (check-in several 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). The remaining participants are either not sure how to enable this feature (3%) or they always disable it (2%). Similarly, most of them use social network applications daily (85%), whereas the rest use them occasionally (14%). As expected, most of the participants read the Terms of Use and Privacy Policies only sometimes (50%) or never (21%).

Accessibility to the user's personal data by other users is a primary privacy concern. This is commonly controlled by defining the visibility of one's profile in the privacy settings of 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, and 23.4% know some of them), as shown in Figure 5.4(a).

While most of the participants stated that they currently feel safe using Swarm (87%), 64% of them believe check-ins can be dangerous and want to control others' access to their check-

ins, as presented in Figure 5.4(b). This can denote that the user may misunderstand how privacy implications can occur in terms of the relationship between the application and the consequences of using it. When asking participants to explain their answers, those who thought check-ins could be harmful mentioned the fact that a Swarm friend could be an attacker "If a 'friend' has malicious intent, I can be tracked down and be subject to physical harm. Also, there is always the risk that data will be stolen and is used without my consent". In addition, many participants reported the risk of advertising their absence from home "Anyone who follows you can know where you are at the exact moment. Robbers can know that you are not home.", and the risk of being stalked and spammed "If someone got access to my Swarm check-in data they could easily use it to help them stalk me or to find out about my friends and harass them and me." Furthermore, some of them brought the possibly of inferring personal information and routines "If someone is skipping work/school, etc, this acts as proof they were skipping." and "Check-ins tell a lot about my activities in Swarm.". Others mentioned the threat of access to and exploitation of their data by others "By publicising your location you provide information not only to individuals (whom you may not want to see), but also to companies, who may use that information about your daily patterns/habits to spam you." and "It could be harmful if the data is accessed by third parties and misused if there is any sort of security breach".

As for those who felt that check-ins are not dangerous, they seem to be less aware of the potential privacy risks that may result from the access to and utilisation of their location data. They also tend to trust the application. Their most common justifications include: "They are safe and fast. We can securely check in" and "I believe that they never share my personal things to any third parties or use it for some other purpose".

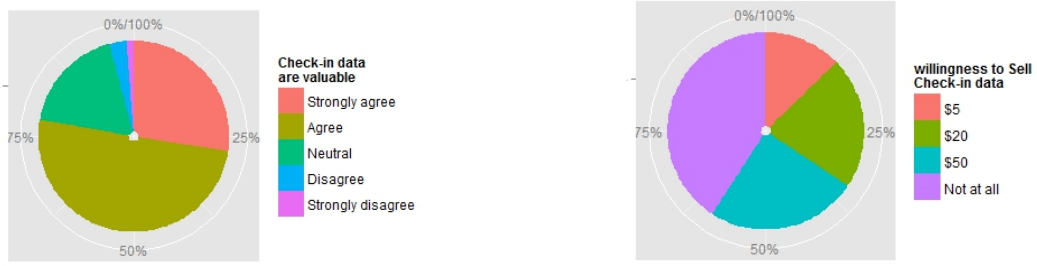


(a) (b)

Figure 5.4: (a) Participants' knowledge of their Swarm friends, (b) Participants' views on location sharing.

Similarly, 78% of participants reported that they value their check-ins information, as shown in Figure 5.5(a), yet 60% of them were actually willing to sell their check-in data for $50

and less, as illustrated in Figure 5.5(b). When examining users' awareness towards their data collection and use as presented in Figure 5.6, 82% of the participants reported that they know who can see their check-in data. Nevertheless, their knowledge boundaries do not make it clear whether they are awareness related to their friends' accessibility only or other parties as well (i.e. other users of the service or third parties). More than half of the participants (59%) know that Foursquare shares their personal data with third-party agencies for targeted-marketing and advertising purposes, whereas fewer (48%) are aware that their data can be used for other purposes. 67% of the participants stated their awareness of Foursquare's collection of their location data even when not using the app.



(a) (b)

Figure 5.5: (a) Participants' view on whether check-ins are valuable , (b) willingness to sell check-in data by participants.

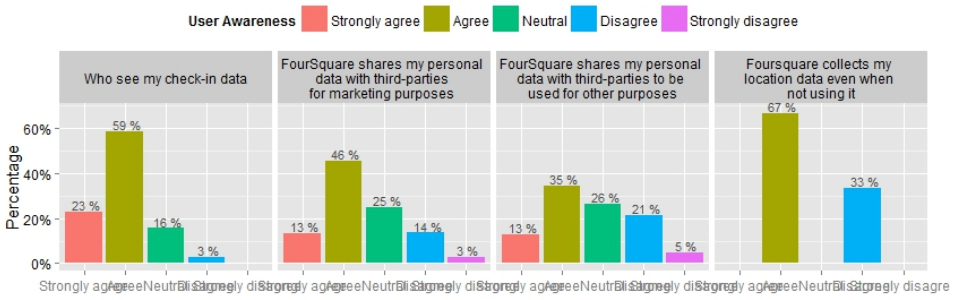
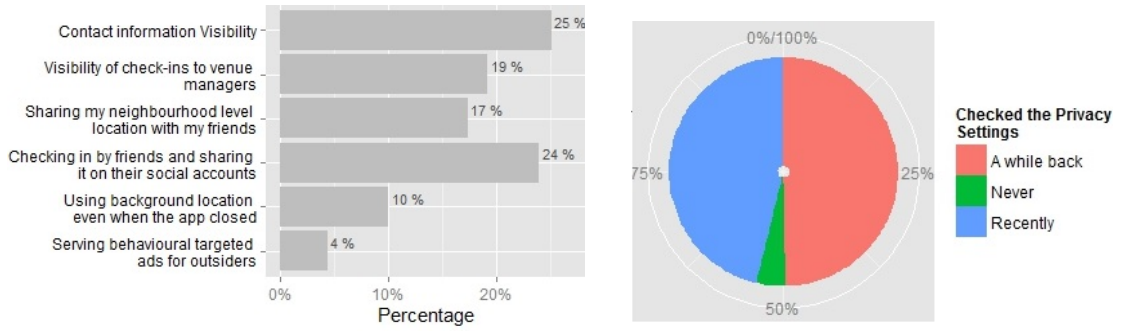


Figure 5.6: Participants' awareness of their data collection and use.

When considering the use of Swarm privacy settings by the participants, 40% have updated only one privacy setting, and 7% have not changed any. The most updated privacy settings, as presented in Figure 5.7(a) are "Who can see my contact information" (25%) and "Enabling my friends to check me in and including my name on their social accounts" (24%). Surprisingly, the least selected ones are "Allowing Swarm to use my background location even when the app closed" (10%) and "Allowing Foursquare to serve behavioural targeted ads for outsiders" (4%). Moreover, the majority of the participants thought that the privacy settings provided were

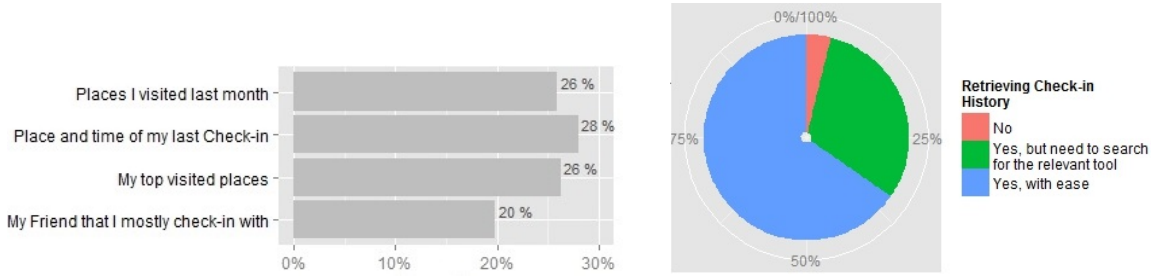
sufficient to protect their privacy (71%), but many (46.15%) also admitted to not checking their privacy settings for a long time as demonstrated in Figure 5.7(b).



(a) (b)

Figure 5.7: (a) The participants' updated Swarm privacy settings , (b) When the participants have checked their privacy settings.

It is interesting to note the seemingly contradictory 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, hence her/his privacy concerns are also limited to this part of her/his data set. This observation is supported by examining responses to the question on which aspects of their location history they are able to recall, where 47% were able to recall only one aspect, and 2.7% remembered nothing of their history. Figure 5.8(a) shows what aspects of location history participants were able to recall. Furthermore, 31% reported that they have difficulty retrieving their check-in history, while 4% are not able to as displayed in Figure 5.8(b).



(a) (b)

Figure 5.8: (a) Aspects of location history participants were able to recall , (b) ease of retrieving check-in history by the participants.

5.5.2 Grouped Analysis of the Check-in Scenarios

In this section, results analyses for all check-in scenarios are grouped and presented based on impacting factors related to the privacy feedback tool. Individual analysis results for each of the scenarios can be found in Appendix C.1

5.5.2.1 Impact of Content on Privacy Perception

5.5.2.1.1 Sufficiency of the Content Provided

Following every scenario, two questions were posed 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 5.9. 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 with the threat level of the check-in scenario (Spearman rank correlation = .245, $p = .000$). Hence, the greater threat the location disclosure poses, the more concerned the participants are about their privacy.

Figure 5.9: Measure of effectiveness, grouped by threat level.

5.5.2.1.2 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 as displayed in Figure 5.9. 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"), Others felt that the privacy settings 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").

5.5.2.2 Impact of Content on User Behaviour

Figure 5.10 demonstrates the effect of content awareness on the attitude of users to modifying their behaviour. 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%).

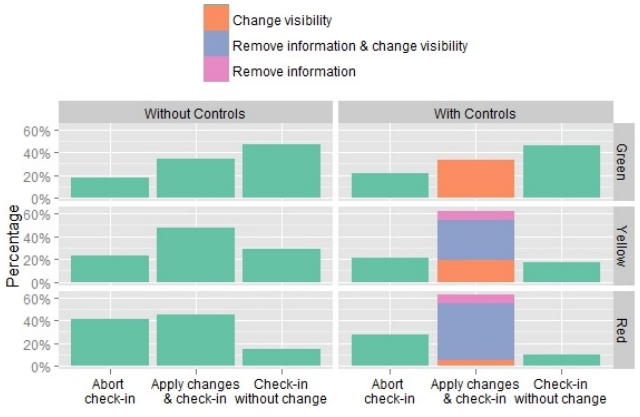


Figure 5.10: Check-in decision with and without privacy controls, grouped by threat level indicator.

5.5.2.2.1 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, the 'proceed with no changes' option was more evident in the Green level scenarios, followed by Amber and Red (representing 47%, 23%, and 12% respectively).

5.5.2.2.2 Privacy Concern and Behaviour

It is useful to observe the impact of the level of privacy concern on the actions participants chose to perform. The participants' check-in behaviour have shown to be significantly influenced by their concern level (Cochran's Q=254.628, $p = .000$), as presented in Figure 5.11. Participants who reported concern about their privacy were the most willing to modify their check-in information or 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 concerns 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 to modify the check-in information or abort the check-in (Spearman rank correlation= .405, $p = .000$).

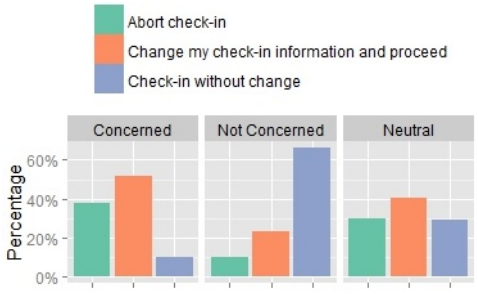


Figure 5.11: Check-in attitude grouped by level of privacy concern.

5.5.2.2.3 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. Scenarios that provide controls were found to more significantly influence the decision of location-sharing action (McNemar-Bowker Test=19.466,

$p = .000$), where 41% (compared to 33%) of participants strongly agree that feedback and control condition were helpful in decision-making compared to the feedback only condition. The difference was more pronounced in the Red and Amber threat level scenarios as shown in Figure 5.12.

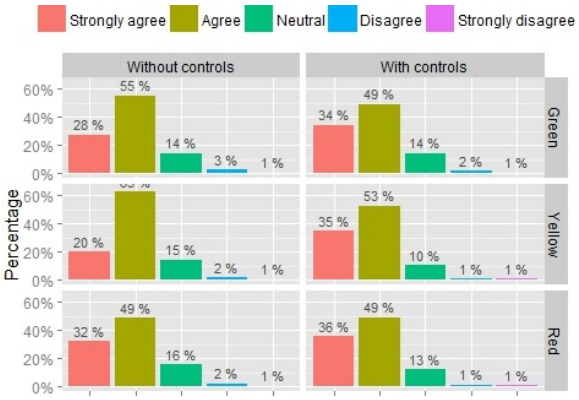


Figure 5.12: Tool support for decision-making based on the availability of privacy controls, grouped by threat level.

5.5.2.2.4 Behaviour Analysis

Figure 5.13 shows the distribution of participants' control actions in relation to particular aspects of their geo-profiles. Most participants decided to remove some of the shared information and to change the visibility of their check-in. The majority chose to remove reference to sensitive places (82%), followed by reference to their interests (77%), their favourite friends (67%), patterns of visit (66%), and current absence from sensitive places (64%).

Overall, 50% of the participants chose to change the visibility of their check-ins when controls were provided, as shown in figure 5.10. Given the options, participants, in general, tend to restrict their check-in visibility to 'friends only', representing 70%, while the remaining 30% would set it to 'private', as displayed in Table 5.4.

Table 5.4: Distribution of participants who chose to modify the scope of visibility in different scenarios..

Privacy level	Scenario No.	Visibility	
		Friends only	Private
Amber (Moderate)	1	77%	23%
Red (Risky)	4	72%	28%
	5	62%	38%

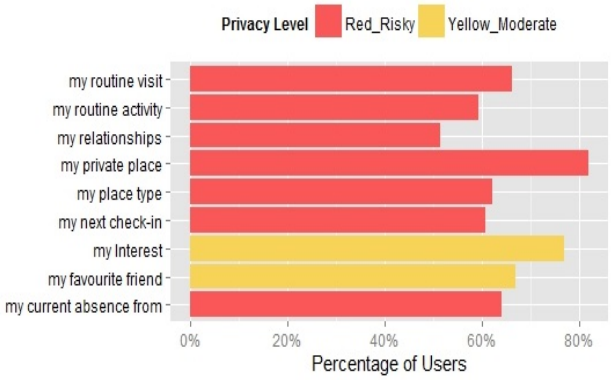


Figure 5.13: Participants' choice of information to remove from their profile, grouped by the check-ins' privacy threat level they were presented in.

5.5.3 Post-Study Phase

5.5.3.1 Location Awareness and Privacy Concern

The overall effect of location awareness on privacy concerns was measured by post-scenario questions (Cronbach's α = .78) and results of which are shown in Figure 5.14. 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 personal information 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 and after 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").

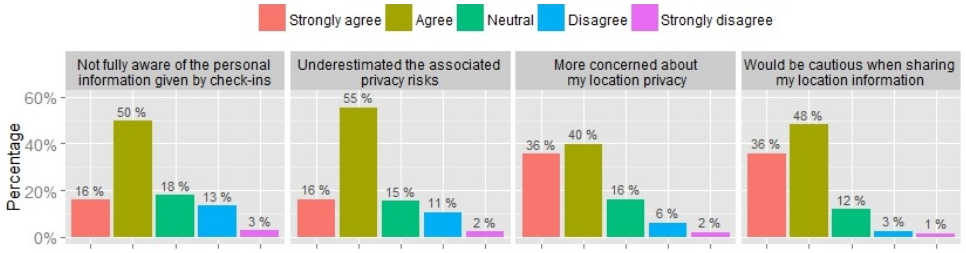


Figure 5.14: General perception of privacy after the check-in scenarios .

Given the user profile that can be created based on their location information, the majority of

participants actually stated their desire to know what information of theirs is collected (81%), who has access to their data (80%), and what their data are used for (81%), as presented in Figure 5.15. The majority were also highly inclined to control these aspects of their data (information collection: 83%, accessibility: 82%, and utilisation: 79%) where they showed stronger agreement towards data controlling than data awareness (45.3% comparing to 35.6%).

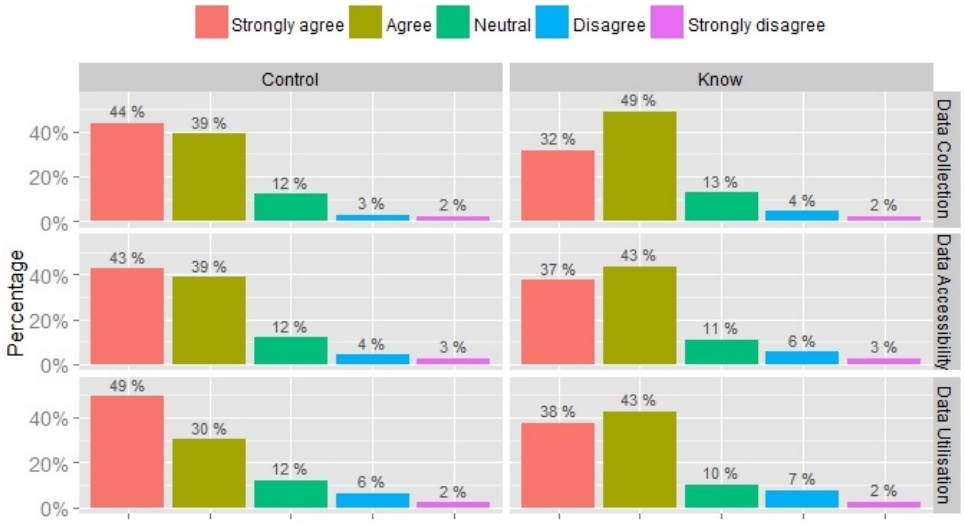


Figure 5.15: Participants' general desire for better knowledge and control of data collection, accessibility, and utilisation.

5.5.3.2 Usability of The Location Awareness Tool

Finally, the overall impression of whether participants consider the tool useful was measured (Cronbach's α = .892). Figure 5.16 gives a summary of the results. This shows the participants' willingness to use such a tool since it allows them to be better informed about their data and manage it as well. Note 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 5.17.

Comments on an open-ended question on their views of 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.", "I really love this tool, and it would make me feel much safer when using Swarm." and "I like the idea of a tool that can help me accurately assess my privacy risks."). Few participants had some suggestions for enhancing the information and controls offer such as "Tip for privacy,

custom alerts when certain conditions are met" and "I would add an option to only make check-ins visible to certain friends (like a checklist)". Others showed their interest in viewing their complete geo-profile ("Show a preview of what it'll look like on my profile and be able to check on and off which you would like to add or remove", and "User should have complete access to the details shared. It should have the options to edit or remove any information that may hamper the user's privacy.").

Figure 5.16: Overall impression of the utility of the location awareness tool .

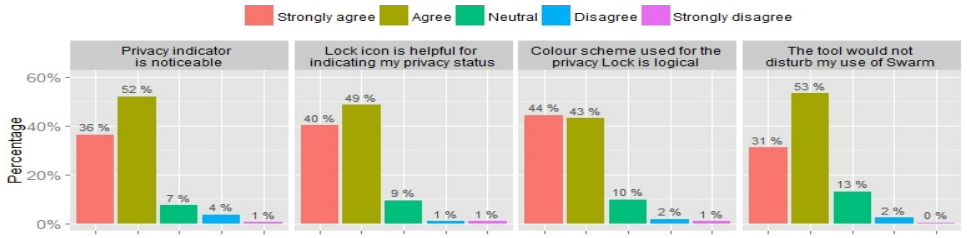


Figure 5.17: Overall evaluation of tool design .

5.6 Discussion

This chapter focusses on enabling location awareness on GeoSNs by providing real-time feedback on related privacy implications of a location-sharing action. The experiment revealed that privacy concerns regarding the sharing of location information are grounded, as users are generally unaware of the extent of the information they are sharing or who is able to access their information. The results also reveal that users trust the applications they use, despite believing that there are risks associated with sharing their location data.

However, as users become more educated about the associated threats to their personal privacy, the results show that they are willing to modify their behaviour significantly, possibly leading to limiting their use of the service (by hiding their profiles) or deserting it altogether (aborting checking-in). This contrasts with Tsai et al.'s [37] findings, where willingness to share increased when users knew who viewed their location, which might be due to the limited awareness offered (accessibility rather than accessibility and content awareness).

It is worth noting that static scenarios of the application were used to gauge users' attitude to privacy, and may have influenced the choices users made in response to the questions asked. However, the aim of this experiment is to capture users' privacy attitudes and behaviour when presented with potential privacy risks resulting from location disclosure rather than capturing whether they would check-in to the particular places introduced in the scenarios in a real-life situation. Hypothetical requests and scenarios are widely utilised to study location privacy attitude and gain generalisable outcomes (e.g. [4, 90, 91, 38, 36, 92]). In addition, the users were initially primed and made aware of the prototypical nature of the test. The proposed level of threat to location privacy model based on logical assumptions is considered preliminary. The purpose of this model is to examine how the use of a privacy level indicator can impact users' privacy attitudes and behaviour. Work on objectively modelling these levels of threat will be carried out in the second stage through investigating unbiased users' location privacy perceptions, as presented in Chapter 6.

Findings from the experiment above can be summarised as recommendations for the design of privacy-sensitive GeoSNs as follows:

- The network needs to provide a full access to the user's geo-profile, allowing the user the opportunity to explore both their captured data and the information inferred from the data. Such accessibility would enable users to understand the link between their sharing actions and potential privacy implications and manage their privacy based on that.

- Users need to be able to remove or modify the contents of their geo-profiles. Such amendment should be treated as if a user deleted a post they shared which should not affect their ability to make use of the application. It is just a way of giving the users the opportunity to hide some of their information that they consider sensitive.

- Users need to be guided on how to optimise their geo-profiles for privacy, i.e. the network service can suggest which aspects of the profile are redundant and may be removed and which aspects are essential for maintaining quality of service.

- A model of privacy threat levels (such as the one presented in Section 5.3.1) needs to be adopted and used to help the user attend to and take appropriate control actions to protect their privacy.

- Users should be told what information from their geo-profile will be shared with third parties and should be given the opportunity to make informed consent on these decisions.

5.7 Conclusion

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

This chapter addresses this problem by a) proposing the design of feedback tools that project a view of the level of threat associated with the disclosure of location information, and b) testing the implications of presenting the feedback on users' perceptions of privacy concerns and their attitude towards sharing their location data on social networks.

The results generally confirm the limitations of users' awareness and their need for more transparent access to their profile content. Results also showed that providing privacy feedback on users' disclosed information and associated privacy risks significantly impacts how they perceive their location privacy, and how they would share their location information accordingly. In addition, it was also clear that users need actionable feedback that allows them to control their content. The 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 GeoSNs.

Chapter 6

Modelling Location Privacy Perceptions in GeoSNs

6.1 Introduction

The previous chapter demonstrated that privacy feedback for location awareness has substantial potential for enhancing location privacy and enabling informed decision-making regarding location disclosure. It offers a preliminary model for mapping the various aspects involved in location-sharing with a representative threat level for aiding users in recognising the privacy related consequences of their actions. In order to provide effective location privacy notification for users, a comprehensive understanding of how users perceive and value the potential privacy risks of their location-sharing activities is necessary. Thus, this chapter aims to gauge users' general perceptions towards location privacy implications to objectively model levels of threat to location privacy, which was initially introduced in Chapter 5. In particular, an in-depth investigation will be carried into users' privacy concerns and sharing behaviour on GeoSNs in regard to three main factors:

- The dimensions of the exposed data including spatial, social and temporal, as presented in Section 5.2.
- Data visibility to other users of the GeoSNs.
- Users' awareness of potential privacy implications resulting from data disclosure.

A user-based online experiment was conducted to objectively capture user abstracted perceptions of location privacy implications. This involved focusing on the data content aspect without intervention from a specific feedback design related to a particular application in order to eliminate any potential bias. Diverse location-sharing scenarios were developed considering the different combinations of the study factors listed above. Moreover, there is no indication to a

possible threat level resulted within the scenarios presented to participants to ensure capturing their unbiased feedback as well .

The results can quantify how data dimensions, visibility and awareness of location disclosure influence perception of privacy. Based on this study outcomes, more robust Location Privacy Threat Levelling Models (LPTLM) that map location-sharing activities to an appropriate threat level (Red, Amber, Green) can be developed. The findings may also be used to provide design insights that allow for better design of privacy notices.

6.2 Experiment

A user-based experiment to examine users' perceptions of implications of location information disclosure was carried out. It aims to capture users' sharing attitude in location-disclosure scenarios that involve different degrees of personal information extraction, information visibility and awareness. Participants' sharing decisions are used as an indicator of their unbiased privacy concerns. The findings can be used to model users' privacy perception towards potential location-based privacy risks in GeoSNs. The study used crowdsourcing means to solicitate a large number of participants who share their location information on social networks.

6.2.1 Study Design

Considering the principles of privacy-oriented experimental design [136], We use between-subjects design (see Section 2.6)to examine the influence of study conditions related to the privacy implication of sharing location data including *data dimension* (Spatial vs Spatial-Social vs Spatial-Social-Temporal) X *visibility scope* (friends only vs public) X *Awareness* (realistic vs attacker's view), resulting in 12 study groups.

Participants are presented with randomly-ordered abstract scenarios that reflect various privacy implications of disclosing location information on GeoSNs, yet they are not mapped to a particular application so as to not bias user' reaction towards a particular location-sharing service and to be able to gauge users' general perceptions. The scenarios are displayed without intervention using a particular design (presentation) style. At the same time, the scenarios offer detailed examples of location-sharing implications based on three aspects: data dimension of the user information, its visibility and awareness of potential privacy implications resulting from data disclosure, to allow the participant to immerse themselves in the scenarios and hence provide more accurate feedback.

The purpose of information sharing with others, in terms of what the information is maybe used for, is not specified in the scenario and will not be examined. The purpose may vary according to a potential adversary . Inferred information can generally be utilised for any purpose whether known or unknown to user. Therefore, we are focusing on how these three study conditions can influence user privacy perception, and leaving the purpose of use up to the participants' expectations.

6.2.1.1 The Data Dimension Aspect

Three primary dimensions of user information on GeoSNs are identified as introduced in Section 5.2 ; 1) the spatial dimension, 2) the social dimension, and 3) the temporal dimension. These dimensions determine the type and volume of information that can be inferred, and hence affect users' attitude to privacy.

In addition, the sensitivity of the place and ultimately the sensitive of the extracted information derived from it varies which also might influence users' perceptions of privacy. Thus, we also considered sensitivity as a sub-factor within the data dimensions in the experiment design. In particular, sensitive information targeted in this study involves knowing user's hobbies, religion and beliefs, political views, sexual life, physical or mental health, ethnic origin, and any offence by location data (California Location Privacy Act of 2012 [1] and the UK Data Protection Act [2]). It also includes inferring home location and the user's absence from it [137]. Moreover, being with a friend at a place, as in the spatial dimension, can also impact the user's privacy attitude, hence this is considered a sub-factor as well. For example, a user might not want to share their location to protect the privacy of the friend they are with.

Therefore, we developed scenarios to reflect three ranges of data dimension that increase in complexity and hence potential risk to privacy. These are spatial only, spatial-social, and spatial-social-temporal. The three dimension include scenarios that vary in a) sensitivity of the data shared and b)whether th user are with a friend (if the social dimension is involved). A detailed discussion of the scenario specifications is presented Section 6.2.3.

[1] `https://www.eff.org/cases/california-location-privacy-act-2012` [Accessed: 26-Apr-2016]

[2] `http://www.legislation.gov.uk/ukpga/1998/29/section/2` [Accessed: 26-Apr-2016]

6.2.1.2 The Visibility Aspect

The visibility of user data in terms of who can access it is an important factor that impacts users' privacy attitude. We chose two basic visibility scopes to incorporate in the study scenarios: friends and public. Friends refers to users' social connection in their GeoSNs such as followers and friends who can be known by the user to some extent. The public scope indicates any other users whether end users of the service or the application (service providers) and third party companies who can use more sophisticated systematic techniques for linking and deriving users' personal information.

6.2.1.3 The Awareness Aspect

Users' awareness of the implications of sharing location information contributes to how they perceive their location privacy. In particular, their sense of privacy may differ when they are aware of what a potential member who is granted visibility to user's profile can actually see. In order to examine how awareness can influence user privacy perception, we presented the scenarios in two settings that address user awareness:

1) The first setting is the Realistic Awareness scenarios , where they show only the context a user is in, such as where they are, how many times they have visited the place, and who they are currently with. In this setting, we aim to capture user sharing decisions and hence privacy perceptions during realistic location sharing scenarios in GeoSNs. The context provided to users reflects their realistic experience and consciousness they would have in mind while using the social networking application in a real-world situation. Users can set the visibility of their profile to a group of users, where all their subsequent posts are automatically seen by this group.

2) The second setting is named the Attacker's View Awareness, where the same scenarios as the first setting are presented in addition to what of their information can be revealed and seen based on their location history. and by which group of users (including potential attackers). This setting aims to capture user sharing decisions and hence privacy perceptions when they are actually aware of what of their information can be exposed to a particular group of users before they make their location-sharing decisions. The users' revealed information basically mainly includes disclosing their frequency of visiting a place, home location and absence from it, co-location and nature of relationship with

friends, and mobility patterns, which can be used to extract other information, such as movement prediction.

6.2.2 Procedure

The experiment begins by presenting to participants hypothetical location-sharing scenarios that vary in study conditions. Participants are assigned to one of 12 study treatments/ groups, where each group incorporates a unique combination of study conditions, including *data dimension*, *visibility scope* (to friends or other users of the application) and *awareness*. The number of participants in each treatment is shown in Table 6.1. Each treatment includes 9 to 10 location-sharing scenarios that vary in sensitivity and presence with a friend . The details of the scenarios are discussed in the next section.

The scenarios are presented where the information displayed in the scenario can be known by the user but not yet shared within the selected visibility scope. We opted to capture users' privacy perceptions indirectly by assessing their willingness to share this information with the chosen visibility group using three options: 'yes', 'maybe' and 'no'. In this way, user reaction is unbiased, since they report about the sharing preferences without any indication of privacy issues. Then, we map 'yes' to being unconcerned, 'maybe' to being concerned, and 'no' to being very concerned, and model their privacy perceptions across all study aspects based on this mapping. The scenarios in each treatment were randomly ordered and presented to each individual participant in order to further ensure that there is no carryover effect in participant feedback. We also balanced the number of sensitive and intensive scenarios in each of the study treatments so as not to skew the results.

After providing their feedback on the scenarios, the participants were presented with an open-ended question asking them to justify why they selected their sharing decision as 'no' or 'maybe' in order to further investigate their reasoning behind their decisions. Then, participants' demographics were collected as well as their social networking experience, including which GeoSNs they use and whether they tag their friends when sharing their locations. Finally, we gauged how they value their online privacy by introducing six statements, where some are positively framed and others are negatively framed, as a way of avoiding any potential influence. These statements measure participants' privacy preferences using a 5-point Likert scale (strongly disagree/ strongly agree) which were selected from published privacy scales used by [29]. We negatively framed two of them to make sure that their inputs are not biased. We created an index variable for the privacy preference in order to show the general sense of privacy for the participants . This was done by reverse-coding the negatively framed statements (i.e. agreeing is considered disagreeing) and then taking the average of the six statements. The mean of the privacy pref-

Table 6.1: Number of participants in each combination of study conditions. Independent variables are Data Dimension (Spatial, Spatial-Social, or Spatial-Social-Temporal), Awareness (Realistic or Attacker's View), and Visibility (Friends or Public)..

Study Aspects	Realistic		Attacker's View	
	Friends	Public	Friends	Public
Spatial	59	59	58	59
Spatial-Social	60	60	59	61
Spatial-Social-Temporal	58	59	62	61

Table 6.2: Participants' preferences in the presented privacy statements..

Statement	Average	SD
I am not concerned that companies are collecting too much personal information about me. REVERSE CODED	2.456	1.2
It usually does not bother me when companies ask me for personal information. REVERSE CODED	2.641	1.17
When people give personal information to a company for some reason, the company should never use the information for any other reason.	4.357	0.92
I have limited the personal information that I post to my social networks accounts.	3.869	1
I don't post to my social networks accounts about certain topics because I worry who has access.	3.807	1.14
If I think that information (including location) I posted to my social networks accounts really looks too private, I might delete it.	4.134	1.04

erences variable was 3.845 (Cronbach's α=0.703), which means that the participants actually place importance on their online privacy. Table 6.2 shows participants' preferences towards the presented privacy statements.

A pilot study was carried out using five volunteers who were users of GeoSNs in May 2016. They were introduced with the complete study scenarios and questions. The aim of the pilot was to ensure that the study is comprehensible and clear before it is disseminated. The volunteers were asked to fill in five feedback question after completion of the study. All of the participants said that the study was easy to understand and the scenarios were easy to follow. For example, they said that " It is easy to understand and clear" and they "had no trouble following the scenarios'. However, they did suggest to improve the wording of some scenarios and their descriptions. They were mainly confused by the nature of the relationship between the two friends examples presented in the scenarios, and they suggested to make it clearer in the

introduction to the scenarios.

6.2.3 Scenarios

The scenarios were developed based on combinations of the three conditions of the study: the data dimension (Spatial, Spatial-Social and Spatial-Social-Temporal), visibility (Friends and Public) and awareness (Realistic and Attackers' view). Each of the 12 treatment groups has a different value for each of the study conditions . The following explains how the study conditions are addressed in the scenarios:

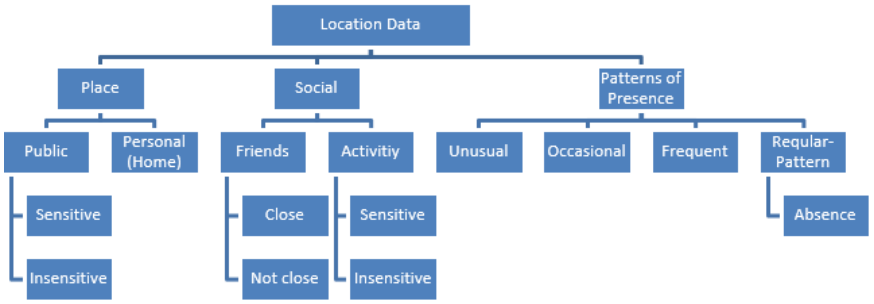


Figure 6.1: The data-dimension taxonomy that shows the information presented in a location-sharing scenario .

To understand how the aspect of data dimension is represented in these scenarios, we developed a taxonomy that contains the elements and their attributes that should be included, as shown in Figure 6.1. Based on this taxonomy, we generated the scenario specifications that should be covered in each of the data dimension combinations:

- In the Spatial only scenarios, all the Place attributes shown in the taxonomy are considered. The Pattern of Presence should be to set to Unusual or Occasional only (see Figure 6.1) where a place can be visited a few times without any particular association with time period.
- In the Spatial-Social scenarios, all the Place and the Social attributes are considered. The Pattern of Presence can be Unusual, Occasional or Frequent (Frequent Pattern of Presence scenarios represent close Friends and favourite activities and less Frequent Pattern of Presence scenarios represent not-close Friends and not-favourite activities).
- In the Spatial-Social-Temporal scenarios, all the Place attributes are considered. The Pattern of Presence can be Regular Patterns, Absence or Unusual in terms of detecting a change in user patters, as shown in Figure 6.1. In terms of the Social element, any activity can be included, and friends can only be those who have a close relationship with the user since it is unrealistic to go out routinely with someone who is not close with the user.

In this study, we selected religion, political views, health and ethnic origins as representatives of sensitive places from those presented in Section 6.2.1.1. Any other public place is considered as non-sensitive. Personal place basically refers to home or residential locations. The treatment including the Social dimension presents scenarios for being with a friend at a place where we also differentiate the closeness of the friend to the user. We want to examine if participant reactions would differ based on this. Therefore, two examples of friends were introduced: Alex as a close friend in real life and on the social network application, and Jack as not a close friend in real life, but is a friend on the social network application. Below are examples of scenarios presented that correspond to the three data dimension combinations:

- **Spatial (Insensitive):** You are in the *Central Shopping Mall* in Town. You are looking for a new pair of shoes. You may have been here occasionally in the past.
- **Spatial-Social (Sensitive with a friend):** You are now in the *Main Hospital* in Town *with Alex*. You are both visiting a friend. You have visited this hospital only once last year.
- **Spatial-Social-Temporal (Personal with a friend):** You are *with Alex* at your *home* at 16 Park Place (an apartment building). He normally visits on Sunday Evenings.

The scenarios also differ based on the awareness aspect. Realistic scenarios show only the context a user is in. However, in the case of the Attackers' View, the same scenarios in the Realistic condition are presented in addition to mentioning what of their information can be revealed based on their location history and seen by the set visibility group . Both awareness types involve all of the data dimension combinations shown in the later examples, as well as their sub-factors in terms of sensitivity and being with a friend. The following are examples of the same scenario when having different awareness in both visibility groups:

- **Spatial-Social-Temporal/Realistic/ Friends or Public:** It is now Saturday evening and you are in the Good Life Pub in Town. You regularly go there on weekends.
- **Spatial-Social-Temporal/Attackers' View/ Friends):** It is now Saturday evening and you are in the Good Life Pub in Town. You regularly go there on weekends. If you share your location track, your *friend connections* will be able to see that you regularly go to this pub on Weekends.
- **Spatial-Social-Temporal/Attackers' View/ Public):** It is now Saturday evening and you are in the Good Life Pub in Town. You regularly go there on weekends. If you share your location track, *other users of the application* will be able to see that you regularly go to this pub on Weekends.

The participants are asked to imagine themselves in the given scenarios where they are using a social networking application that records their location history and interactions with others and their profile visibility is set to one of the visibility scopes (Friends or Public). Capturing users' privacy attitude towards location information is carried out by asking them about their willingness to share their location given the presented location-sharing scenarios so as not to bias their responses by triggering them to think about their privacy. Participants are offered three levels of sharing decisions, namely 'yes' which is mapped to being unconcerned, 'maybe' which is mapped to being concerned and 'no' which is mapped to being very concerned. Here are examples of the type of questions asked in the data dimension combination group considering the different awareness and visibility groups whether with a friend or not:

- **Friends (alone)/ Realistic:** Would you share your location now with friends?
- **Friends (being with a friend)/ Realistic:** Would you share your location now with friends and tag Alex as well?
- **Friends (alone)/ Attackers' View:** Would you share your location *track* now with friends?
- **Friends (being with a friend)/ Attackers' View:** Would you share your location *track* now with friends and tag Alex as well?
- **Public (alone)/ Realistic:** Would you share your location now with other users?
- **Public (being with a friend)/ Realistic:** Would you share your location now with other users and tag Alex as well?
- **Public (alone)/ Attackers' View:** Would you share your location *track* now with other users?
- **Public (being with a friend)/ Attackers' View:** Would you share your location *track* now with other users and tag Alex as well?

Under Realistic awareness scenarios, participants are asked whether they would share their current location to mimic real-world experience of using GeoSNs. However, in the Attackers' View scenarios, they are asked about their willingness to share their whole location track in order to capture their reaction to the privacy implication of location disclosure, since the scenarios indicate what of their personal location information can be known if they choose to disclose their track. The only difference in the visibility condition is that participants are asked whether they would share their location with either Friends or Other users of the application (Public). The scenarios and questions related are provided in Appendix D.

6.2.4 Recruitment and Participants

The experiment was conducted in June 2016. Participants were recruited using Amazon Mechanical Turk (MTurk) and were limited to those who are users of social networking applications and have shared their location from their social network accounts. 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 location sharing decisions. A qualification test was introduced by the start of the study to determine whether the workers meet the qualifications required for the study and can proceed to complete it, or they do not, and thus automatically aborts them.

The MTurk workers must also have a 95% or higher approval rate for at least 500 tasks to be able to participate. This is to make sure that they provide valid feedback according to the study instructions. We also ensured that that any eligible worker can participant only once in any of the 12 versions of the study by using a MTurk mechanism. In particular, we assigned a qualification to those who completed the study and then specified in our MTurk task that any worker who has this qualification cannot take part. The various study treatments' surveys were disseminated in different time periods throughout the day to ensure that any eligible worker from any county can participant in our study in order to maintain the representativeness of our sample.

747 participants took part in the study, and 32 did not meet the qualification of sharing location information with others on social networks. The remaining 715 participants were able to complete the survey in an average of 6.14 minutes. The sample was young (Mean= 33.35 years old, SD= 9.88), and the proportion of males and females was equal (50% each). Most were from North America (72.31%), Asia (18.04%), and Europe (6.15%).

As for their social networking experience, the majority use social network applications frequently (several times a day) (69.79%). The most used platform for sharing location information is Facebook, followed by Twitter, Instagram and Google+, representing 95%, 55%, 53% and 41% of participants, respectively, as demonstrated in Figure 6.2. In addition, most participants tag their friends when sharing location information on Social Networks (always: 15.66%, sometimes: 78.74%). 22.66% of participants enable "Location Services" or other similar location features on your mobile applications frequently (always on), 70.35% enable them moderately (when required by an application), while only 4.48% disable such features.

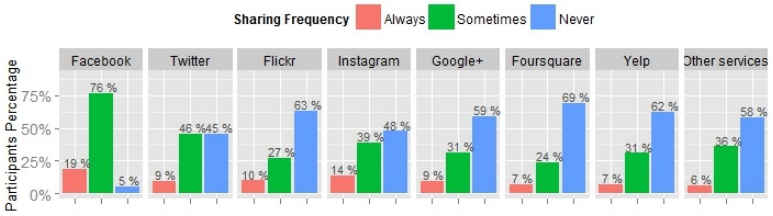


Figure 6.2: Participants' location sharing attitude on GeoSNs .

Table 6.3: Percentage of selected sharing decisions (Yes, Maybe or No) in each conditions group.

		Realistic			**Attackers' View**		
		Yes	Maybe	No	Yes	Maybe	No
Spatial	**Friends**	*56.10%*	19.49%	24.41%	42.59%	24.83%	32.59%
	Public	45.25%	27.29%	27.46%	40.34%	28.98%	30.68%
Spatial-Social	**Friends**	47.65%	26.74%	25.61%	46.67%	26.85%	26.48%
	Public	40.56%	26.11%	33.33%	36.07%	30.97%	32.97%
Spatial-Social-Temporal	**Friends**	50.09%	34.61%	28.23%	44.81%	30.60%	30.24%
	Public	41.17%	29.87%	36.43%	37.70%	23.13%	*50.27%*

6.3 Results

Analysis of the survey results including data manipulation and presentation were achieved using the R statistical programming language. SPSS was also used to carry out statistical tests. Three main independent/predictor variables represent the study conditions: Data Dimension (Spatial, Spatial-Social, or Spatial-Social-Temporal), Awareness (Realistic or Attacker's View), and Visibility (Friends or Public), as well as one dependent/outcome variable representing the participants' location-sharing decision (yes, maybe, no), which indirectly indicates their privacy attitude. Table 6.3 and Figure 6.3 demonstrate the percentage of selected sharing decisions (Yes, Maybe or No) in each conditions group.

The Chi-square test of independence was used for examining the impact of the study conditions on users' privacy attitudes. Spearman's Rank-Order Correlation was also used to determine the strength and direction of the correlation (if any) between the study condition and users' perceptions. To explore the relationship between the different levels of the study conditions on users' privacy attitude, we opted for an ordinal logistic regression model, where the levels of the outcome variables is coded as yes=1, maybe=2, and no=3. To interpret the regression results, positive coefficients (>0) indicate greater likelihood of willingness to share location (not concerned) ; coefficients equal to 0 mean no additional likelihood on top of the baseline, and negative coefficients (<0) indicate lower willingness to share (higher likelihood of being concerned). The results from the model are presented in Table 6.4.

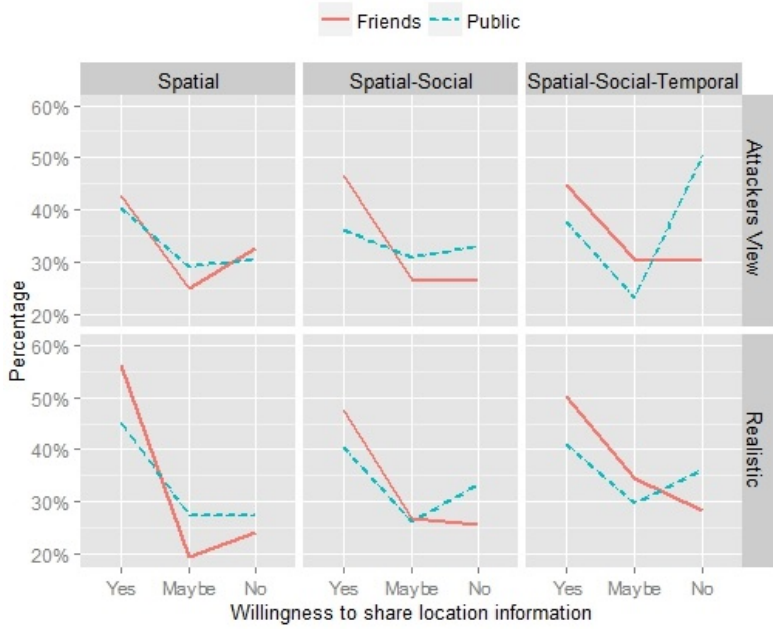


Figure 6.3: Percentage of selected sharing decisions (Yes, Maybe or No) in each conditions group.

Table 6.4: Results of ordinal logistic regression model.

Condition	Estimate	Odds ratio	Std. Error	P(Sig.)	95% Confidence Interval	
					Lower Bound	Upper Bound
Data Dimension (bassline= Spatial-Social-Temporal)						
Dimensions=Spatial	.252	1.287	.054	<.0001	.147	.358
Dimensions=Spatial-Social	.151	1.163	.055	.006	.043	.259
Visibility (bassline=Public)						
Visibility=Friends	.320	1.378	.045	<.0001	.233	.408
Awareness (bassline=Realistic)						
Awareness=Attackers' View	-.217	.805	.045	<.0001	-.305	-.129

As expected, the data dimensions, visibility and awareness of location-sharing activities showed to significantly impact participants' privacy perceptions (p<.0001). We found positive correlations between being concerned and having more privacy-threatening scenarios. In particular, the visibility aspect has the strongest impact on privacy perception followed by the awareness then the data dimension. Interestingly, sub-factors including place sensitivity and whether being with a friend at a place also showed to significantly influence privacy attitudes, where the sensitivity of a place reduces a user's willingness to share by 31%, and presence of a friend reduces it by 8%. What follows are several analyses that better demonstrate how participants'

perceptions vary based on location-sharing factors.

6.3.1 Impact of Data Dimension

The data dimensions of the derived information in location-sharing scenarios has statistically significant impact on users' willingness to share (Pearson Chi-Square= 22.72, p<0.0001). In particular, there is a moderate positive correlation between the participants' sharing attitude and the data dimension of the extracted information (Spearman's rho=0.53, p<0.0001), which means having more data dimensions involved in a given location-sharing action increases users' concern about location privacy and reduces their willingness to share.

When examining the relationship between user perceptions across all data dimension combinations using the logistic regression, participants presented with Spatial scenarios were most likely to share their location, and those presented with Spatial-Social scenarios were more likely to share their location than participants in Spatial-Social-Temporal scenarios, as expected. This finding suggests that participants exposed to scenarios that reveal more data dimensions, and hence personal information, were more likely to express privacy concerns than those in less revealing dimension scenarios. This is indicated in participants' justifications of why they were reluctant to share their location, where most of them refer to privacy and safety issues, saying, for example, "I may not want people to know all of my personal business about where I am and what I'm doing" and "I wouldn't share my location in any place where my personal privacy or safety concerns were an issue".

Participants in the Spatial scenarios are the least concerned about their privacy (Maybe 25.1%:, No: 28.8%), whereas privacy concerns increase for the Spatial-Social scenarios (Maybe 27.7%:, No: 29.6%) and the highest concern rate is shown in Spatial-Social-Temporal scenarios (Maybe 27%:, No: 33.2%). Participants who answered maybe or no to locations sharing in Spatial-Social scenarios were more specific in explaining their choices by indicating particular concern with identifying places that are frequently visited, where they said "I would not share anything too personal, like places I frequent that are in my neighbourhood" and "If it's a place that I commonly go, then I would not share". Similarly, participants who refused or hesitated to share locations in Spatial-Social-Temporal scenarios justified their attitude by referring to their concerns about others finding their movements patterns and the consequences of such inference, saying, for example, "I am concerned about my safety. Some people could see the pattern of my whereabouts, and use that information to stalk me or my friend", "If things are routine, that is giving someone your schedule and they could track you and get to you if they wanted to", and "Because there are some places you just do not need to let others know where the location is. These days people could try to come to your home and rape you, murder you, or even kidnap

you". Some were worried about their privacy regarding absence from their personal places, such as their home, as they explained, "Sharing my location for places I visit on a regular basis advertises that I am not at home on those days and times" and "I would not let anyone know where I go on a regular basis. This is a good way to get your home robbed".

6.3.1.1 Sensitivity of Location-sharing Scenarios

The sensitivity factor of the location as well as what information it reveals was considered when designed the location-sharing scenarios, as this has the potential to influence how people react toward sharing a given location. Locations that are categorised as sensitive are based on privacy acts, and those referred to as personal mainly include home locations (as discussed in the data dimension aspect). Generally, observing the sensitivity of location and the potentially inferred personal information across all data dimension combinations has a statistically significant effect on users' willingness to share (Pearson Chi-Square= 729.903, $p<0.0001$). A positive correlation was found between the participants' attitude level and the sensitivity of the derived information (Spearman's rho=0.312, $p<0.0001$). This implies that users' location sharing decision and hence location privacy concerns are influenced by the sensitivity of the place or associated information. As expected, participants are the most likely to share their location in non-sensitive scenarios (Yes: 58.4%, Maybe: 27%, No: 14.6%).

However, users' willingness to share decreases significantly (by 31%) when sensitivity increases, indicating that participants are the most concerned about their privacy and least likely to share their personal locations (i.e. their home) (Yes: 27.4%, Maybe: 21.9%, No: 50.7%), followed by their sensitive locations (Yes: 33.2%, Maybe: 27.7%, No: 39.1%). Participants' justification for their reluctance to share also suggested that they are willing to share public or interesting locations but not private or sensitive ones. They explained this saying "Sharing location information for public places is mostly OK with me but sharing personal location information related to religion, political affiliation or a friend's house via location info is something I try not to do" and "I really wouldn't want to share medical location places or anything having to do with my culture, faith or home. Those are private issues that I do not mention on social networks". Several participants mentioned that they would hesitate to share their locations if "they're boring" or "they are not of interest to most people". Observing participants' justifications shows that they consider the location sensitivity in the first place, and some also take into account the interestingness or usefulness factors of the location to others.

Examining the participants sharing attitude individually towards the categories of sensitive information showed that the top two information categories that participants both hesitated and denied sharing are health (Maybe:22%, No:55%) followed by home(Maybe:27%, No: 51%) as

illustrated in Figure 6.4. The third and fourth categories that participants refused to share are religion and political views (38% and 35%, respectively).

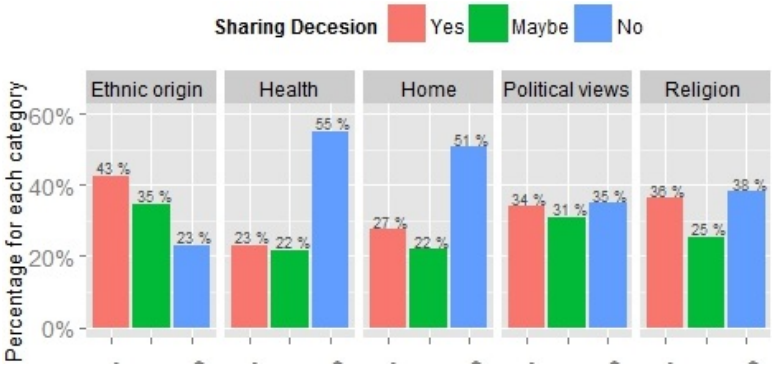


Figure 6.4: Sharing decisions for sensitive and personal places .

6.3.1.2 Sharing Behaviour When being with friends

Whether a user is with a friend can also influence their attitude toward location sharing. The possibility of a friend's presence in a place with a user is a sub-factor of the social dimension. We captured its impact by presenting approximately the same number of scenarios where 1) a user is alone at a place, and 2) being with a friend whenever the social dimension is involved . An overview of participants' sharing decisions when at a place with a friend, across all data dimension combinations, showed that it has a statistically significant impact on their willingness to share the location that their friends are tagged in, and hence their privacy perceptions (Pearson Chi-Square= 46.363, $p<0.0001$). This association has a positive correlation between sharing attitude and whether the user is with a friend (Spearman's rho=0.082, $p<0.0001$). Participants are less likely to share their location if they are with a friend than when at a place themselves (38% said yes when with a friend, compared to 46% when alone).

The reasons for this difference in attitude are explained by participants to be mainly because they think the information is sensitive and needs to be protected, or they have to seek their friend's permission before sharing, as they said, "To protect the privacy of other people I was with or visiting" and "If I do tag friends, I like to ask their permission first". Examining participants' willingness to share their location when with a close friend or just a friend from a social network connection, revealed that it is significantly associated with participants' sharing attitude (Pearson Chi-Square= 1.04.255, $p<0.0001$). Participants are less willing to share their location when they are with a close friend (No: 38%) than when with an acquaintance (No: 21%), as demonstrated in Figure 7.7. This might be because a user can be with a close friend in any Spatial-Social and Spatial-Social-temporal scenario whether the visit to a place is occasional or

routine. Hence, the scenario can be more privacy-threatening, which makes the participant less willing to share. It can also be because participants want to protect the privacy of their close friend. Interestingly, some participants who selected not to share their location with a friend from a social network justified their choice by saying, “Me along with my close friend, location sharing is ok, but there is no point to share the information of me with an unknown guy to others” and “I don’t know Jack that much, so tagging him would be weird”.

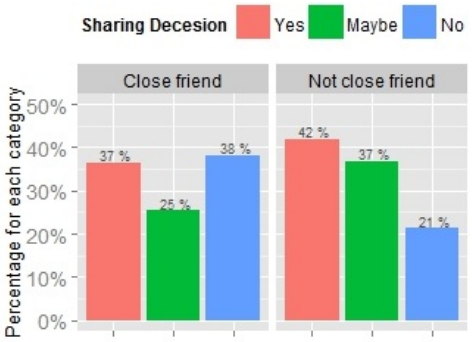


Figure 6.5: Sharing decisions when with a friend categorised by their closeness to user.

6.3.2 Impact of Visibility Scope

The visibility aspect of the information derived in location-sharing scenarios has a statistically significant impact on users’ likelihood to disclose their location (Pearson Chi-Square= 50.204, $p<0.0001$). In fact, there is a very strong positive correlation between the participants’ privacy attitude and the visibility of the extracted location-based information (Spearman’s rho=0.85, $p<0.0001$). This suggests that participants are less likely to share their location with other users of a social network (public) than with their friends. Moreover, observing the relationship between the user’s privacy perception across visibility scopes using the logistic regression, participants who were asked whether they would share location with friends were more likely to share than those who were asked about sharing with the public as shown in Table 6.4. This implies that disclosing location-based information with the public triggers higher privacy concerns (34% said no to sharing with public compared to only 27% with friends). Participants who were reluctant to disclose their location with other users justified their attitude by referring to their desire to protect their privacy from unknown people, saying, “I didn’t want to disclose my location for strangers to know, where I was”, “I don’t want people to know where I am or potentially stalk me”, and “Some occasions seem too personal to share with the public”.

Considering the data dimension and visibility conditions shows that these two combined also have an effect on participants’ attitude to disclosing their location information in GeoSNs. Generally, the least concerned participants who are willing to share their locations are those who

were presented with Spatial only and sharing with friends scenarios (Yes: 49.34%), while the group who are the most concerned about their location privacy are the participant in Spatial-Social-Temporal scenarios to be shared with the public followed by those in Spatial-Social scenarios to be shared with the public (No: 43.35% and 33.15% respectively), as demonstrated in Figure 6.6. The visibility aspect seems to trigger more privacy concerns than the data dimension aspect where 29% of the participants refused to share their location with the public in Spatial only scenarios, compared to only 27% who denied sharing with friends in Spatial-Social-Temporal scenarios. It seems that participants were able to recognise the more privacy-threatening location-sharing scenarios and made their sharing decision based upon, saying, for example "I am not comfortable with strangers having access to my address and access to my routines ".

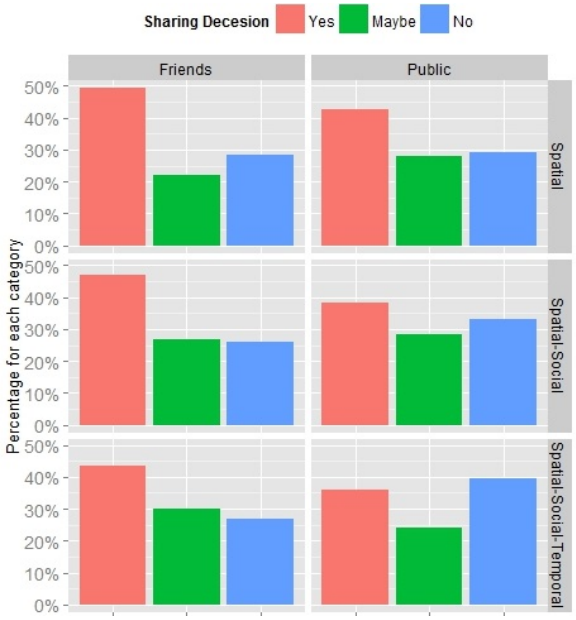


Figure 6.6: Percentage of selected sharing decisions (Yes, Maybe or No) categorised by visibility and data dimensions conditions .

6.3.3 Impact of Awareness

Users awareness of location disclosure implications has a statistically significant impact on users' willingness to share (Pearson Chi-Square= 23.340, p<0.0001). This association shows a moderate to strong positive correlation between the participants' attitude and the visibility of the inferred information (Spearman's rho=0.58, p<0.0001). This means that participants who had the attackers' view on location-sharing scenarios were less likely to disclose their locations than those presented with realistic scenarios as shown in the regression results Table 6.4.

Participants' justification for refusing to share their location or a hesitation to do so indicated

three main issues. The first is unwanted access to their location. For instance, they said “I would be concerned with who is seeing the post” and “I wouldn’t want to broadcast my history with the place” . The second cause is undesired tracking of their movement, for example, “I didn’t want to be tracked in sensitive areas” and “I don’t want to be tracked and I don’t want someone to notice patterns to where I go”. The last reason regarded their concerns about personal information inferences. They explained saying “I prefer to keep certain things private - politics, health info, and any other information that could be used to deduce other things I prefer to keep private”, “The location mentioned was of a sensitive and personal nature. This was either due to the nature of the location itself, or my reason for being there which others may be able to guess or speculate about based on the location”, and “Sharing some locations would allow other users (who I may not want to share that data with) to interpret or assume things about me that I would not necessarily want to be public knowledge”.

Considering data dimension condition to examine users’ privacy perception with the awareness condition showed that participants are most willing to disclose their location in spatial only and realistic scenarios which are the least privacy-breaching (Yes:50.68%). Generally, participants with awareness of the accessibility of information revealed are less willing to disclose their location comparing to those in realistic view among all data dimensions combinations, as illustrated in Figure 6.7, which can suggest that awareness has a stronger impact than the data dimensions revealed. Participants are the most concerned about their privacy (deny sharing) when presented with Spatial-Social-Temporal scenarios from the attackers’ view (No: 37%).

Figure 6.7: Percentage of selected Sharing decisions (Yes, Maybe or No) categorised by data dimensions and awareness conditions .

The visibility condition is also included in with the awareness aspect in this examination. The general trend shows that visibility influences concerns about the awareness aspect, where the participants are the least willing to share their location with the public when witnessing the attackers’ view, this is followed by realistic awareness (No: 36% and 32%, respectively), as presented in Figure 6.8. Exploration of the relationships between all study conditions in terms of their mutual impact on users’ privacy perceptions reveals that the participants most likely to share their location are those presented with Spatial:Friends:Realistic scenarios (Yes: 56.10%),

followed by those in Spatial-Social-Temporal: Friends: Realistic and Spatial-Social: Friends: Realistic (Yes:50% and 48% respectively), as shown in Figure 6.3. This observation suggests that participants were least concerned about their location privacy when sharing with friends in realistic experience and the extracted information is at minimum (data dimension=Spatial only), and their concern increases when more personal information are revealed. On the contrary, participants denied location disclosure the most in Spatial-Social-Temporal: Public: Attackers' View scenarios (No: 50.27%), followed by Spatial-Social: Public: Attackers' View and Spatial:Public:Attackers' View scenarios (No: 33% and 31% respectively). This also denotes that they were the most concerned about their privacy when being aware of what of their information are revealed to whom and sharing with the public, and when all data dimensions are exposed, which lead to derive the most privacy-threatening personal information.

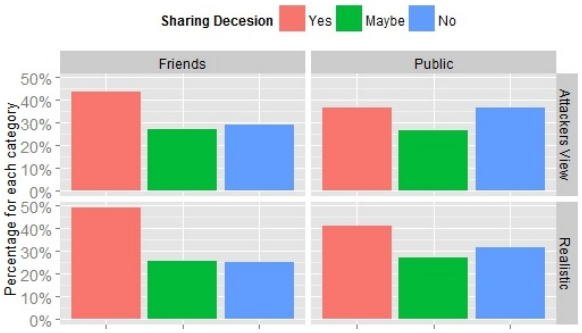


Figure 6.8: Percentage of selected sharing decisions (Yes, Maybe or No) categorised by visibility and awareness conditions .

6.4 Discussion and Implications

The validity of this study was carefully considered since we used hypothetical location-sharing scenarios. Such an approach has been effective in yielding generalisable outcomes in a considerable number of location-sharing studies (e.g. [38, 36, 92]). In fact, using hypothetical situations provides an advantage to this study by removing any restriction associated with using a particular GeoSN application, where its functionalities and specific environment affect users' location-sharing behaviour, and hence shift the focus from the core aim of this study (factors that affect users' location privacy perceptions). Furthermore, use of a between-subject study design for the different study conditions as well as showing randomly-ordered scenarios for each treatment ensured unbiased responses.

The scenarios were also developed in a way that enables the participants to realistically imagine themselves in these situations by providing them with the context of a location-sharing scenario that does not necessarily mimic what a GeoSN shows, but rather the user's experience when they

use it. In fact, the participants' explanations as to why they were reluctant to share their location reflects that they truly put themselves in the given scenarios and provide feedback based on that. This is shown specifically when they refer to themselves. For example, "When I was in political meeting, I would not share location without the permission of others" and "I answered maybe or no to places that I frequent because someone might see a pattern". They also referred to their real experiences, "I have a friend that checked in everywhere she went. Her profile was public so anyone could see it. Her house was burglarised twice in one month because everyone knew she wasn't home.", "I started to think about how if it was a regular routine that I was somewhere on a specific night, it was going to give other people potentially too much information about my habits and whereabouts", and "tagging a friend. I've done it in the past".

Users sense of privacy can vary and hence their perception of threat. Therefore, the factors that can impact users' privacy attitude, data dimensions, visibility, and awareness of the latter two factors were considered and identified based findings from studies presented in previous chapters before carrying out this experiment. Then, users' privacy perceptions were captured in relation to these logical factors to confirm the extent of their impact. In addition, large number of participants were recruited from different backgrounds and use experiences of GeoSNs to ensure yielding generalisable input and limit any potential bias. Participants justifications for their sharing decisions showed also that they understood the potential risks in the presented location-sharing scenarios. More importantly, indication of privacy threat level should be combined with showing the user feedback about what of their personal information can be extracted and who can see it to enable them to better understand and assess their privacy status and eliminate any potential biasing in the threat level indication.

6.4.1 Location-sharing Privacy Perception

The results demonstrate the general trend that users of GeoSNs actually value their privacy and suggests that they consider the various factors that revolve around location-sharing activities when making their sharing decisions as a way of managing their location privacy. Participants showed that they avoided sharing their locations and expressed greater concern when dealing with more privacy-threatening situations contributed by the study factors.

The data dimensions involved in the information extracted from location-sharing activities were showed to have significant impact on users' sharing attitude and privacy concerns. Location-sharing activities that reveal more personal information of users are correlated with less willingness to share and higher privacy concerns (particularly in Spatial-Social-Temporal scenarios). Overall, sensitive locations have shown to trigger higher sense of privacy where participants felt such places can reveal more personal information than they desire. Similarly, being with

a friend in place reduces users' likelihood of sharing their co-location to protect the privacy of both parties.

The visibility of the derived location information also influences users' location-sharing decisions, where participants' privacy concerns showed a strong association with sharing their locations with the other unknown users. This observation indicates that they want to restrict their information access in order to protect their privacy. The contextual information provided in the location-sharing scenarios suggests that presenting the privacy implications of sharing location information (as in the attackers' view) raised users' privacy awareness as implied by their explanations for their sharing decisions, and enabled them to make informed consent about their location information disclosure. The practice of informed consent is supported when users are aware of the privacy consequences of their online activities in terms of what data of theirs are disclosed, who can access it, and for what purpose, so they can make informed decisions [94]. Even the realistic scenarios show some kind of awareness in terms of frequency of previous visits to a place and with whom, while in real-world situations, GeoSNs do not display this information when users are about to share their location.

6.4.2 Privacy Design Implications and Proposal of Location Privacy Threat Levelling Model

The findings of this study provide a valuable lesson in terms of what aspects significantly influence users' attitudes to privacy, and hence should be considered when designing privacy notice . Data dimensions, including sensitivity, visibility, as well as awareness factors of the shared location-based information showed to have a clear impact on how users perceive their location privacy and ultimately decide upon their level of information disclosure.

Based on Westin/ Harris privacy segmentation index that examines users' feelings towards their privacy, in 2003, 26% were privacy fundamentalists who have a strong sense to protect their privacy and very concerned, 64% were privacy pragmatic who try to balance having the benefit with their personal privacy, and 10% were the privacy unconcerned [86]. This segmentation gives an overall users' perception towards privacy , whereas our study focuses on the domain of location privacy in GeoSNs as well as modelling users' privacy perceptions considering sharing factors. We can map this index to our findings to reveal the users' segmentation by corresponding 'yes' to privacy unconcerned, 'maybe' to privacy pragmatic and' no' to fundamentalists in each combination of disclosure factors.

We utilised users' perceptions to develop a LPTLM. This includes three privacy levels, as identified in Section 5.3.1: Green, where it is safe to disclose location information, Amber, where

Table 6.5: Participants' sharing decisions based on the data dimensions, visibility and sensitivity of disclosed information in the case of Realistic awareness.

Visibility		Friends			Public		
Data Dimensions	**Sensitivity**	*Yes*	*Mybe*	*No*	*Yes*	*Mybe*	*No*
Spatial	Insensitive	73%	18%	9%	67%	26%	7%
	Sensitive	45%	21%	34%	31%	28%	41%
Spatial-Social	Insensitive	67%	23%	10%	60%	24%	16%
	Sensitive	32%	30%	38%	25%	28%	47%
Spatial-Social-Temporal	Insensitive	57%	30%	13%	49%	31%	20%
	Sensitive	36%	31%	33%	31%	26%	43%

Table 6.6: Participants' sharing decisions based on the data dimensions, visibility and sensitivity of disclosed information in the case of Attackers' view.

Visibility		Friends			Public		
Data Dimensions	**Sensitivity**	*Yes*	*Maybe*	*No*	*Yes*	*Maybe*	*No*
Spatial	Insensitive	61%	23%	16%	55%	32%	13%
	Sensitive	30%	26%	44%	31%	27%	42%
Spatial-Social	Insensitive	57%	25%	18%	55%	32%	13%
	Sensitive	38%	28%	34%	21%	30%	49%
Spatial-Social-Temporal	Insensitive	52%	32%	16%	48%	28%	25%
	Sensitive	36%	27%	37%	25%	16%	59%

caution should be exercised when disclosing location information, and Red, where it is dangerous to disclose location information. In this model, we consider the three main factors of data dimension, visibility, and awareness, as well as data sensitivity due to its significant impact on user perception as observed. Participants' sharing decisions based on these aspects are presented in Table 6.5 in the case of Realistic awareness and in Table 6.6 in the case of Attackers' view.

One way of proposing this model is to directly map users' sharing attitude to the privacy level where 'Yes' is mapped to the Green level, 'Maybe' is mapped to the Amber level and 'No' is mapped to the Red level. The results of users' perceptions considering these aspects can be used directly and simply in a visual privacy indicator that shows the ratio of the three privacy levels in a given location-disclosure situation based on similar users' experience, as shown in Figure 6.9(a) when awareness is realistic and in Figure 6.9(b) when awareness is from the attackers' view.

Alternatively, the overall percentage of users' sharing decisions in each situation of the considered four aspects can be used to limit the privacy level to one or a maximum of two in order to provide a clearer estimation of privacy. If the majority of users with a threshold $\geq 60\%$ selected a certain sharing decision, then the privacy level mapped to this decision would be the proposed privacy level in the model. For example, 61% of users selected to share their location in the

(a)

(b)

Figure 6.9: Visual privacy indicators considering the dimension, visibility and sensitivity in the case of (a)Realistic awareness and (b) Attackers' view.

Table 6.7: The proposed LPTLM based on directly mapping privacy levels to sharing decisions by considering their proportions.

Awareness		*Realistic*		*Attackers' view*	
Visibility		*Friends*	*Public*	*Friends*	*Public*
Data Dimensions	**Sensitivity**				
Spatial	*Insensitive*	Green	Green	Green	Green∼Amber
	Sensitive	Amber	Amber	Amber	Amber∼Red
Spatial-Social	*Insensitive*	Green	Green	Green∼Amber	Green∼Amber
	Sensitive	Amber	Amber∼Red	Amber	Amber∼Red
Spatial-Social -Temporal	*Insensitive*	Green∼Amber	Green∼Amber	Green∼Amber	Green∼Amber
	Sensitive	Amber	Amber	Amber	Red

Spatial/Insensitive/Friend/Attachers' view scenarios, hence the suggested privacy level would be Green. Otherwise, it would be between the privacy levels assigned to the two highly-selected decisions if they are next to each other in regard to their ranking. For instance, in Spatial/Insensitive/Public/Attackers' view scenarios, the two decision chosen the most were 'Yes' (55%) and 'Maybe' (32%). Therefore, the proposed privacy level would be between Green and Amber. If the two highly-selected decisions are not adjacent in regard to their ranking which in this case are 'Yes' and 'No', then logically the privacy level would be Amber since it falls in the middle. Table 6.7 demonstrates the resulting LPTLM.

Table 6.8: The proposed LPTLM considering users' willingness to share based on average responses in both awareness types.

	Friends		**Public**	
	Insensitive	***Sensitive***	***Insensitive***	***Sensitive***
Spatial	Green	Amber	Green	Red
Spatial-Social	Green	Amber	Amber	Red
Spatial-Social-Temporal	Amber	Amber	Amber	Red

Another approach to developing the LPTLM is to consider the percentage of users who are willing to share their location in the different scenarios given. After taking the average results of sharing decision in both Realistic and Attacker's view awareness, where:

- A Green classification is used with a threshold of $\geq 60\%$ for the 'Yes' value.
- An Amber classification is used with a threshold of < 60% and > 30% for the 'Yes' value.
- A Red classification is used with a threshold of $\leq$ 30% for the 'Yes' value.

Table 6.8 presents the proposed LPTLM based on this approach.

A third technique for proposing the LPTLM is to evenly distribute users' responses to 'Maybe' between 'Yes' and 'No'. Participants' sharing decisions based on the data dimensions, visibility, awareness and sensitivity of disclosed information after this distribution are presented in Table 6.9. Then, the proportion of participants who were willing to share is considered to find the proper privacy level corresponding to the given location-sharing factors where:

- A Green classification is used with a threshold of $\geq 70\%$ for the 'Yes' value.
- An Amber classification is used with a threshold of < 70% and > 50% for the 'Yes' value.
- A Red classification is used with a threshold of $\leq$ 50% for the 'Yes' value.

The resulting LPTLM is shown in Table 6.10.

All of the previous three versions of LPTLM can be used in any privacy-aware system as they are all considered equally potent. All of them are developed by utilising the majority of participants' responses as a threshold to derive appropriate threat levels. Hence, as noted, most of the derived threat levels in correspondence to the sharing factors across these LPTLMs are similar. The proposed models as well as users' privacy perceptions can be used to offer suggestion for designing effective privacy notice as the followings:

Table 6.9: Participants' sharing decisions based on the data dimensions, visibility, awareness and sensitivity of disclosed information after splitting 'Maybe' responses between 'Yes' and 'No' .

Awareness		*Realistic*				*Attackers' view*			
Visibility		*Friends*		*Public*		*Friends*		*Public*	
Data Dimensions	**Sensitivity**	Yes	No	Yes	No	Yes	No	Yes	No
Spatial	*Insensitive*	82%	18%	80%	20%	72.5%	27.5%	71%	29%
	Sensitive	55.5%	44.5%	45%	55%	43%	57%	44.5%	55.5%
Spatial-Social	*Insensitive*	78.5%	21.5%	72%	28%	69.5%	30.5%	71%	29%
	Sensitive	47%	53%	39%	61%	52%	48%,	36%	64%
Spatial-Social -Temporal	*Insensitive*	72%	28%	64.5%	35.5%	68%	26%	62%	39%
	Sensitive	51 %	49%	44%	56%	49.5%	50.5%	33%	67%

Table 6.10: The proposed LPTLM based on evenly distributing users' responses to 'Maybe' between 'Yes' and 'No'.

Awareness		*Realistic*		*Attackers' view*	
Visibility		*Friends*	*Public*	*Friends*	*Public*
Data Dimensions	**Sensitivity**				
Spatial	*Insensitive*	Green	Green	Green	Green
	Sensitive	Amber	Red	**Red**	Red
Spatial-Social	*Insensitive*	Green	**Green**	Green	**Green**
	Sensitive	**Red**	Red	Amber	Red
Spatial-Social -Temporal	*Insensitive*	**Green**	**Amber**	Amber	**Amber**
	Sensitive	Amber	Red	Red	Red

- *Factors to Consider for Privacy Notice:* Participants showed different privacy attitudes towards different location-sharing scenarios. This suggests that when designing for privacy, a user must be informed about what dimensions of their data are exposed as well as the content of the disclosed information. They also need to be aware of who can see and access this information and clearly mark what is considered as sensitive information to allow them to make informed consent about their information sharing.

- *Use of Privacy Indicator:* A visual privacy indicator can be shown in the privacy notice as a way of providing easy privacy cues that can be coloured based on the given information disclosure factors. Three sub-indicators can also be shown for visibility, sensitivity and level of extracted information, to provide a detailed privacy status regarding each of these factors. Such indicators can efficiently impact user behaviour towards making informed sharing decisions.

- *Use of Privacy Score:* Privacy scores can be presented to users in order to help them assess their privacy status by calculating it based on the factors involved in a location disclosure action. This score can offer an effective approach to influencing user sharing

behaviour by increasing their privacy awareness.

- *Similar Experiences:* A privacy notice can show a user what others did in similar location-sharing situation which might aid in the process of decision-making.
- *Personalisation:* Findings indicate that configurable privacy settings should be offered in the privacy notice to users, since, although the majority may lean toward a certain attitude, others have different privacy views, as seen in the study results. Hence, privacy settings can enable a user to personalise the notice based on their personal privacy preferences.

6.5 Conclusion

This chapter presents an in-depth user-based study that investigates factors that influence users' location-sharing attitude in order to understand how users perceive location privacy and allow them to make informed consent for location sharing. Results showed that users' location-decisions are impacted by the dimensions of the exposed data and its sensitivity, visibility to others, and awareness of potential privacy implications. Users were less willing to share and hence more concerned about their privacy whenever the presented location-sharing scenario posed a greater risk to their privacy, especially when they are aware of the hidden implications. These factors should be considered in the design of any location privacy awareness system. The study outcomes were used to propose several versions of LPTLM that demonstrate how the results of users' privacy perception can be utilised to suggest a privacy level for a user based on the factors involved in a location-sharing task.

Chapter 7

Towards Holistic Geo-Profile View for Privacy Awareness on GeoSNs

7.1 Introduction

This chapter will address this issue using location privacy awareness by following the same aspects used in Chapters 5 and 6 which are the dimensions of the exposed data, their visibility to other, and awareness of the related privacy risks which all shown to contribute to the way in which the respondents perceive location privacy and behave with regard to location-sharing. However, it aims to propose a usable privacy-oriented interface that provides holistic view or, in another word, full access to a user's location-based profile resulting from his/her location sharing on GeoSNs. It offers a holistic view of the users' data whether directly collected or implicitly inferred about them including who can see them that can be accessed by users when needed. In addition, users' profile can be extracted from multiple GeoSN accounts to show a wider overview of what the users' are giving away of their information online. The location privacy awareness approach is used here in a different context than Chapter 5 that provided real-time and task-specific notifications. The design focuses on presenting privacy awareness information simply and directly using several visualisation methods, to clarify the meaning of the information shown. Employing visualisation techniques was shown to be an effective way of educating users about their data disclosure to online services, showing them the privacy implications and influencing their attitude to location-sharing [33, 122, 123, 124, 92]. The evaluation is carried out to measure the impact on users' privacy attitude and behaviour in a series of semi-structured interviews with users of GeoSNs, whose feedback is based on their own real-shared data.

7.2 System Overview

A Geo-Profile Visualisation system is proposed for supporting location privacy awareness. It provide full access to user's location-based profile derived from their location disclosure on GeoSNs. It aims to enhance users' awareness of their accumulated location-based containing information explicitly collected by GeoSNs and implicitly inferred about them on the basis of their shared data. The presentation of a geo-profile is privacy-oriented. Information is structured and displayed to serve the purpose of informing users about the privacy implications of location-sharing in a direct and simple way without leaving the users to discover these implications for themselves. The secondary goal of this system is to enable users to effectively manage their location privacy on GeoSNs by enhancing their awareness of what personal information has been collected and can be inferred about them, and who can view it so they can alter their location-sharing behaviour according to their attitude to privacy.

7.2.1 System Framework

It is assumed that the application stores a basic user location profile that records their spatiotemporal track of place visits and related contextual data, and that the application is also capable of deriving implicit information from this data, including, for example, the strength of the user's relationship to specific places, and their visit patterns to different places. Location privacy threats are triggered based on the content of a user's location profile, and consequently privacy alerts are generated. Figure 7.1 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.

7.2.1.1 Data Collection and Profile Generation

The data collection and profile generation module is responsible for collecting current check-in information as well as generating and updating the user's location-based profile. First, the current check-in data is acquired from user's social network accounts, represented using the check-in data model, as shown in Figure 7.2. Then, the information of the given check-in is transferred to the user profiling engine, where the user's location profiles are generated and updated. This engine processes the new raw check-in data along with the user history in terms of the previously collected check-in information to derive new personal information through inference cycles.

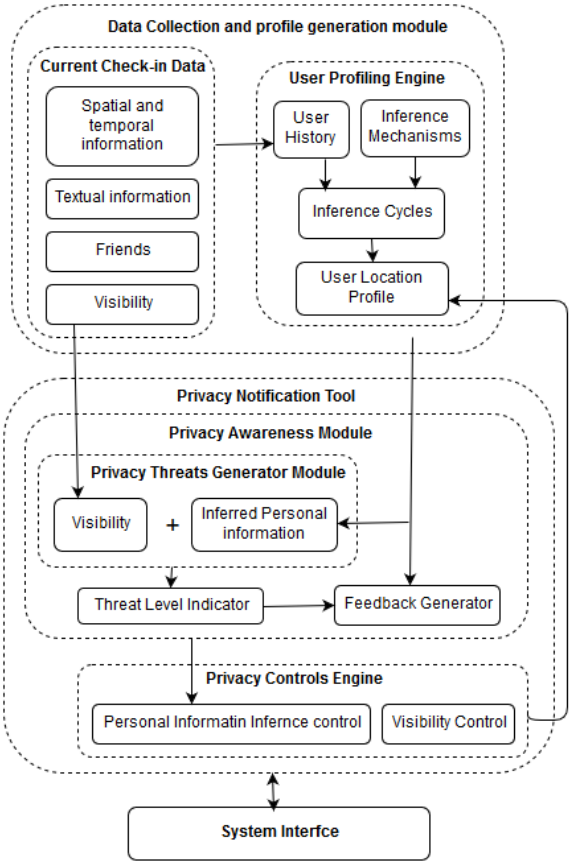


Figure 7.1: Components of the Location Privacy Awareness tool.

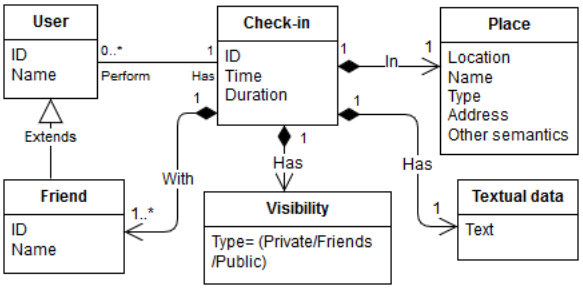


Figure 7.2: UML diagram of the check-in data model.

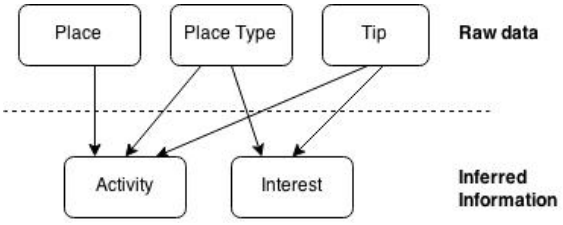


Figure 7.3: Links between place model elements.

The recently collected and inferred information is then used to update the user's location profile with new associations with places, interests, and other users. The place model defines links

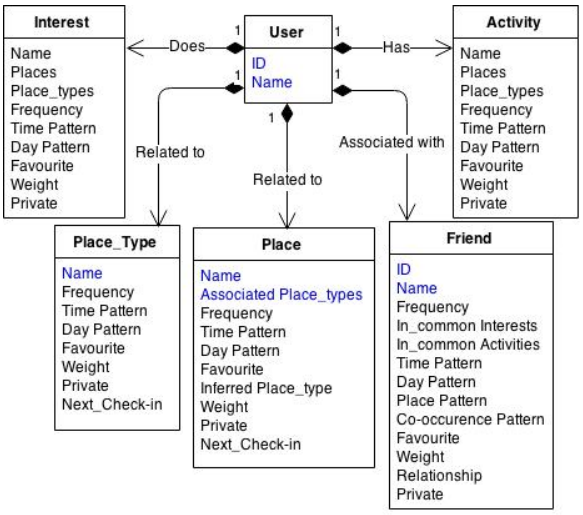


Figure 7.4: UML diagram of the user location profile model..

between check-in data and personal information, in which place, place type and tip are used to extract the user's activity, whereas the place type, and tip are used to infer the user's concepts and interests, as presented in Figure 7.3. Figure 7.4 shows the user location profile model. The attributes marked in blue are given directly, whereas the rest are inferred. The *Frequency* attribute has a cumulative value based on the number of times an element occurs in the user's history. Possible values for common attributes in the user location profile model are as follows:

- Frequency= (daily/frequent/occasional/ monthly/yearly)
- Time Pattern= (mornings/lunchtime/afternoon/evenings/night)
- Day pattern= (Monday/Tuesday/Wednesday/Thursday/Friday/working-days/weekends)
- Favourite= (yes/no)
- Weight: degree of association to user
- Private= (yes/no)

Hence, the generated user location profile can include information about:

- The user's relation to places (and place types):
  - Extracting favourite places
  - Revealing private places
  - Identifying absence and presence in private places
  - Inferring patterns of visits

  - Inferring type of association to a place (inferred user-related place type)
  - Predicting future movements and transitions between places

- The user's interests and activities
  - Extracting favourite interests and activities
  - Revealing sensitive (private) interests and activities
  - Deriving degree of association with interests and activities
  - Inferring pattern of interests and activities

- The user's association with friends
  - Extracting in common interests and activities
  - Inferring co-visiting patterns
  - Determining favourite friends and the type of relationship
  - Revealing private friends

7.2.1.2 Privacy Notification Tool

Once the user's location profile is enriched with the newly collected and inferred personal information, the data processing is handled by the privacy notification tool. The Privacy Awareness Module first identifies location privacy risks associated with a check-in in relation to the extracted personal profile and its visibility to others using the Privacy Threats Generator Module, which is used to estimate the threat level. A snapshot of the user's location profile is extracted through the Feedback Generator to provide a user information summary related to the current check-in. The information generated from this unit is then transferred to the tool interface to be presented to the user. Lastly, the Privacy Controls Engine provides the user with appropriate settings in regard to the derived personal profile and its visibility to others, as means of enabling privacy management.

7.2.2 The System Design

The system design should be developed in a way that serves the goal of the system, which is to raise users' awareness of the privacy implication of sharing their location-oriented information on GeoSNs by showing them the extent of the data that can be collected and information that can be inferred from them. Thus, the design is developed from a concern for privacy focusing

on this information. It also considers the usability aspect, in terms of the effectiveness and efficiency of the system in making it easy for users to find and understand needed information.

The following paragraphs discuss the design principles used for this system.

7.2.2.1 Browser of Geo-profile

This main browser's design shows the user as a central node which has outward links to other nodes representing the main concepts of the user's profile, namely, the visited places, interests and activities, and co-locations with friends. Once clicked, these nodes are expandable to sub-nodes which represent sub-categories. They show different extents of personal information collection and inferences, namely, a list of extracted information about a concept (e.g. the visited places or interests), the user's favourites or tops within this concept, and the patterns of this concept. The sub-nodes are ultimately linked to show a relevant privacy-awareness interface containing details of the user's information and their associated implication. Table 7.1 shows the geo-profile nodes, sub-nodes and what each of them displays.

This design uses the Fully Connected navigation model allowing a user to jump from any node to any other to view his/her their sub-nodes [138]. This visualisation approach is used to symbolise the user profile, the privacy implications and the links between them, which allow users to directly recognise the purpose of the system and of the information presented. Each sub-node is colour-coded depending on whether the information it shows is directly shared by the use or inferred from his/her shared data as a way of indicating to him/her who this information was acquired from. Figure 7.5 presents the design of the main window (browser) of Geo-Profile Visualiser when clicked on all nodes.

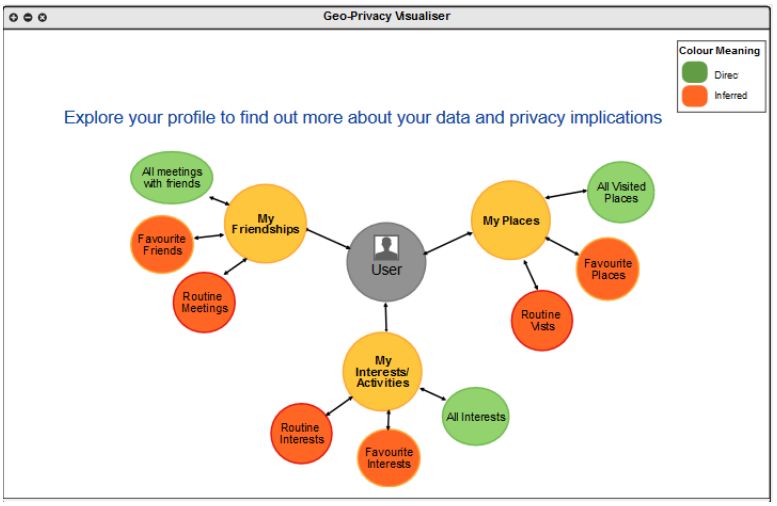


Figure 7.5: The design of the main window (browser) of Geo-Profile Visualiser when clicked on all nodes..

Table 7.1: Geo-profile structure and information displayed.

Main Node	Sub-Node	Attributes	Related to	Information Extraction Level
My Places	All Visited Places	Name, Coordinates, Date, Time, Type, Main type, Address, Visibility, Source		Direct
	Favourite Places	Name, Type, Main type, Address, No. of visits	All Visited Places	Inferred
	Routine visits	Name, Type, Main type, Time pattern, Day pattern, No. of visits	All Visited Places	Inferred
My Interests/ Activities	All Interests	Interest, Associated places, Associated place type, Associated timestamp	All Visited Places	Direct
	Favourite interest	Interest, Associated places, Associated place type, No. of visits	All Interests	Inferred
	Routine interests	Interest, Associated place types, Time pattern, Day pattern, No. of visits	All Interests	Inferred
My Friendships	All Meetings with Friends	Friend, common interests, Meeting place, Meeting timestamp, Meeting address	All Visited Places	Direct
	Favourite Friends	Friend, Meeting place type, No. of meetings	All Meetings with Friends	Inferred
	Routine meetings	Friend, Time pattern, Day pattern, Place pattern, No. of meetings	All Meetings with Friends	Inferred

7.2.2.2 Privacy Awareness Information Viewer

This viewer is basically the interface which appears when any sub-node is clicked on. It presents information of the following kinds. Figure 7.6 shows basic interface designs of the three sub-nodes that can be found in all of the main nodes, including the all directly collected information, favourites and patterns of each node:

- **Privacy Warning:** It displays a direct warning message that explicitly alerts the user to a privacy risk in terms of the type of information that is exposed by his/her data and its visibility to others. The warning also provides a link to the other privacy-awareness interface which holds information used for deriving current information, presented as a way of indicating the relationships between the users' data.

- **Personal Information Extracted:** It shows the relevant geo-profile information derived in three different approaches to makes it easy for users to explore and understand their information. Information is presented in a tabular format that gives a detailed and direct view. Information presented in the table can also be filtered according to its attributes (columns) to allow users to explore the data further. Information is also presented in a map-based visualisation that displays related places which can be clicked to show more specifics about a place in a tooltip. In the *favourites* and *routines* sub-nodes, the information, in addition, is represented using graphs to allow more visual exploration.

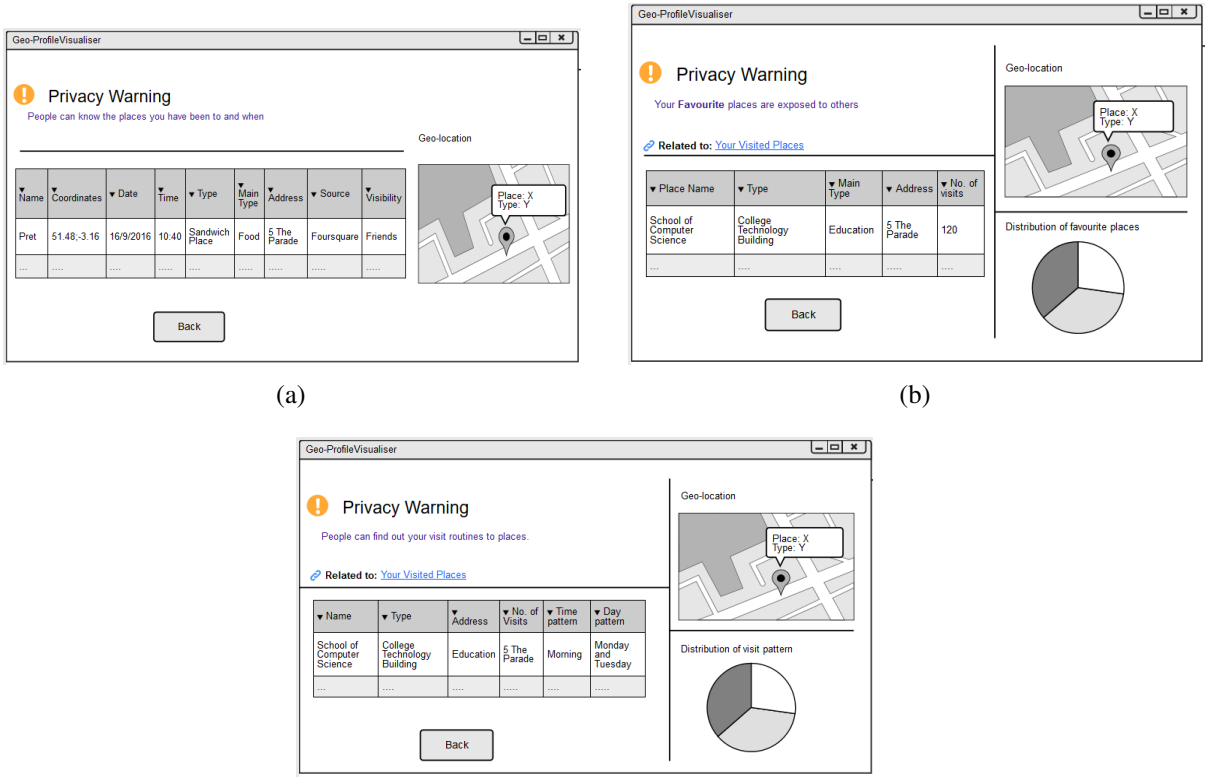


Figure 7.6: The information viewer designs for a) All visited places, b) Favourite place, and c) Routine visits..

7.3 Preliminary Testing For The Proposed Design

Before finalising the system design and evaluating it, a preliminary test is needed to ensure the usability of the system and to discover any potential problem with its design.

7.3.1 Cognitive Walkthrough

Usability inspection methods provide relatively cheap and efficient means of finding deficiencies in a user interface [139]. Heuristic evaluation is one popular inspection approach, which checks usability problems against a set of usability principles [140]. However, it is more important at this early stage to focus on how simple it would be for an inexperienced user of such a system to obtain the needed information. Cognitive walkthrough is one of the usability inspection methods that concentrates on testing how easy it is to learn to use a system, specifically, how to complete a task through exploration, assuming no previous knowledge [141]. This method was chosen for the present evaluation since it can be used at the first stage of design development and can be undertaken by one person at a time [139].

The evaluation was carried out in form of defining the task of finding different users from the system information provided, such as favourite places or patterns of visits, and recording the steps required for finding it and whether it was successful. Details of this evaluation can be found in Appendix E.1. The result of this test was that users' information was found easily in two steps.

7.3.2 Focus group

Focus group discussions are a well-established research method that allows a particular topic to be explored in detail. It offers some of the advantages of a questionnaire by enabling qualitative data to be gathered in the form of discussions which can offer in-depth insights on a subject [142]. Discussion between the moderator and members of a focus group also provides knowledge not only about people's thoughts but about the way in which they reason and form opinions [143]. Thus, a one-hour focus group session was held in February 2016 to review the design of the system and to discuss and gauge people's opinions on it. Seventeen volunteers were recruited to this group from Cardiff School of Computer Science and Informatics, nine females and eight males. All of them use social networking applications and, more to the point, 10 of them share their location information on these applications.

The session started by introducing the concept of a geo-profile visualiser, its purpose and how it worked. Then slides about the system design interfaces were shown. Next a discussion was led by the moderator on the potential of such a system, the usability of the design and the comprehension of information presented. The participants were also encouraged to add their ideas for adding to or altering the system. By the end of the session, a feedback sheets were handed to them were they were asked to record their opinions about a) validity and potentials of the system, b) clarity and usability of the design as well as information presentation, and c)

any other suggestions. Overall, the participants spoke of a positive impression of this awareness system and how useful it would be to use it for "privacy assessment". For example, one user of location services said "The system is very useful as it educates users about their privacy and raises their awareness of such risks"; another said "I believe this system has potential for people like me to be cautious of location sharing". Another added that it was "very practical and useful for the user to use it on a large scale". Other users of social networks who do not currently share their location also commented positively. For example, one participant said that it was "very interesting for people who use location-based social networks". As regards the design content feedback, they generally said "It is clear and to the point" and there was "nothing to be added". They liked the design of the geo-profile and said things, for example, "The spider diagram is quite simple to understand; being able to click on nodes and having more information". They also showed particular interest in the information-extraction indicator. For instance, one participant said "I like the coloured icons showing how the data was acquired" and another added "I liked the differently coloured indicators. Normally users pay more attention to visual indicators and hence they can be more careful about their privacy level".

Suggested Design Edits by the participants:

- Improve the text of the "colour meaning" of the information extracted and place it at the bottom of the screen to make it more noticeable
- Improve the presentation of the privacy warning text to make it more explicit and easier to understand what information is revealed and who can see it
- Showing what the node is displaying in a tooltip
- Show more a detailed graph for the routines
- Include a search option to search the results shown in the tables

7.4 User Geo-Profile Generation and Prototype Implementation

The first step for generating users' profiles is to collect their shared data on GeoSNs (which is Foursquare in this experiment as discussed in next section). The participants check-ins and their well as friends on the application information were collected using Foursquare's APIs. The participants were asked to download this information, save it in a text document and send it to the author. They were sent a PDF document with a simple step by step guide to completing this process. Those who needed help with this process were also offered personal assistance.

Apigee Console [1] is a platform that provides direct API calls to many social networks such as Facebook, Twitter and Foursquare; to authorise this website users simply need to log in with the social network account that they want to download their data from. It was used for collecting participants' data because of its simplicity and availability. Before the study could begin, the participants' check-in data had to be collected and processed in order to generate their geo-profiles and prepare suitable personalised tasks for each of them.

The users' data were retrieved in JSON format. A script (in Python) was developed to generate the participants' geo-profile by converting the data collected in JSON format into a rational database in .csv format. Figure 7.7 presents the UML diagram for the basic geo-profile structure demonstrating the relationship between its elements.

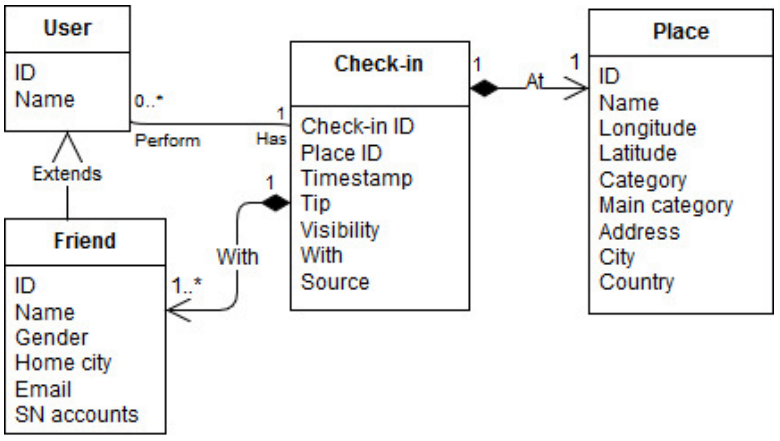


Figure 7.7: UML diagram for the geo-profile..

A desktop application was developed using Python as a prototype of the Geo-Profile Visualiser; it used the proposed system framework and design presented earlier in Section 7.2.2. PyQt4 was used for the design of the Graphical User Interfaces (GUI). Figure 7.8 shows the main screen of the Geo-Profile Visualiser when the "My Places" node is clicked on. More details about the information given in each node and sub-node are provided in a tooltip when the user hovers over them in order to facilitate navigation. If information of a sub-node is not available or cannot be extracted, such as there are no co-locations with friends or no visit patterns, a warning massage is displayed to user when clicking on it explaining that. The prototype uses the already-extracted .csv files of a user's basic geo-profile from a particular user directory to populate the system with user information. The Panda library was used for data analysis and mining which dealt directly with the .csv files. Information in each sub-node was basically presented in a table that can be sorted based on its attributes. Search functionalities are also provided in each sub-node privacy-awareness interface from which the information can be filtered according to relevant aspects of the data such as place name, type and data of check-in. The basic information in each node:

[1] Apigee Console, `https://apigee.com/console/foursquare` [Accessed: May. 2017]

"All Visited Places", "All Interests" and "All Meetings with Friends", is directly retrieved from the users' basic geo-profile. Interests are derived from the types of place visited by the user.

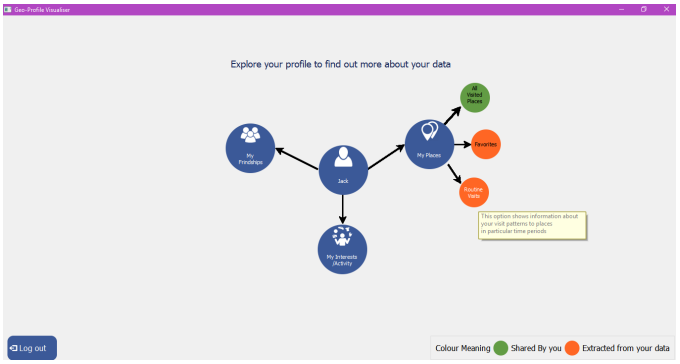


Figure 7.8: The main screen of the Geo-Profile Visualiser when the "My Places" node is clicked on..

However, more mining was needed for the other sub-nodes. In the "Favourite" sub-node, a simple query was used to find the users' favourite places, interests and friends, based on frequency (i.e. at least 5 times). The top 20 results are shown in descending order. As regards routines, a pattern was detected if 20% of a user's visits (occurrences) related to places, interests, or co-locations at particular times, days or both. At least 5 correlations had to be shown. For instance, If a user visited place X more than 30 times, 6 of them on a Friday morning, then it counted as a routine of visiting place X with a time pattern of 'morning' and a day pattern of 'Friday'. Please note that the focus of the study is not on the algorithm used for pattern extraction, which was used only to give an idea of the type of information that can be inferred from the location data. There are many studies in the literature that demonstrate and evaluate pattern-extraction algorithms (e.g. [45, 9, 8]).

The MatPlotLib library was used for generating the graphs in the sub-nodes 'Favourites' and 'Routines'. Pie-charts show the top 5 resulting favourites clustered on the basis of frequency, as presented in Figure 7.9. More detailed graphs are used to represent the 'Routines' sub-nodes where the subject of the pattern is displayed as a colour-coded point in relation to a time period on the x axis and in relation to a day of the week on the y axis, as displayed in Figure 7.10. Any graph can be zoomed into by clicking on it when it is shown in a separate window.

Google APIs were employed to produce the maps with a user's places plotted on them. A map can be dragged in any direction and can also be zoomed into. The user can also click on any row of a table to show the location information on the map. A tooltip is shown when a place marker clicked on to display more information about it. For example, in "Routine Visits" are shown the name and type of the place as well as the pattern related to it. An Internet connection is needed to load the maps.

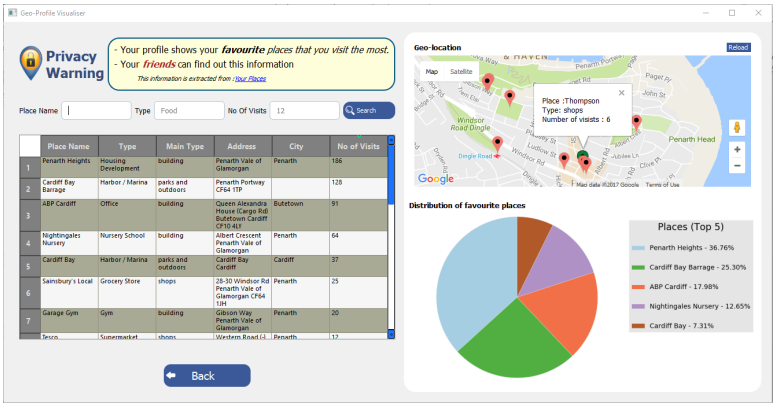


Figure 7.9: A screenshot of a user's Favourite Places interface..

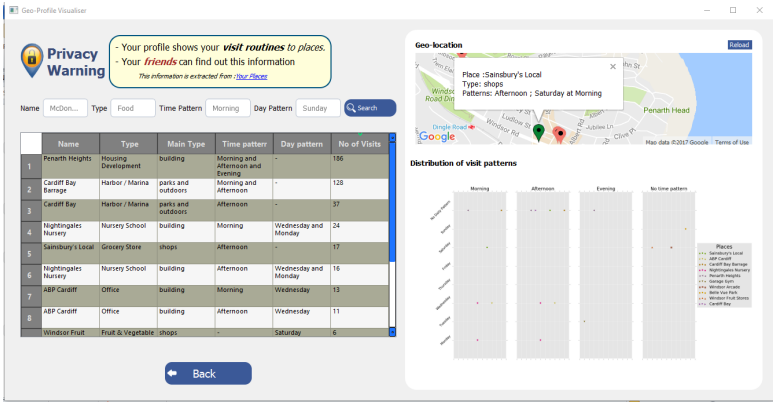


Figure 7.10: A screenshot of a user's Routine Visits interface..

7.5 User Study Evaluation

A user-based study was conducted to assess the impact of the geo-profile information presentation whether directly shared or derived from users' privacy awareness and attitude. In particular, the assessment considered the effectiveness and efficiency of the proposed geo-profile visualisation system for location privacy awareness that shows offline feedback on the user's aggregated geo-profile and related privacy implications. The usability of the prototype was also evaluated. The main objectives of the study are, first, to capture the impact of privacy perception in terms of a user's awareness of information collection, inference and accessibility, together with privacy concerns and attitude. This can be done by:

- asking the participants to carry out the pre-defined task of finding out more about their location-based information (whether directly collected or inferred) and observing and scoring them as they did so
- showing them statements and asking them to react with Likert-scale based answers (between Strongly Disagree and Strongly Agree) that capture their privacy awareness and attitude

- asking the participants open-ended questions to allow for in-depth discussion of their reactions, privacy concerns and possible changes to location-sharing behaviour

The second objective is to test the usability of the design (in particular, the aspect of usable privacy in terms of enhancing the presentation and management of privacy information) by:

- observing how easy it is to find information and how many steps are required
- showing them statements and asking them to react with Likert-scale based answers (between Strongly Disagree and Strongly Agree) about the system's usability

7.5.1 Method

Considering the principles of privacy-oriented experimental design [136], to eliminate any possible bias in the results, a between-subjects design was used to examine the influence of presenting privacy-oriented geo-profile information on users' privacy awareness and attitude. Participants were randomly assigned to one of the study groups; a no-awareness (baseline) group and an awareness group. In the no-awareness group, participants were asked to performed tasks about finding information related to their mobility using the offered tool by a GeoSN for accessing users' data, whereas in the awareness group, participants performed their tasks using the proposed geo-profile visualiser. Participants in both groups access and view their own shard location data and provide their feedback accordingly.

Again, Foursquare and its check-in application, Swarm, were chosen as the platform for this study due to its popularity and the representativeness of the GeoSNs. It also has been used in broadly similar previous studies in the literature [26, 16, 71, 45]. Hence, Swarm History, with which Swarm lets its users view and search their check-in history, was used for the no-awareness group. Figure 7.11 shows a screenshot of Swarm History.

Semi-structured interviews were used to allow for participants' reactions and the reasoning behind them to be explored in greater depth than the use of questionnaires permits. Quantitative and qualitative methods (a mixed approach) were employed in this study to better capture participants' feedback through open- as well as close-ended questions, which enabled statistical tests to be made. Participants were already users of this Swarm application but were not under any pressure to use it merely for the study purposes. Moreover, their actual shared location (check-in) data were used to ensure the validity as well as the accuracy of the study outcomes.

The experiment had three phases:

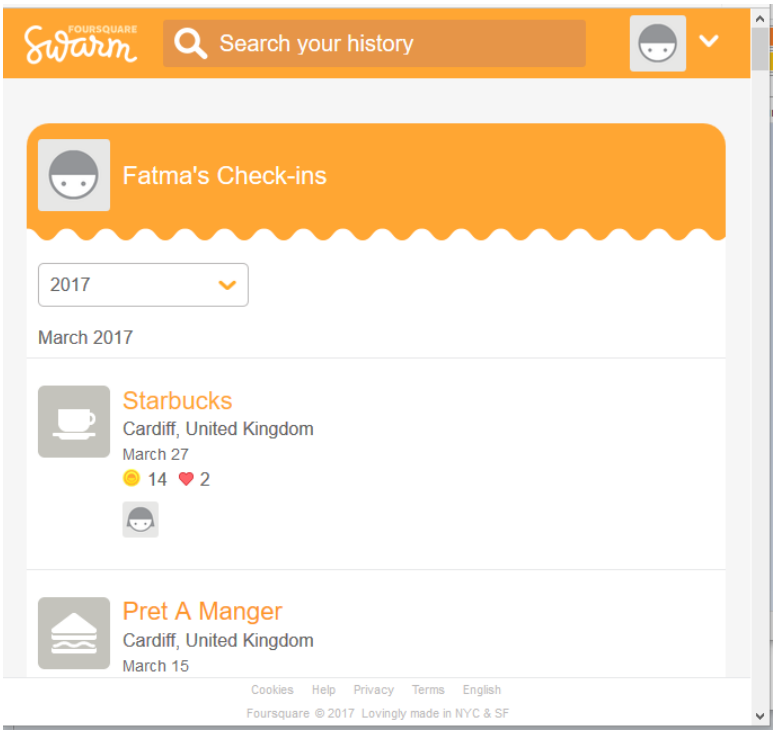


Figure 7.11: A screenshot of Swarm History..

1) **Phase 1:** Selecting participants who were regular users of the Foursquare/Swarm application and had checked into many places before

2) **Phase 2:** Collecting the check-in data that they shared on the application, and processing it to generate personalised geo-profiles for them, using the prototype. Then, personalised tasks were defined for each participant based on his/her geo-profile since information that can be derived may vary, depending on the user's location-sharing behaviour.

3) **Phase 3:** Conducting the study itself, in which the participants were interviewed and used the tool selected for each group.

7.5.2 Recruitment and Participants

The experiment was conducted during February-April 2017. The participants were solicited by email, on social media (including messages on Swarm application), and with face-to-face invitations. Participants had to be users of the Swarm application at the time and to regularly check into places in order to have suitable experience and the dataset required for the study. They had to be willing to provide personal check-in data for use in this study and able to attend an interview in person. All of these factors increased the difficulty of finding participants. However, enough time was spent to recruit representative participants with diverse backgrounds

and experience as users.

Two participants did not attend the interview and were assumed to have withdrawn; thus 26 participants in total completed the experiment; 13 in the no-awareness group and 13 in the awareness group. The participants checked in between 36 and 14911 times. Seven participants had fewer than 100 check-ins; 11 had between 100 and 1000, and the remaining 8 had more than 1000. Twelve of them have visited between 17 and 83 places, and fourteen have checked in between 117 and 493 places. When considering the categories of their places, nineteen of participants have between 15 and 98 place categories, while the reaming have between 110 and 154. Their ages ranged between 19 and 41 (average =27, SD=5.311). Eleven of them were students in Computer Science (9 postgraduates and 2 undergraduates), three were students in the Social Sciences (2 postgraduates and 1 undergraduate), three undergraduates in Engineering, two in Business, one postgraduate in Medicine, one in Art, one in Mathematics, one in Software Engineering, one in Mechanical Engineer, one Commercial Manager, one lecturer in in Computer Science. The female participants (16) outnumbered the males (10). Half of them were Asian, nine European, three from Africa and one from the Caribbean.

7.5.3 Study Procedure

The study was carried out as a series of semi-structured interviews that started by ascertaining a brief demographic background. Then two Likert scale questions were asked, to be answered in a range from Strongly Disagree to Strongly Agree; the intention was to capture the participants' sense of safety in using a Swarm application and their level of concern over online privacy. These were followed by closed- and open-ended questions to gauge the participants' social networking experience and online privacy perceptions in terms of their knowledge of and attitude to data collection, use and control.

The second part involved the use of a location-data access tool specified for each group: Swarm History for the no-awareness group and a Geo-Profile Visualiser tool for the awareness group. For each group, the interviewer started with a brief description of what the tool shows or provides and for a few minutes allowed the participant to explore his/her profile using the tool. Then the participants were asked to carry out pre-defined tasks that were personalised for them on the basis of their generated geo-profile (7 tasks on average). They were encouraged to think aloud and were observed as they used the tool to perform the tasks; in this stage the computer screen was video-recorded. The performance of the tasks is scored as a total success, a partial success or failure. The aim of the tasks was to find whether it was possible to find all a users' shared and possibly inferred data, the accessibility of their data and the ease and explicitness of accessing it. In addition, they were designed to measure the degree of privacy awareness res-

Table 7.2: A list of all the possible tasks allotted to the participants..

	Task
1	Find the place, place type and time of any recent check-in of yours
2	Find one of your favourite places (most visited), and how many times you visited it
3	Find a routine place that you visit at a regular time or day
4	Find who can see your information, such as places you visited and favourite places
5	find an interest/activity of yours and where it took place
6	Find one of your favourite interests (most visited), and see how many times your it took a place
7	Find a routine interest of yours that takes place at a regular time or day
8	Find what information about you is directly shared by you or inferred on the basis of the information you shared, such as your routine interest/ activity are
9	Find the place and address of a checking/meeting with a friend of yours
10	Find one of your favourite friends that you check-in with a lot and how many times you checked-in together
11	Find a routine meeting with one of your friends that usually takes place in a particular place, day or time

ulted from finding such information in terms of being informed about hidden knowledge about themselves. Table 7.2 provides a list of all the possible tasks allotted to the participants, each participant receiving his/her own sub-set, depending on his/her own geo-profile. To record their impression of performing the tasks, the participants were asked to react to a series of statements in a set of Likert-scale answers that ranged from Strongly Disagree to Strongly Agree. The aim was to capture the level of accessibility of the data provided to users, the impact of the tool on the level of awareness of their data disclosure and its relationship to the privacy implications, as well as the influence of the latter on their attitude to privacy. Open-ended discussion of the same aspect measured in the above scales followed, to allow the participants' views to be investigated in more depth in order to understand the impact of their level of privacy awareness and consequently their attitude.

Participants of the awareness group were asked to provide feedback on the usability of the Geo-Profile Visualiser tool using a five-item Likert scale, from Strongly Disagree to Strongly Agree which had been developed to test some specific aspects of this tool. A standard and widely-adopted usability evaluation scheme, the System Usability Scale (SUS) [2] was also used. It included both positively- and negatively-framed statements for which the score of the former was the scale position minus 1 and of the latter was 5 minus the scale position. The scores were then summed and multiplied by 2.5. The total score was then compared with the standard

[2]System Usability Scale, `https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html`[Accessed: May. 2017]

average score of 69.

Finally, the members of both groups answered again two post-study questions, the same as the two Likert-scale questions that captured the participants' sense of safety using the Swarm application and their level of concern over online privacy. The purpose was to examine any hint that the tool they had used had influenced their answers. They were also asked to rate on a 5-point Likert scale their willingness to share their location profile, as a way of capturing their attitude to privacy. The members of the awareness group were asked three further questions about a) whether they thought their initial understanding of the collection of location information and privacy implications had been limited; b) if they were tempted to change the amount of location information that they shared and how; and c) whether they would want to use the Geo-Profile Visualiser and for what purposes. These questions were not applicable to the no-awareness group, which used only what the application provided. The study question may be found in Appendix E.3.

7.5.4 Pilot Study

Pilot tests were first carried out on four volunteers from the school in order to ensure that participants could easily follow all the study stages and understand its requirements so that they could provide proper feedback. These pilot tests were also aimed to examine the experimental set-up, in terms of the video recording of the computer screen as the participants interacted with the tools, the voice recording throughout the interview, and the time needed for completion. Below is the list of questions used to capture the volunteers' feedback on this study:

- Were the questions both closed- and open-ended easy to understand?
- Was it clear to you what you needed to do in the tasks?
- Did you need more clarification on any question before you could answer it? If yes, which ones?
- Were you able to follow all the study stages?
- Do you have any additional comment you would like to add about this study?

They all agreed that the study stages overall were clear and connected and they could follow them throughout. They reported that they understood the questions and the tasks and what feedback had been needed from them. However, they made some comments on the wording and scope of some of the questions which needed improving, as shown in Table 7.3, below.

Table 7.3: Suggested edits from the pilot study .

Question	Comments	Edits
Have you ever used tools/applications that help you manage your online privacy?	Confusing, not clear, give example	Have you ever used tools/applications that help you manage your online privacy? **(e.g. browser add-ons)**
Who do you think can access your check-in data?	In which application/service	Who do you think can access your **Swarm** check-in data?
Have you ever checked your privacy settings?	In which application/service	Have you ever checked your privacy settings in **Foursquare/Swarm**?

7.6 Results

In this section, the study results are analysed and presented. The Fisher Exact test is used for examining whether the impact of the type of tool used (Swarm History and Geo-Profile Visualiser) on users' privacy awareness and attitude is significant. It is an alternative to the Chi-square test of independence to provide more accurate results when the sample size is small. The related Cramer's V test was used to examine the strength of this association regardless of sample size, by comparing the difference between the means of two samples where 0.1 is considered small, 0.3 is medium and .5 is large [144].

Moreover, the Friedman Chi-Square test and ordinal logistic regression were used to find the significance and direction of influence of the tool provided to perform their tasks, on some aspects that were measured before and after the actual study.

7.6.1 Pre-Study: Technological Experience, Privacy Awareness and Attitude

More than half the participants were experienced users of web applications and technologies (54%), while 31% were very experienced and the rest were somewhat experienced. Most felt safe using the Swarm application (3.8 on average out of 5), yet they were generally concerned about their online privacy (4 on average).

Some aspects of the users' attitude to online privacy were captured to provide a rough idea of their sense of privacy, as shown in Figure 7.12. Two-thirds of them had deleted a post shared on their social network accounts due to privacy concerns and had checked their shared check-in on Swarm History, mostly to see where they had been or remember the name of a place that they had visited. 65% of the participants had regretted sharing certain information on a social

network account, which indicates that users can share information without being conscious at the time of its potential consequences. In addition, 38% of them have used tools that helped them to manage their online privacy, ad blockers being the most used. Three of them mentioned using other tools including a blocker for sharing on social networks from other websites, a manager for website trackers, and a twitter add-on that deletes all tweets before a certain date. Moreover, 35% of the participants have requested to delete their data, for instance their profiles from an online service.

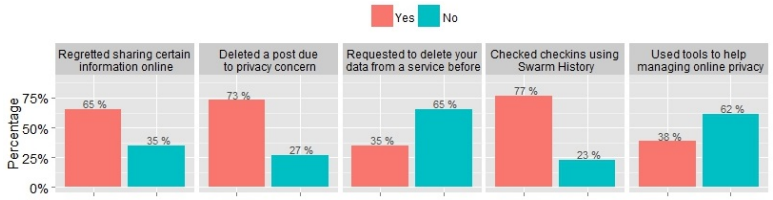


Figure 7.12: Participants' attitude to online privacy..

About 60% of the participants were not aware of some of the Foursquare/Swarm practices of collecting and sharing users' data, as presented in Figure 7.13. In particular, 85% of them were not aware that the application could share their data with third parties for other purposes than marketing and 47% of those were certain that the application could not do so. In addition, 58% of the participants were not conscious of sharing their data for marketing and advertisement purposes, while 36% of them thought that the application could not do so. More than half of the participants were aware that the application could collect their location even if they were not interacting with it. They said that they knew this mainly because the application had given them suggestions based on places that they had not checked into or asked them later whether they been to a place they visited without checking in. The remaining 38% were not aware of this practice.

Figure 7.13: Participants' attitude to online privacy..

At this stage, the participants' perceptions about their location data accessibility, potential use and control were also examined. When generally asked who they thought could access their data, almost all the participants said 'The application and their friends on it', as shown in Figure

7.14. 10 of the 26 participants also said that other users could also see their data. Seven of these participants explained in what situations this could happen: when they went to their profile information (2); when sharing their check-in on other social networks (2); when other users were at the same time in the same place as the participants (2); and if they were the mayor of a place (1). Only seven participants were aware that third parties could access their data, and only one participant thought that her data could be accessed by herself alone. This suggests that the participants are not fully aware of all the parties that can view their data. Figure 7.15 demonstrates the participants responses when they were asked what could be known about them when they shared their check-ins. Most of their answers concerned knowing where they went and when (7); their interests (7); and where they worked, studied or lived (7), which is more or less the basic information they disclose when sharing their locations. Fewer participants gave more thoughtful answers, including anyone else in the same place (2); their activities (2); if they were the mayor of the place (2); and co-locations with friends (2). Again, most of the participants seemed to have a limited view of what personal information could be inferred about them when they shared their location data.

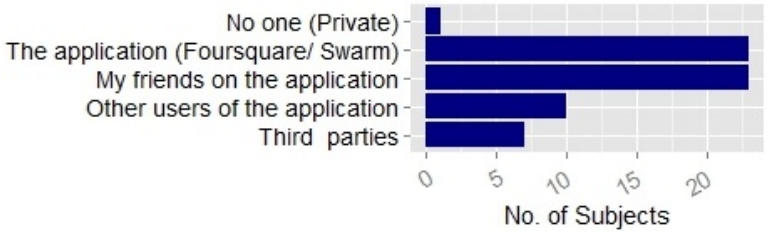


Figure 7.14: Participants' perception about their data..

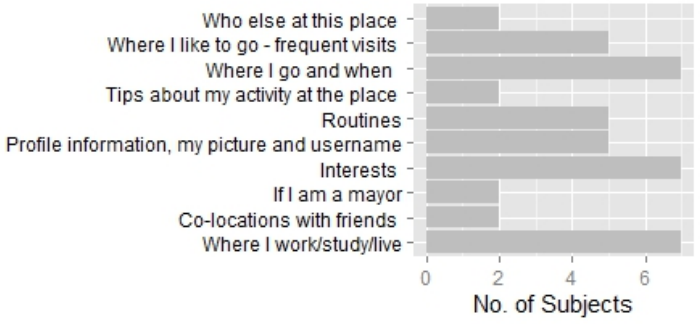


Figure 7.15: Participants' views on possible information extraction based on their check-ins..

Regarding data control, 65% (17) of the participants said that they had never checked the privacy setting provided by the application. 89% of those who had checked this did so rarely (8 out of 9). 16 of the participants said they could control who could see their check-ins (private, friends or public) and one of them even added that he could choose certain people to share with, but

this is not possible in the applications. Five thought that they had no control over their data. P11 said “I have to give a location, since I want to use the service” and P9 added “I don’t have much control since I have to share my location to use the service. Not much privacy associated with the app.” P18 commented “Even if I delete them, there may be backups”. Two participants said they could control the level of the location they shared. Two of them mentioned that their check-in history was already private and there was no need for more control and another two had no idea of the answer to this question. When asked to indicate what they could control among the options provided in a list (see Figure 7.16), most of them (21+) thought that they could delete all their check-ins and delete their profile, but this option is not available. Many of the participants (up to 22) were not aware of an important feature of their check-in data: that they can control it with the privacy setting. For instance, they can reveal the location when they are not using the application, get targeted ads outside the application and get checked in by friends. Surprisingly, some participants thought that they could control aspects that are not in fact available in the application, namely, collecting location while using the application (4) and getting targeted ads inside the application (5). Thus, the participants often seem to be unaware of which features of their data they can control.

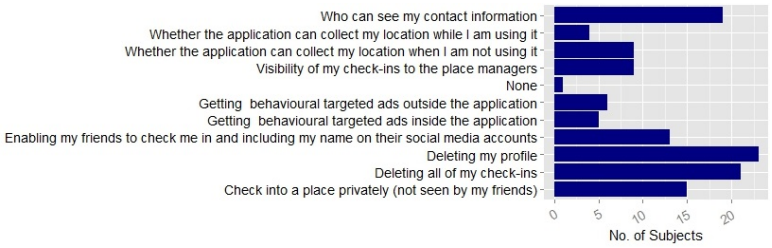


Figure 7.16: Participants’ feedback about what they can control of their privacy on Foursquare Swarm..

7.6.2 Task Outcomes

When the participants were asked to perform their personalised tasks of finding privacy-related information about their shared location information by means of the tool provided for their group, this tool made a significant impact on their success in performing the tasks(Pearson Chi-Square= 36.737096, p<.0001). In addition, this association has a large effect size (Cramer’s V= 0.742620, p<.0001). Each participant’s performance of the tasks is presented in Figure 7.17 for the Swarm History (no-awareness) group and in Figure 7.18 for the Geo-Profile Visualiser prototype (awareness) group. On average, the participants of the former group were able to successfully retrieve only 25% of the information they were asked to find, 10% found some of it (partial) and the remaining 65% of tasks failed to find any. However, the average successful score showed a considerable increase to 97% of the latter group; 11 out of these 13 participants

were able to perform 100% of their assigned tasks successfully. Not a single failed task was recorded for this group.

Figure 7.17: Task results for each participant (P1-P9) in the Swarm History (no-awareness) group..

Figure 7.18: Task results for each participant in the Geo-Profile Visualiser prototype (awareness) group..

In a comparison of the two groups' capacity to find relevant information, the participants of the Swarm History group were able to successfully find 78% of the basic information they shared, including their visited places, interests and co-locations with friends, as well as 47% of their favourite (or most frequented) places, interests and co-location as demonstrated in Figure 7.19. 81% of the finding-patterns tasks were completely failed when no hints were provided and complete failure was recorded for 100% of the data extraction task (indicating what information about a user had been directly shared by him/her or inferred on the basis of the shared information) and for the data accessibility task (indicating who could see their information). However, in the group using the Geo-Profile Visualiser prototype, almost all the types of information asked for in the tasks were successfully found (basics, favourites, patterns and data accessibility). Three participants slightly struggled to find the indication of the data extraction level.

The participants using the Swarm History to carry out their personalised tasks showed that they had difficulty in locating their information in the Swarm History. These participants thought that finding details about their shared data was not explicit or easy and also took a relatively long time. They had to navigate many steps in their history, using search and follow, to reach certain information. Sometimes the steps that a participant took did not lead them to the information required and they had to try another route. Even finding the simplest information was a struggle

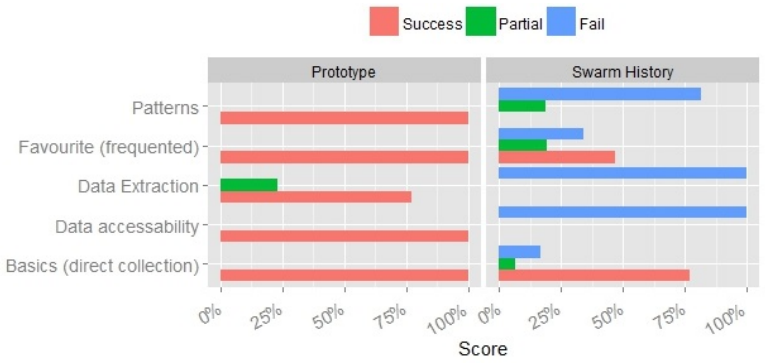


Figure 7.19: Capacity to find relevant information in both groups based on the type of tasks..

for some participants, for example, they found the type of a place by interpreting the icon shown next to the place name. This confused some participants, since different icons can be used for the same type of place. Many participants also relied on using their memory of where they had gone to provide an answer or at least to help them finding it. In contrast, performing the tasks using the Geo-Profile Visualiser prototype was easy, fast and straightforward. Almost all the participants were able to find the needed information from the beginning using only two steps.

7.6.3 Information Awareness and Privacy Attitude

This section examines quantitatively (by Likert-scaled answers) participants' opinions of and attitudes to their ability to find information related to their privacy, using the tool provided for their group. These data are supported qualitatively (by data derived from open-ended questions). The results of statistical tests of the quantitative data, including the participants' rating on 5-point Likert scales of some impact statements are shown in Table 7.4.

7.6.3.1 Impact on Privacy Awareness

The type of tool used significantly affects both the participants' ability to view the information they share when they check-in and their understanding of it. These two associations also have a large effect size whereby the participants of the prototype group tend to agree more strongly (4.8), whereas those in the Swarm History group tended to seem more neutral about these statements (3). When asked whether they found the tool helpful for accessing their data collected, those in the no-awareness group generally reported that Swarm History showed them a basic and general view of their shared check-ins which some found vague. Almost all of them (with one exception) wanted more explicit details about their check-in data. Some of them said that they had to do more searching to reach the desired information and that the presentation needs

Table 7.4: Analysis results of participants opinion towards their ability to find information in both groups using some statements..

Statement	Swarm History Average Results	Geo-Profile prototype Average Results	Fisher Exact Test	Cramer's V
This tool helps me understand the information I share when I check-in	2.8	4.7	<.0001	0.858
This tool allows me to view the information I share when I check-in	3.3	4.9	0.0001	0.788
This tool helps me understand the possible information that can be extracted about me when I check-in (e.g. patterns of visits, top interests)	1.8	4.8	<.00000 1	1
This tool allows me to view the possible information that can be extracted about me when I check-in (e.g. patterns of visits, top interests)	1.8	4.8	<.00000 1	1
This tool allows me to know who can access my data	1.3	4.1	<.0000 1	0.936
This tool motivates me to be more in control of my online data (e.g. deleting posts, updating privacy settings)	2.5	4.2	0.0034	0.704
I am satisfied with the way Foursquare/Swarm collects and stores my data	3.2	2.9	0.058	0.575
This tool makes me more concerned about my privacy	2.6	4.4	0.0011	0.741
This tool encourages me to alter the way I share my location information to protect my privacy	2.4	4.2	0.0067	0.693

to be improved including more filtering and ranking features. For example, P7 said "It should be enhanced and presented it in a better way" and P8 added "I need to do a lot of work to find out information.". Others felt that the application had collected more information about them than it showed. For instance, P12 said that "Very limited, the app collects more things" and P11 mentioned "It is not complete truth but I have to trust them under the privacy policy". As regards the awareness-group participants, they all found that the Geo-Profile Visualiser showed a more detailed and direct view of their location data, which they could find easily. P17 said that it is "More detailed than swarm app, my profile is grouped in a nice way which makes it easier for me to find information". P18 added that "They even the coordinates of the place. It is scary that they know that specific. I am surprised by the number of check-ins I have. Swarm

app doesn't present these details but focuses on displaying information in a fun way."

In terms of data inference, the type of tool used also had a significant impact on helping participants to see and understand the information that could be extracted about them when they check-in (e.g. patterns of visits, top interests). This association has a large effect size- the participants of the Geo-Profile group tended more to strongly agree (4.8), while those in Swarm History group tended more to disagree with these statements (1.8). Asking them how helpful the tool was in finding the kind of information that can be extracted from their data revealed that all the participants in the no-awareness group reported that Swarm History did not show them this kind of information. They said that "It gives the bare minimum" and they have to work it out in writing "Using pen and paper" or, as P12 stated, "I have to make my own assumptions". All the participants showed an interest in knowing about and viewing their information (possibly extracted). Some of them wanted to know because of their concern for privacy. For example, P3 said "It would be enlightening about my pattern since it is something I was not conscious of, it would teach me how much I share and to change the way I share if I am not aware of it". P2 mentioned that "It would be good to have, due to privacy concerns, so I know what I am sharing and what others can know about me". A few others referred to using the extracted information for other life-management purposes. P5, for example, said "Patterns will help in day planning and time management" and P7 added "I can even discover things about myself, I might be aroused by the amount of time I spend in certain places". Examining the feedback on this question from the awareness-group participants indicated that all of them found the Geo-Profile Visualiser helpful and thought it interesting to view what can be inferred about them and they expressed their need to be aware of such information. They said that there was "too much" extracted information and that "Some of them [the details] I didn't realise" and "I find it is very enlightening. I had no idea that such information can be inferred about me." P18 pointed out how the Swarm application encourages users to check into a place by focusing on the game-playing aspect; they said "The Swarm app is motivating and makes it appealing to use and share location by providing stickers and being a mayor of places". P11 also indicated that Swarm hides possibly inferred information in order to keep users using the app. He said "The app shows the bare minimum just to comply with legal requirements since they have to give users some access to their data. Swarm is based on gamification. It does not show more of the users' data to them, such as daily patterns, so users won't get freaked out about it".

The type of tool used has a significant impact on informing the participants about who can access their data. This association showed also a large effect size. The participants in the prototype group tend more to agree (4.1), whereas those in the Swarm History group seemed more to disagree with this statement (1.3). Further elaboration on inputs from the no-awareness group revealed that all of them would in fact like to know who can view or has viewed their information. Two of them were particularly interested in knowing how accessible their data

were to third parties. P8 pointed “The app is deliberately not showing who can access my data so I don’t get scared and stop using it”. All the awareness-group found the Geo-Profile Visualiser helpful in revealing the accessibility of their data to protect their privacy. Four of them also wanted to know who else apart from their friends could see their data.

7.6.3.2 Impact on Privacy Attitude

The tool used in each group was shown to have a significant impact on motivating participants to be more in control of their online data by, for example, deleting posts or updating privacy settings. This association also has a large effect size. Members of the Swarm History group were on average not motivated to perform this action (2.5), compared to the Geo-Profile Visualiser group (4.2). In addition, participants’ satisfaction with the way Foursquare/Swarm collects and stores their data was significantly different between the two groups and also had a large effect size. About half of the awareness group were not satisfied, compared to only one participant in the no-awareness group.

The participants’ privacy concerns were also significantly impacted by the tool used and a large effect size was noted where the members of the awareness group were generally concerned (4.4), while those in the no-awareness group were not. The members of the Swarm History group explained why they were not concerned; it showed only the basic check-in information that they had chosen to share. Interestingly, four participants were concerned about who could access their data. For instance, P12 said “I am more concerned since I don’t know who can see what of my data because the app collects more but it is not showing it and I should know.” P2 added “I am more concerned about who can see my information and it does not tell me that”. The participants who used the Geo-Profile Visualiser generally mentioned that they were concerned because the revealed personal information was beyond what they had shared and they had not realised that this was possible. For example, P17 said “It shows who can see my data. My information can be inferred and I am not aware. Extracted interests and patterns can spoil my privacy.”. P22 also added “I am scared. It reveals more than I expected. People can use it to do something bad, hurt me or stalk me”. Moreover, the type of tool used showed a significant impact on encouraging participants to protect their privacy by altering the way they share their location information. This association has a large effect size. In particular, the Geo-Profile Visualiser group on average showed more willingness to change how the share their location information (4.2) than the Swarm History group (2.4).

7.6.4 Usability of Geo-Profile Visualiser Prototype

The usability of the Geo-Profile Visualiser prototype was evaluated by the members of the awareness group. Swarm History is part of a commercial and well-established GeoSN application and hence its usability is beyond the scope of the present study. The usability results are shown in Table 7.5. At first, five specific usability statements that measure particular aspect in the Geo-profile prototype were put to the participants. The results capture the participants' experience with the prototype in terms of its simplicity, information organisation and presentation and its adoption for learning about and managing online data. The participants are inclined to strongly agree with these statements (4.7 on average). Using the standard SUS for usability, the calculated score was 87.12 according to this scale, which is considered above average (68).

The participants' feedback when asked open-ended questions about what they liked in the prototype design emphasised how well it was structured and presented for their profiles. For instance, P 17 said "User-friendly, nice grouping to related concepts, provides good details and different ways of presenting data" and P25 stated "the colour scheme is informative, maps and graphs make information easy to grasp at a glance, and visualisations are useful and help to understand". However, among their dislikes or wanted to be improved in the prototype, P17 and P18 complained that the colour meaning for the data extraction level was not clear. P22 said that "It would be nice to have a different view for the check-ins made by my friends".

7.6.5 Post-Study: Impact on Sense of Privacy and Safety

After the participants of each group had used the tool provided to find privacy-based information about their location profiles and their reactions to privacy awareness and attitude had been captured, they were asked to rate again how safe they felt in using Foursquare Swarm and how concerned they were about their online privacy. In the awareness group, the Geo-profile Visualiser prototype showed that it had a very significant impact on the participants' sense of safety before and after using it (Friedman Chi-Square = 12.000, p = .001). The participants felt significantly less safe after using the prototype than before using it (ordinal regression coefficient= -3.975, p= 0.000309). Initially, they had felt safe using Foursquare Swarm (3.8 on average), yet their rating dropped to 1.8 on average showing that they no longer felt safe after using the prototype. As regards the no-awareness group, the Swarm history had a less significant impact on their sense of safety before and after using it (Friedman Chi-Square = 5.000, p = .025). In particular, the change in the participants' attitude to feeling safe in using Foursquare Swarm was not significant either before or after using Swarm History (ordinal regression coefficient= -1.102, p= 0.146). They still generally felt safe in using the application (before: 3.8, after: 3.4). Figure 7.20 presents the results of the participants' ratings of the sense of safety that they felt

Table 7.5: Average score for usability statements..

Scope	Usability Statements	Score
Specific to Geo-Profile Visualiser prototype	It was simple to use the tool	4.8
	This tool provides me with easy access to my location profile (e.g. favourite places, co-locations with friends and my visit routines)	4.8
	This tool organises and presents my location profile in an effective way	4.7
	The graphic representation of my data in this tool helps me to understand the content easily	4.5
	The map representation of my data in this tool helps me to understand the content easily	4.8
	I would use such a tool to learn about and manage my online data	4.5
Standard SUS	I think that I would like to use this system frequently.	3.8
	I found the system unnecessarily complex.	1.2
	I thought the system was easy to use.	4.8
	I think that I would need the support of a technical person to be able to use this system.	1.5
	I found the various functions in this system were well integrated.	4.2
	I thought there was too much inconsistency in this system.	1.6
	I would imagine that most people would learn to use this system very quickly.	4.6
	I found the system very cumbersome to use.	1.3
	I felt very confident using the system.	4.6
	I needed to learn a lot of things before I could get going with this system.	1.5

using Foursquare Swarm before and after the actual experience of using the tool provided for the tasks.

There was no significant impact in either group on the level of concern over privacy among the participants either before or after using the tool provided. In other words, the participants were generally still concerned about online privacy. However, those in the awareness group showed a slightly greater increase in their level of concern (before: 4.2 on average, after: 4.4 on average) compared with the members of the no-awareness group (before: 3.8 on average, after: 3.7 on average). Figure 7.21 demonstrates the participants' ratings regarding their concern over online privacy pre and post the actual experience of using the tool for the tasks. Asking the participants whether they minded sharing with other people than their friends the location profile presented showed that both tools had a significant impact on their sharing decision and also had a large effect size (Pearson Chi-Square=13.516, p=.004, Cramer's V= .721). The participants

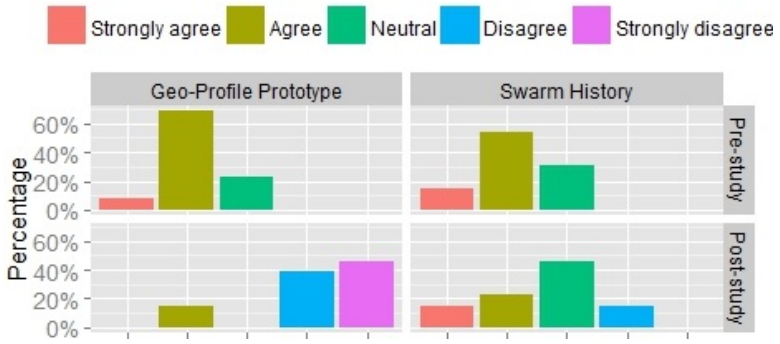


Figure 7.20: Participants' sense of safety using Foursquare Swarm before and after the actual experience of using the tool in both groups..

who accessed their location profile using the Geo-profile Visualiser prototype strongly minded sharing it with others (1.1 on average), while those who accessed their profile using Swarm History tended to be generally neutral about sharing it (2.7 on average).

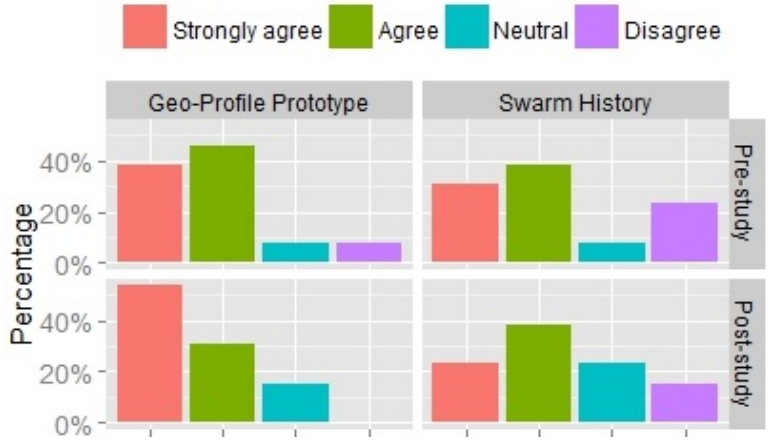


Figure 7.21: Participants' level of concern over online privacy pre and post the actual experience of using the tool in both groups..

The members of the awareness group were asked particular open-ended questions to discover in more detail how the Geo-profile Visualiser prototype had impacted on their privacy awareness and possibly their location sharing behaviour. All of them agreed that at first before using the prototype, they had had a limited understanding and awareness of the detailed collection of their data and also of the related privacy implications including possible information inferences, which seemed to shock some of them. For example, P14 said "I was not aware of such detailed data collection and extraction. I thought it was just sharing place and time. It is scary that other people can know the visit pattern at my house". P16 added "I thought it was just check-ins. I was not aware that it can lead to deriving my patterns or favourites. I now understand location sharing implications better". In particular, four participants criticised the application for not supporting privacy awareness. For instance, P22 said "It is helpful to improve my awareness

and to know better. I am more concerned because people can get my data and know my patterns such as going shopping Sunday morning and here I am. Swarm History is limited and not detailed enough. It is not helping me to be aware." P19 also reported "Now I am more aware. Swarm is very vague about data collection and use. Even privacy settings are enabled by default. It should force us to look at them before using the app."

In addition, almost all the members of the awareness group stated that they would change the way they shared location data. Figure 7.22 shows their responses when asked how they would change their sharing behaviour after being coded and clustered. The most often chosen strategy for protecting their privacy was to check-in less frequently and not check into private places such as home. Other strategies mentioned were, for example, not to check into routine places, to check-in only in a general area or if it was their first visit to a place, and to check into a fake location. In fact, four participants mentioned that they might stop sharing their location altogether since it is "unnecessary information about me," as P19 argued. Only P20 showed no willingness to change sharing behaviour and said "I would continue to share as before. I am not a celebrity, self-conscious about what to share in the first place."

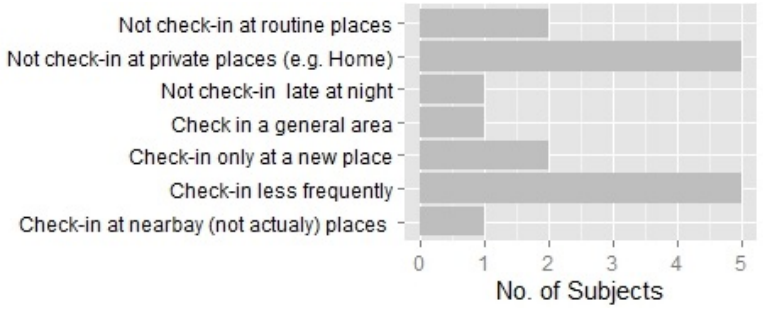


Figure 7.22: Participants' responses towards changing their sharing behaviour (clustered)..

Furthermore, all of the participants showed an interest in using the Geo-Profile Visualiser mainly to explore their profile by viewing what was collected or extracted about them; they wanted to use this information to manage their privacy by learning what to share. For example, P26 said "To explore what data about me the app collected and to see what information can be inferred about me to try and manage my privacy." P19 added "I would use it right at the start of using Swarm to see what they collected and what could be extracted, and to learn about and modify the way I check-in. Then, once in a while after that, to see how my mobility changes and see if I need to change anything".

7.7 Discussion

In this section, the study implications are discussed in terms of their validity and quality together with the impact of the information content and presentation on users' privacy awareness and attitude.

7.7.1 Validity of The Study and Its Outcomes

The validity and quality of the study were taken carefully into account in its design. First of all, the platform used in this experiment was a public GeoSNs that is widely used by thousands of people worldwide and not a restricted proprietary application with limited features of interaction. Thus, it reflects the users' authentic experience with GeoSNs and can provide applicable insights. In addition, the sample used in this study consisted of genuine users of Foursquare Swarm (selected as a case study), who had already been using it for a while. No pressure was put on them to use it for the study purposes alone. The participants were diverse in their demographics and generally in their experience online, in particular with Foursquare Swarm. Moreover, the results were based on the participants' shared data online. These factors ensure a yield of accurate feedback, since the participants communicated their thoughts and impressions on their own data about an application that they were familiar with.

Between-subject study design was used to capture unbiased results and a control group that used the data-access tool offered by the Foursquare Swarm was included as a baseline to compare the proposed prototype against. Interviews were also employed to capture the reasoning being the participants' inputs and to observe their behaviour when using the tool provided. The proposed system that provides an interface for accessing user geo-profile derived from GeoSNs (in the form of the Geo-Profile Visualiser prototype) was successfully applied to the data retrieved from Foursquare Swarm and can be generalised on the basis of other GeoSNs as well.

This system shows the users different aspects and complexity of their disclosed location data whether directly collected by the application or inferred based on such collection. It enables users to see their visited places, interests and activities, and co-locations with friends as well as the degree of association with and patterns of these types of information. Such accessibility to the user's geo-profile can answer all of the information extraction questions presented earlier in Section 4.3.3. This Geo-profile Visulisation tool enables users to explore their extracted personal profile based on their sharing actions on GeoSNs to learn more about their mobility habits and be aware of what other users of the service might be able to see and know about them in order to allow them to understand potential privacy implications and manage their location-sharing actions accordingly. Although the implemented prototype used the application API to

retrieve users' location data, as discussed in Section 7.4, friends have full access to the user's disclosed location data and can used them to derive more personal information about the user. Many users do not know all of their friends on the GeoSNs (see Section 5.5.1) which allows for potential privacy risks. Friends of users can also indirectly disclose these users' location by tag them without their explicit consent which reveal more personal informant than what they intend to share. This system shows also the users this unintended location information disclosure. In addition, some users share their location publicly which makes it visible and accessible by other users of the application without the need to use the API. Some of users' location information are set to be public by default without their control. In the case of Foursquare Swarm, for example, a user location is seen by other users who are at that same location or if a user is the most frequented visitor of a place. The Geo-profile Visulisation tool displays this location-based information that is accessible by the public.

7.7.2 Impact of Information Content and Presentation on Privacy Awareness and Attitude

The study results showed that the Geo-Profile Visualiser prototype has very promising potential for enhancing users' awareness of their location data exposure and associated privacy implications, which consequently allows them to make informed location-sharing decisions and effectively manage their location disclosure.

The Geo-profile Visualiser provides an aggregated and centralised view of users' shared location information and related the privacy risks. It not only informs the users who can see their data but more importantly tells them what further personal information can be derived from their location sharing. Informing the users of both their data extraction, whether explicitly collected or implicitly inferred and well as the accessibility of their data by others has a stronger impact on users' privacy awareness and attitude than informing them only about who can access their data. When users were aware of certain information extracted about them that they did not want to be known, they would worry about who could actually view their data. This is evident from the responses from members of the awareness group, who wanted to know exactly who could access or had accessed their check-in data and reported that they were concerned about others finding out their favourites and routines. The participants' privacy awareness was significantly improved and their privacy attitude was also significantly affected, in terms of their concerns, managing their online privacy using the available settings or by other means (e.g. deleting posts or not to check into certain places), and changing their sharing behaviour to protect their privacy. This impact confirms the results of related studies that used visualisation for privacy awareness purposes [33, 122, 123, 124, 92]. Related studies on the privacy awareness of users

of web and social applications usually focuses on one type of awareness, either visibility to others (e.g.[32, 104, 38, 36]) or the data disclosed (e.g.([21, 123, 102, 124]), whereas in this proposed system, both of these aspects of awareness are included together with data inferences from the data disclosed. All of these aspects are accessible to users on one platform as a way of particularly addressing the awareness of location privacy in GeoSNs.

The tool provided by GeoSNs for users to access their data (Swarm History, in this case) shows nothing beyond what the users have chosen to share and hence hides from them the full awareness of exactly what data about them is collected and what more information can be revealed about them or who exactly can access their data, as demonstrated in the results. The current means supplied by the application allows for a partial and relatively vague view of users' data, as participants reported. Users have a right to full access to their information, whether shared voluntarily or inferred. The tool offered by the GeoSNs – Foursquare Swarm – focuses on "game-playing" and "fun" aspects, as the participants mentioned, which actually encourages them to share more information about their location and diverts their attention from many possible privacy implications.

The participants in both groups were keen to know exactly which of their friends viewed their data. They were more interested in knowing who else, including third parties, could access or had accessed their location data. This detailed level of accessibility information is held by the application itself and is capable of telling users this. The only way that the Geo-profile Visualiser prototype can be improved in this respect is by showing the user who besides their friends has the right to view their data. This is stated in the application privacy policy but usually in a general and unclear way. In addition, some participants (about three) had often checked-in places by their friends. They were surprised to learn this and annoyed that these co-check-ins had usually been done without their direct consent. When asked about their awareness of sharing their data with third parties, many of them stipulated that they wanted to be asked first for permission.

The design also plays an important role in achieving a proper level of privacy awareness. The participants' feedback was in favour of the proposed design, which helped them to find information about their privacy easily and unequivocally. They mentioned that their profile was sectioned and presented clearly. The findings emphasise the recommendations for the design of a privacy-sensitive GeoSNs suggested in Chapter 5 Section 5.6. In addition:

- Users should be immediately Informed about any information about them shared by others using user tags to allow them take action to manage their privacy as they pleased.

- Users should be able to select individually which friends to share their data with.

- Users should be asked for consent to share their data with third parties; they also should be informed whenever a third party accessed their data and shown exactly what data were shared and what they would be used for.
- The GeoSN should inform the user whenever the privacy setting is updated and the options provided should be disabled by default in order to force the users consciously to give their consent for all practices listed in the settings.

7.8 Conclusion

In this chapter, another design for addressing location privacy awareness in GeoSNs is proposed. This system provides users with an interface for accessing their geo-profile extracted from their available data on GeoSNs. It focused on showing them the privacy implications of their location data disclosure in terms of personal information that could be inferred and their data visibility to others. Initial usability testing and a focus group were arranged to ensure that the system could achieve its goals and to improve it as needed. A user-based study was carried out to run a full evaluation of the system using subjects who already used GeoSNs and supplied their own shared data.

The proposed Geo-Profile Visualiser showed that users' privacy awareness can be significantly enhanced in terms of being able to know more about their data disclosure and the related privacy risks. This knowledge consequently demonstrated a marked potential to impact users' attitude to privacy by enabling them to take appropriate actions to manage their privacy. The findings suggest that users need more privacy-oriented GeoSN tools giving them a full view of the data about them that are collected and extracted. The findings also indicate that users should enjoy more transparency about the practices carried out on their data and should actively consent to them.

Chapter 8

Conclusion and Future Work

This chapter concludes the thesis by summarising the main contributions in relation to the hypothesis , and discussing possible future work.

8.1 Thesis Summary and Contribution

GeoSNs provide many useful and convenient ways to communicate with others and find information easily and enjoyably. Users can be profiled based on their interactions and submissions to these applications. Such profiling can offer more personalised services to users, yet it can also pose many threats to their privacy. A review of related literature presented in Chapter 2 revealed that privacy feedback proved to be useful in enabling user awareness in web and social networking applications, yet the design of privacy feedback in GeoSNs is needed. The thesis hypothesis was:

The lack of personal location privacy on GeoSNs can be addressed by enhancing user awareness of the information they share and its implication on their personal privacy. A framework for the storage and presentation of personal location information as well as its privacy implications needs to be supported by GeoSNs. The effectiveness of the framework can be evaluated by measuring its impact on users' attitude and behaviour when interacting on these networks.

In order to investigate the hypothesis and fulfil the objectives presented in Section 1.2, a number of steps were required. The main aim of this work has been to enable location privacy awareness by supporting privacy-oriented GeoSNs. To achieve this aim, we needed in the first stage of work to study the link between location information disclosure and risks to privacy, and assess users' awareness in this matter. Thus, the privacy implications of location- sharing in GeoSNs were first examined in relation to main factors contributing to the location privacy problem including data collection, accessibility, exploitation, and security to gain a better understanding of the extent of location privacy issues, as discussed in Chapter 3. In addition, potential threats to privacy required to be explored for individual users, in particular, to further investigate and validate the privacy implications of location disclosure on GeoSNs. Hence,

statistical analysis was conducted to show the possible derived information from typical datasets collected by GeoSNs for different types of user. The above investigations mark the first contribution in this work. Results showed that there is high feasibility of the inference of rich personal information about users and their mobility, including spatiotemporal movement tracks and patterns, absence and presence in places, degree of association with specific places, and co-location patterns with other users. They also demonstrate the need for examine user awareness of such potential privacy risks. To address this need, A user-based study was carried out to evaluates the degree of user awareness and attitudes to privacy implications of sharing location information on GeoSNs using representative samples of users of these applications. It included assessing user knowledge of privacy terms and conditions commonly used on GeoSNs that state practices carried out on their data, their awareness and reaction towards possible personal information inferences, and their preferences to control access to their personal information. This study is presented in Chapter 4 and represent the second contribution of this wok. Findings demonstrated that users show a significant lack of knowledge about how their data is used by the applications and suggest that users do not fully appreciate the privacy risks resulting from location information disclosure on GeoSNs. The study revealed that there is a strong necessity to improve users' awareness of their collected and derived information, how it is stored, and to have the ability to control visibility of their location data. In this stage, the first two objectives in regard to researching the problem of location privacy awareness on GeoSNs and assessing the lack of awareness of privacy implications by users when using these networks are fulfilled.

Guided by the outcomes of the first stage which mainly stress the need for enhancing the visibility of users' profile information on GeoSNs, the second stage of this work focused on proposing and evaluating feedback design solutions based on the limitations identified in the first stage by providing explicit presentation of location information exposure and associated locations privacy risks. Firstly, we proposed and evaluated a privacy feedback design based on investigating the level of awareness with respect to extended user geo-profiles to preliminary test the impact of location awareness on users' attitude and sharing behaviour, as demonstrated in Chapter 5. This marks the third contribution of this work. This feedback provides real-time privacy notifications that are related to a specific location-sharing action, and offers some controls over information disclosure in order to enable users to make informed decisions before they publicise their location information. It offers awareness in regards to the visibility of data to other users and content extracted, and indicate an threat level to privacy based on these aspects. To ensure the usability and effectives of feedback, the design aspect was considered to optimise it. Evaluation was conducted through a user-based experiment for assessing its impact on privacy attitudes and potential location-sharing behaviour using realistic scenarios and mock-ups of the Foursquare Swarm application. The results indicated that users' privacy awareness of their information disclosure was significantly enhanced by the presentation of the associated threats

to their personal privacy in the privacy feedback. Their sharing behaviour was significantly impacted as well based on the provided awareness where they were able to make informed location-sharing decisions that includes in some situations limiting location information they share or deserting it all together. This experiment outcomes generally highlighted the promising potentials for employing privacy feedback to improve users' location privacy awareness and ultimately their behaviour. In addition, the initially proposed threat levels to location privacy showed to be effective in influencing users' attitude and their ability to manage their data with respect to privacy. Therefore, a data-driven approach was proposed to objectively model levels of threat to location privacy for providing effective location privacy notification through investigating users' perceptions towards location privacy implications, covered in Chapter 6 and linked to the fourth contribution. In particular, an extensive user-based experiment was carried out to study users' location sharing behaviour on GeoSNs in regard to three factors: the dimensions of the exposed data, data visibility to other users of the application, and users' awareness of potential privacy implications resulting from data disclosure. Diverse location-sharing scenarios were used without indication to privacy levels or mapping to certain GeoSNs to eliminate any potential biasing. The results confirmed that all of these factors, including sensitivity of the location, have a significant impact on users' sharing decisions and appreciation for their privacy. Users were less willing to disclose their location and hence more concerned about their privacy in the situations that pose more risk to their privacy, especially when they are aware of the hidden implications. The findings of this experiment were then utilised to model level of threat to location privacy which can be incorporated in privacy-aware systems to improve users' privacy awareness and ultimately allow them to make informed location-sharing decisions. Although such a model is based on the input from large sample of representative GeoSNs' users, sense of privacy may vary among individual and hence a privacy-aware system can use the propose model as a default and then adapts based on the individual privacy preferences.

As the above experiments in the second stage demonstrated the effectiveness of providing detailed location privacy feedback in enhancing uses' awareness and ultimately enabling them to take proper actions to protect their privacy, we noted that there is still a need for validating these effects using real users' data with an implemented system that provide such privacy feedback. Thus, we proposed and evaluated a usable and privacy-oriented interface that provides access to users' profiles resulting from their location sharing on GeoSNs, whether explicitly collected by the application or implicitly inferred, discussed in Chapter 7 and represent the fifth contribution. The design of this interface focuses on supporting user awareness of the potential privacy implications of their location disclosure on these applications, and taking into account the usability aspect. A prototype was implemented for evaluation, and Foursquare Swarm was used as a source of real users' data. Evaluation was run as a comparison between participants who used Foursquare Swarm (no-awareness group) and other participant who used our implemented pro-

totype (awareness group) for accessing and finding out about their own information. Observing and interviewing the participants revealed that the privacy awareness of the 'awareness' group was significantly improved in terms of knowing their data disclosure and the related privacy risks, whereas there was almost not improvement in the 'no-awareness' group. The enhanced awareness enables them to effectively manage their privacy by modifying how they share location information according to their privacy preferences. For example, they would reduce check-in frequency in certain places to hide their mobility routines. These three experiments carried out in the second stage address the third and the final objectives related to identifying methods for extending or modifying the design of GeoSNs to enable user awareness of their information and privacy threats associated with sharing this information, and evaluating the effectiveness and usability of these proposed solutions. By following the above research methodology, we evaluated the hypothesis to be true. Lacking personal location privacy on GeoSNs can be addressed by enhancing user awareness of the information they share and its implication on their personal privacy, using a framework (privacy feedback design) for the storage and presentation of personal location information. This framework showed to significantly impact users' privacy attitude and behaviour. Thus, GeoSNs need to support such a privacy-aware framework that enhances users' location privacy while enjoying the services provided by such applications.

8.2 Reflection on Some Related Issues

Sharing location information on GeoSNs provides many useful services and convenient benefit for the users. For example, finding the best rated Italian restaurant nearby or telling their friends where they currently are. However, location disclosure can lead to undesirable privacy implications such as revealing users' absence from their private places or their movement patterns. That does not mean that location information should not be shared at all on these networks. Users can always balance the trade-off between benefit of sharing location and privacy implications. This is where the importance of location privacy awareness (feedback) methods lays. Presenting users with location privacy feedback informs them of the potential risks (and its level) of their sharing actions in terms of what personal information can be inferred about them and who can see or access their data. Such feedback allows users to make informed decision when to share or not according to their own privacy preferences. Therefore, they can selectively share their locations while enjoying the service. Users at least have the right to know any possible privacy compromises they make by sharing their location information. Users generally would still share their location information as long as they gain benefits from doing so and would only hide locations that they consider threat to their privacy Nevertheless, the service providers of GeoSNs should not be significantly affected by enhancing location awareness as the majority of users would continue sharing their location but based on their sense of privacy which varies

among individual. Thus, there would not be a particular aspect of location sharing (e.g. place type or sharing time) that would be heavily influenced by the general use of GeoSN applications as a result of improving users' awareness. In addition, other users of the application would still have access to plenty of location-based information that can fulfil their needs.

The availability of users' location data within a GeoSN and hence the extent of privacy implications depends on two main factors. The first is the type of GeoSNs used. In LBSNs, all of users' shared information revolve around location data, whereas location data can represent only a limited portion of users' shared information in LESNs as location is not the core of these applications. Hence, using LBSNs can pose more serious privacy threats, such as revealing mobility routines and relationship nature with friends, than LESNs since the former can lead to construct denser and more diverse location history of users. The second factor is related to how frequently users use the application and share their location. More use of location services in GeoSNs would results in denser spatiotemporal history of user data collected and the greater certainty and diversity in the information extracted from this data. At the same time, sparsity of data resulting from less use of the application can impact the density and the certainty of the derived information. This sparsity can occur in the spatial aspect where a user visits limited number of places, in the temporal aspect where a user visits places in certain times only, or in both aspects. Consequently, it might be not possible in such situations to infer users' patterns or extract wider range of their interests and activities which in turn enhances their privacy. however, having sparse data do no means that there are no privacy implications of disclosing location. For example, a user can check-in to her home location which is considered a sensitive place for most users. Thus, users of GeoSNs need to be informed about their data collection and potential privacy implications of their interaction with the application and allow them to decide upon what of their sharing actions can cause risk based on their own privacy preferences.

8.3 Future Work

This work provides insight into means of enhancing location privacy management on GeoSNs through supporting privacy awareness. Future work can further explore a number of areas, including:

- **Investigating the effect on users' perception of privacy when using more detailed privacy settings**
  Some studies have proposed rule-based privacy settings that are mapped particularly to offer more precise control over location privacy in location-sharing applications (see Section 2.6.1.2). They demonstrated that using location-oriented privacy settings can impact

users' perceptions of privacy and hence their sharing attitude as well. Thus, investigating the impact of using more detailed privacy settings that are tailored to address the nature of location-sharing in GeoSNs on uses' privacy attitude and behaviour would provide useful insights for enhancing users' location privacy in these networks.

- **Examining how does this work can be generalised to other GeoSNs**
  In this work, both LBSNs and LESNs as types of GeoSNs were considered when examining and proposing solutions for the problem of location privacy. Few examples of GeoSNs (particularly Foursquare and Twitter) were used for the evaluations which was sufficient to test and demonstrate the needed effect. Nevertheless, the interaction features and environments can slightly vary across different GeoSNs. Therefore, it would be interesting to examine how this work can be generalised to other GeoSNs and what the needed amendments or adaptation, if any, are for this work's proposed solutions and recommendation to be optimised.

- **Conducting longitudinal experiments to observe actual changes to privacy behaviour**
  This work involved user-based studies to examine the influence of the proposed location privacy awareness tool, as presented in Chapters 5 and 7. In particular, they investigated the impact on users' privacy awareness and concerns, as well as their behaviour towards location-sharing and privacy management when using the proposed tools. Further examination of how users' behaviour towards their location sharing while using GeoSNs would change after these users are provided with the proposed location privacy awareness solutions, would be worthwhile.

- **Deriving users' geo-profile from multiple GeoSNs**
  As discussed in Chapter 7, people can be users of more than one GeoSN, hence their geo-profile can be aggregated from their multiple accounts. This worked mainly utilised Foursquare as the data-source for evaluation, since using one source was appropriate for the initial testing of the privacy impact of the proposed solutions. Foursquare also offers a typical representation of a GeoSN. A promising, yet challenging direction for future work would be deriving users' geo-profiles from their multiple accounts across GeoSNs to infer richer personal information, and to study how such a comprehensive view can influence users' privacy attitudes and behaviour.

- **Extracting other types of personal location-based information**
  This work focused on examining how the proposed location privacy awareness tools can affect users' privacy attitude and hence their behaviour using earthier simulated realistic scenarios (Chapter 5), or real data with applied inferences, such as top places and movement patterns (Chapter 7). These methodologies were sufficient for capturing the participants' feedback needed for the evaluating our proposed solutions for location privacy

awareness. Additionally, exploring how users react towards other types of inferences on their data, such as time spent in a place, friendship detection, and future movement and place transition predictions could provide valuable insight.

Appendix A

The Survey Questionnaire on Examining Users' Privacy Awareness, Concerns and Attitude in GeoSNs

This survey aims to examine the privacy concerns of users of online social networks, in particular users' concerns towards their location information. Therefore, in order to participate in this survey, you should have used the location features on these networks (e.g. added location to your posts/pictures or checked into places).

Three main aspects are addressed in this questionnaire including your awareness for these services' terms of use, how you feel about your personal information being used by these applications, and how you might prefer to control access to your personal information on these applications.

This survey is done as part of research undertaken in Cardiff University, UK, to study the privacy implications for users of Social Networking applications. No personal information are collected in the survey and the results of the survey will only be used for research purposes.

It will take no more than 10 minutes of your time. Your participation is much appreciated. Thank you.

You will have a chance to win one of $\pounds 10$ vouchers when you complete this survey The survey will end on 9th May and the winners will be then randomly selected and contacted. You need to enter your email address by the end of this survey.

A.1 Use of location information on Social Networks

1) Please select your age category

○ 14 or younger

○ 15-24

○ 25-34

○ 35-44

○ 45-54

○ 55-65

○ 66 or older

2) Please select your gender

○ Male

○ Female

3) How often do you use Social Networking applications?

○ Frequently (several times a day)

○ Moderately (several times a week)

○ Occasionally (once a week or less)

4) How often do you use location information on Social Networks?

Service	Always	Sometimes	Never
I add locations to my posts and pictures on Facebook	○	○	○
I add locations to my tweets on Twitter	○	○	○
I use location to tag my pictures on Flickr	○	○	○
I add locations to my posts or I check-in Google+	○	○	○
I photo map pictures on Instagram	○	○	○
I check-in on Foursquare	○	○	○
I check-in on Yelp	○	○	○
Other services	○	○	○

5) Have you linked any of your accounts on these applications - for example sharing your tweets on Facebook?

○ Yes

○ No

A.2 Your knowledge of Terms of Use and Privacy Policies for Social Networking Services

1) Have you read the Terms of Use and Privacy Policy documents for the applications that you use?

   ○ Yes

   ○ No

2) Are you aware that the terms and conditions of the applications that you use may contain the following terms of use:

   - The application collects and stores your precise location (as a place name and/or a GPS point), even if you mark your location as private, for a possibly indefinite amount to time.

     ○ Yes

     ○ No

   - The application can use your location information in any way possible including sharing it with other applications or partners for various purposes (commercial or non-commercial).

     ○ Yes

     ○ No

   - If you share your location information, your friends and any other users are able to access and use it in any way possible.

     ○ Yes

     ○ No

   - The application can collect other personal information, such as your personal profile information and browsing history from other web applications.

     ○ Yes

     ○ No

A.3 Perceptions of Possible Inferences of Personal Information

Consider the following statements. Each details an aspect of personal information that social networking applications can tell about you as an individual, given they have access to your location information.

For each statement, select whether or not you were previously aware this type of data could be found and used, and then state how comfortable you are with social networking applications being able to find out this information about you.

1) I can guess where you your home is.
   Your awareness

   ○ I am aware that this statement is possible, or,

   ○ I did not think that this statement was possible.

   Your reaction

   ○ I am ok with this statement.

   ○ I am uncomfortable about this statement.

   ○ This statement makes me very worried.

2) I can guess where your work place is.
   Your awareness

   ○ I am aware that this statement is possible, or,

   ○ I did not think that this statement was possible.

   Your reaction

   ○ I am ok with this statement.

   ○ I am uncomfortable about this statement.

   ○ This statement makes me very worried.

3) I know which places you visit and at what times.
   Your awareness

   ○ I am aware that this statement is possible, or,

   ○ I did not think that this statement was possible.

Your reaction

- ○ I am ok with this statement.
- ○ I am uncomfortable about this statement.
- ○ This statement makes me very worried.

4) I can tell where you normally go and what you do in your weekends.
Your awareness

- ○ I am aware that this statement is possible, or,
- ○ I did not think that this statement was possible.

Your reaction

- ○ I am ok with this statement.
- ○ I am uncomfortable about this statement.
- ○ This statement makes me very worried.

5) I can tell you where you go for lunch or what you do after work.
Your awareness

- ○ I am aware that this statement is possible, or,
- ○ I did not think that this statement was possible.

Your reaction

- ○ I am ok with this statement.
- ○ I am uncomfortable about this statement.
- ○ This statement makes me very worried.

6) I know your favourite store (your favourite restaurant, your favourite coffee shop, etc.)
Your awareness

- ○ I am aware that this statement is possible, or,
- ○ I did not think that this statement was possible.

Your reaction

- ○ I am ok with this statement.
- ○ I am uncomfortable about this statement.
- ○ This statement makes me very worried.

7) I can guess what you do when you are in a specific place.
Your awareness

○ I am aware that this statement is possible, or,

○ I did not think that this statement was possible.

Your reaction

○ I am ok with this statement.

○ I am uncomfortable about this statement.

○ This statement makes me very worried.

8) I can guess when you are AWAY from home.
Your awareness

○ I am aware that this statement is possible, or,

○ I did not think that this statement was possible.

Your reaction

○ I am ok with this statement.

○ I am uncomfortable about this statement.

○ This statement makes me very worried.

9) I can guess when you are OFF work.
Your awareness

○ I am aware that this statement is possible, or,

○ I did not think that this statement was possible.

Your reaction

○ I am ok with this statement.

○ I am uncomfortable about this statement.

○ This statement makes me very worried.

10) I know who your friends are.
Your awareness

○ I am aware that this statement is possible, or,

○ I did not think that this statement was possible.

Your reaction

- ○ I am ok with this statement.
- ○ I am uncomfortable about this statement.
- ○ This statement makes me very worried.

11) I know when and where you meet up with your friends.
Your awareness

- ○ I am aware that this statement is possible, or,
- ○ I did not think that this statement was possible.

Your reaction

- ○ I am ok with this statement.
- ○ I am uncomfortable about this statement.
- ○ This statement makes me very worried.

12) I can guess which of your friends you see most.
Your awareness

- ○ I am aware that this statement is possible, or,
- ○ I did not think that this statement was possible.

Your reaction

- ○ I am ok with this statement.
- ○ I am uncomfortable about this statement.
- ○ This statement makes me very worried.

13) Other people can know where you are at any point in time.
Your awareness

- ○ I am aware that this statement is possible, or,
- ○ I did not think that this statement was possible.

Your reaction

- ○ I am ok with this statement.
- ○ I am uncomfortable about this statement.
- ○ This statement makes me very worried.

14) Other people can know what you are doing at any point in time.
Your awareness

○ I am aware that this statement is possible, or,

○ I did not think that this statement was possible.

Your reaction

○ I am ok with this statement.

○ I am uncomfortable about this statement.

○ This statement makes me very worried.

A.4 Your Attitude to Privacy on Social Network

Given the statements you have just witnessed about the data social networks may be able to find out about you:

1) Would you consider changing the way you share your location information?

○ Yes

○ No (Go to the next section)

2) How would you change the way you share your location information?

○ I will share my location less frequently.

○ I will stop sharing my location information.

A.5 Your Attitude to Controlling Your Personal Information

If you were able to control the way your location information was used by social networking applications, please select which things you would like to be able to control.

1) I would like to be able to turn off location sharing for specific durations of time.

○ Yes

○ No

2) I would like to turn off location sharing when I visit specific types of places.

○ Never

○ Occasionally

○ o All the time

3) I would like to decide how much of my location information history is stored and used by the application; for example use only my check-in history for the last 7 days.

○ Never

○ Occasionally

○ o All the time

4) I would like to see the predicted personal information that the application stores about me based on my location information .

○ Never

○ Occasionally

○ o All the time

5) I would like to decide how people see my current location; for example, exact place name, or a rough indication of where I am.

○ Never

○ Occasionally

○ o All the time

6) I would like to decide who can download my location information data.

○ Never

○ Occasionally

○ o All the time

7) I would like to know, and control, which information can be shared with other Web applications.

○ Never

- ○ Occasionally
- ○ o All the time

8) I would like to make my location information private; seen only by myself and by the people I choose.

- ○ Never
- ○ Occasionally
- ○ o All the time

Appendix B

Examination of Information Presentation tools and The Remaining Tasks' Analysis

In this appendix, a detailed examination on the tools available for information presentation and visualisation on GeoSNs is provided, along with the analysis results of all location-based tasks provided by Foursquare/Swarm, the selected case study, is presented.

B.1 Information Presentation and Visualisation Tools Available for Users

Investigating how users can view and retrieve their information submitted on GeoSNs can provide insights into the extent of support for privacy awareness which enable for assessing users' current awareness and hence the need for enhancing it as a route for addressing location privacy on these networks. Generally, users of such a service have three main ways to access their information namely, via the user interface, using tools provided by the service and finally using the service API. The user interface is the most basic means that allow users to view their information and interact with the service. Through the user interface, users can view and edit their profile information as well as view the content published by them or others form their social network via the main interface of the service. Users also can use the service API that provides high-level access to almost all of their data in more structured and directed way. These APIs can be utilised by linking them to any computer program. One simple way to use them is by using Apigee Console [1] which offers a web-based interface for accessing the most common web services and social networks.

However, these tools do not provide an interactive and comprehensive presentation of users' location-related data available online or project to the users potential privacy implications of

[1] `apigee.com/console` [Accessed: July, 2016]

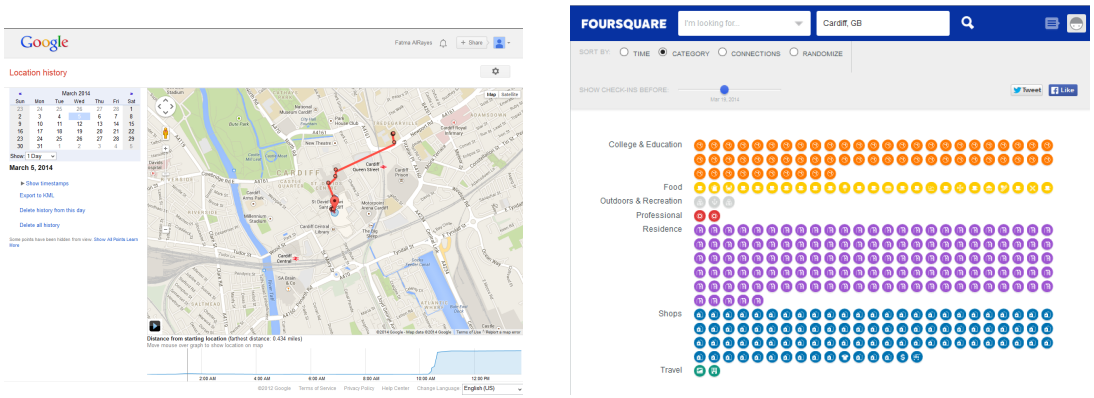
their information disclosure. This section aims to examine and explore the available presentation methods of location data shared on GeoSNs.

B.1.1 Provided by the Service

Here, available data presentation tools offered to users by two examples of common online service providers are explored. Google enables their users to access and manage their data distributed across their various services through Google Dashboard [2]. In the Dashboard tool, all of Google services associated with a certain user are aggregated and presented in one uniform interface. It shows a preview of the important information as well as links for managing the setting and privacy, and viewing the service for each of the different Google applications used by a user. In particular for the location information, Location History service can be accessed via the Dashboard where users are presented with an interactive interface that visualises their map-based mobility tracks if they enabled location history on their mobile devices as illustrated in Figure B.1(a). These tracks as well as the timestamps for locations are shown for each date selected and the user has the option to download them. Such a sophisticated web platforms that provides many essential web services for users collects diverse and large amount of data about the users. However, Google is not providing powerful tools for the users that enable them to adequately view all of their data collected and understand how they can accessed and utilised. The Dashboard just facilitates accessing the different Google services and altering their settings. It is only showing users their mobility tracks without any indication towards the type and amount of personal information inferences that can be made, or who can view their tracks which may jeopardise their privacy.

Foursquare is also one of the most popular location-based social networks which allow users to access and view their check-ins data using in several means. Using the main service application, users can access their check-in history where they can view text-and map-based check-in history. They can search for their visited places based on the month, area, location, category of the place as well as which person they have been with. Nevertheless, the map is displayed in small box and it is inconvenient to navigate through, whereas the text feed of the history is just a listing of the places that a user has been to. In addition, Foursquare developed a visualisation tool, infographics [3], for users where the can view their check-ins places which can be clustered by user based on time of check-in, place category, connection between places and also view randomisation of the check-ins as shown in Figure B.1(b). However, It does not aim for providing comprehensive privacy awareness as it is still offers showing a small window of their information that can be retrieved.

[2] `www.google.com/settings/dashboard` [Accessed: August, 2016]
[3] `foursquare.com/infographics/4sqday` [Accessed: Oct. 2014]



(a) (b)

Figure B.1: (a) A snapshot of Google's Location History. (b) A snapshot of Foursquare's infographics..

B.1.2 Provided by Third Parties

This section is concerned with examining information presentation tools provided by third parties of GeoSNs or individual developer and available to the end user as a way of enabling them to view their online-shared location information. There are mainly three public projects, namely, Foursquare Timemachine, 4sqmap.com and Creepy. These project are reviewed next where their purpose, functions offered and limitations are discussed.

B.1.2.1 Foursquare Time Machine

Foursquare Time Machine [4] is an interactive visualisation project that utilise by the available check-ins information of Foursquare's users to show them their check-in history as map-based video trip in starting from the first check-in along with some basic statistics.

It is a result of a collaboration between Foursquare and Samsung to produce the project for the aim of advertising for Samsung Galaxy S4 [5] where their logo can be seen in this project's website. Foursquare Time Machine mainly focuses on presenting users' check-in tracks in an interesting manner. It was running from June, 2013 until the late of 2014 (not available anymore).

Functions Provided:

The followings are the functions that are provided by this system to the user:

[4] `foursquare.com/timemachine` [Accessed: Oct. 2014]

[5] `http://business.foursquare.com/special-projects/foursquare-samsung` [Accessed: Aug. 2016]

1) The system enables the users to log in with their Foursquare accounts in order to fetch their check-in data.

2) In 'My History' tap, the system shows the users a video trip of their spatiotemporal check-in history on a map that is placed on the centre starting from their first check-in.

3) The system provides statistical graphs of users' check-ins that are presented on the left side which dynamically calculate top places and categories visited as well as check-in percentage of each day of the week while the check-ins history is running.

4) The system enables the users to pause the check-in history trip and navigate forward and backward to view their check-ins in an individual basis while details of the selected check-in is shown on the left panel.

5) In 'The Next Big Thing' tap, the system offers users with predictions in form of a list of places that they may be interested in visiting based on their history, as it offered in form of recommendations on Foursquare.

6) In 'Share My Stats' tap, the system generates an infographic that include the user's check-in history in a graphical and statistical form including top places and place categories on several temporal periods which can be shared on other social networks.

User Scenario and Evaluation:
In this section, a user-based evaluation is conducted by going through Foursquare Time Machine and using its functions in order to assess to what extent the presented information can improve users' privacy awareness about the possible personal information inferences and accessibility by others.

When first the user launches this tool, the user should click on 'Fetch My History' button in order to start using the tool by allowing it to access all of the user data available on Foursquare as shown in Figure B.2. The presented introduction of this seems to promote for the interestingness and usefulness of this tool to encourage using it as it *"bring[s] your Foursquare history to life and help you discover great new places"*. This encourages users to use the services with suppressing the fact that there are privacy implications of disclosing such information online.

1) *'My History' Tap*
   Once the users' data are fetched, the tool shows a map where the first check-in of the user is marked as well as the date and time of this check-in is show on the top right corner. The sequential number of this check-ins, the place name and coordinates, photos of the place (if any) as well as the place of the next check-in is presented in the a left-sided panel. In

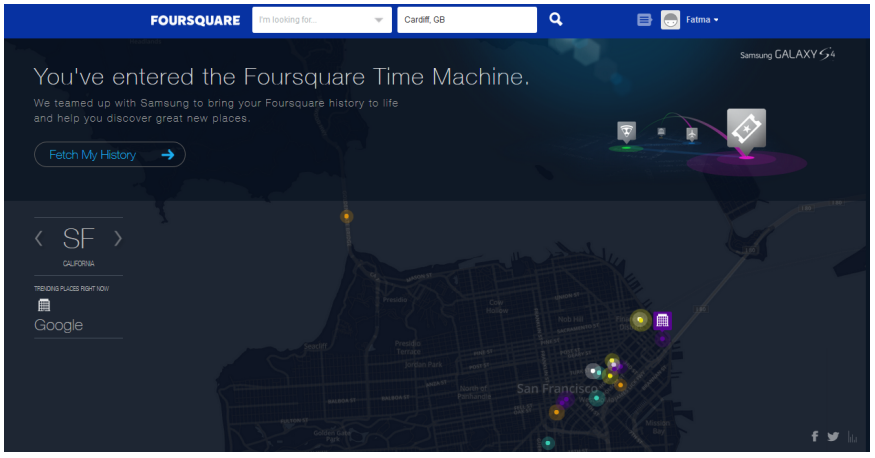


Figure B.2: The main window of Foursquare Time Machine..

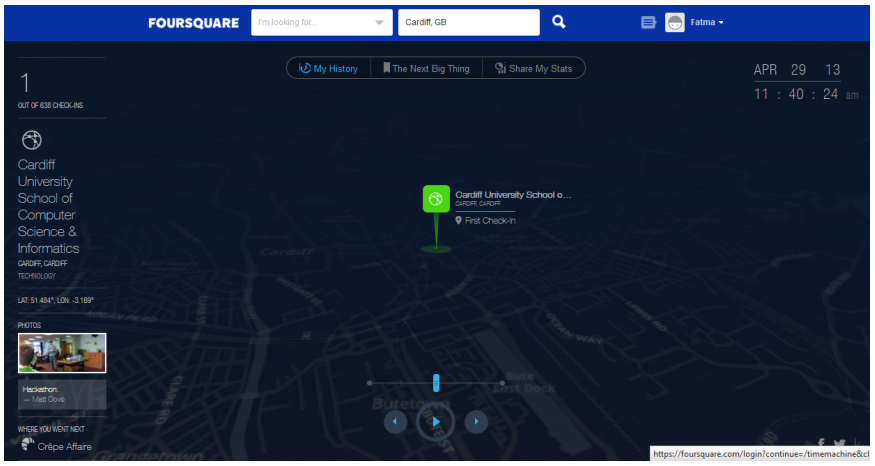


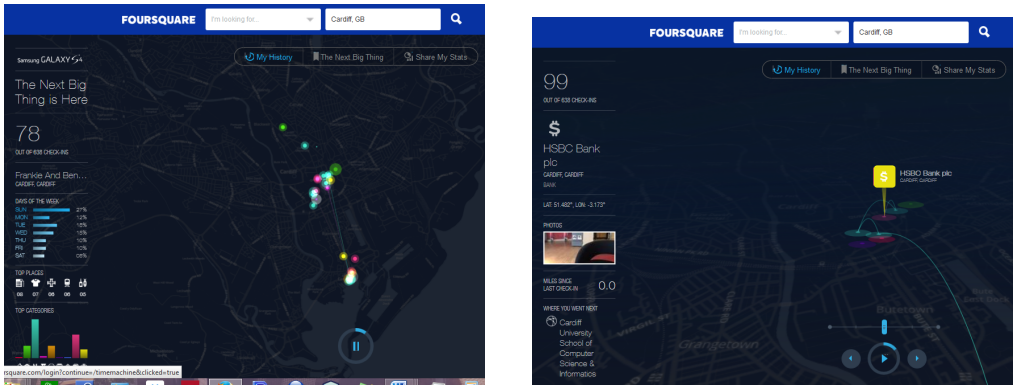
Figure B.3: The first screen shown when fetching the user history..

the bottom of the map a video play bottom is provided to the user is illustrated in Figure B.3.

When clicking on the play button, the tool runs a video trip of the user check-in history by showing sequentially the venues visited by the user while the date and time information is dynamically changed to correspondents the temporal information of the currently shown check-in. Moreover, the tool presents on the left panel statistical graphs of users' check-ins which dynamically calculate top places and categories visited as well as check-in percentage of each day of the week how that reflects how far of the user check-ins is presented. When pausing the history trip, the left panel show information of the current check-in as in the first check-in. The user can also go backward and forward with the history to view the check-in individually as demonstrated in Figure B.4.

Going through the user history shows the location mobility of the user in terms where the user have been to and on which date and time beside showing basic check-in frequency

statistics of based on the day and place. However, the user cannot filter the location history based on temporal or spatial criteria in order to explore the collected location history. Therefore, the user does not have means to have closer look into the data and understand how they can be utilised and hence appreciate the potential privacy implications of location-data sharing. In addition, Foursquare Time Machine might not show all of the user history (as in this case showed only the main area where most of the check-ins happened) which can hide a valuable and sensitive part of the location history that the use cannot see. It seems that presenting the user history motivate the user to enjoy this experience rather than to provide the user with means for discovering the possible use of their data and hidden privacy risks.



(a) (b)

Figure B.4: Snapshots of the tool when running and pausing the user history..

2) *'The Next Big Thing' Tap*

When navigating to this tap, the user is offered with a list of places that this user might be interested in visiting based on their mobility history as shown in Figure B.5. Although this list is created by utilising the users' shared location data, it can be vague to them on what basis these place were recommended, and hence fail to recognised that can pose a threat to their privacy.

3) *'Share My Stats' Tap*

When clicking on 'Share My Stats' tap, the tool generates an infograhic that include the user's check-in history in a graphical and statistical form similar to what is presented in Figure B.6(a) including top places and place categories visited on several temporal periods which can be shared on other social networks. However, in this user case, the tool was frozen many time while attempting to generate the statistics without giving any kind of error massage or feedback on the encountered problem as demonstrated in Figure B.6(b). This infographic presents only some limited statistics that is not be sufficient for providing full privacy awareness.

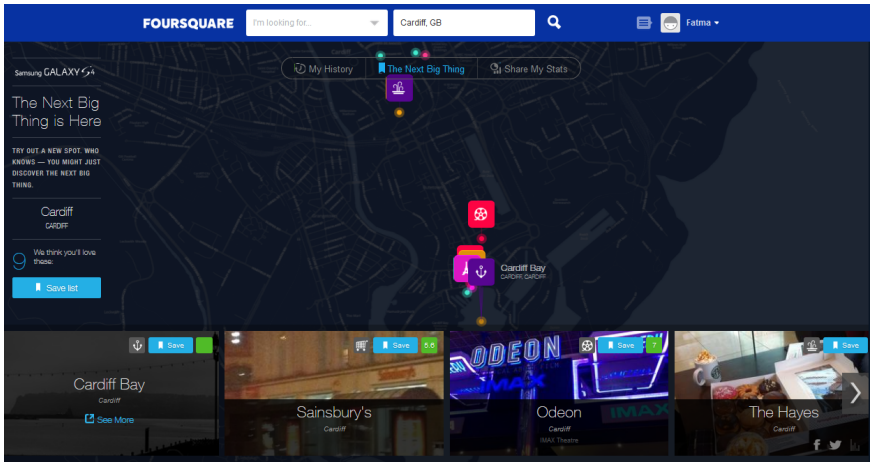
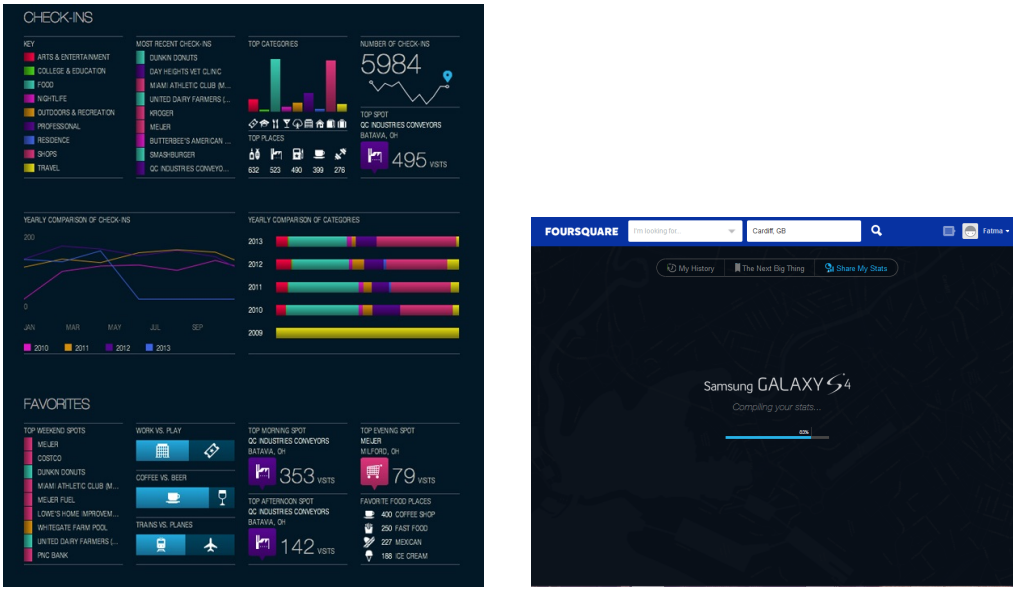


Figure B.5: A snapshot of the 'The Next Big Thing' task..



(a) (b)

Figure B.6: (a) An Example of the infographic. (b) A snapshot of when 'Share My Stats' feature freezes during generating the infographic..

Limitations:

There are some limitations in this project. Firstly, the main purpose of it was to advertise for Samsung Galaxy S4 and it is essentially emphasising on playfulness aspect of the visualisations. It motives users to be more interested in viewing their mobility data in this dynamic and colourful visualisation, and may be even more interested in using Foursquare. In addition, it is presenting relatively rough and overall visualizations of the users' rich location-based information. Moreover, the visualisation can be incomplete as it focuses only on showing the areas of where the user has high check-in frequency in which may not present all of users shared data. Furthermore, this project lack the flexibility to view users' check-in history by specifying

a certain date or period of time as well as the capability to provide more complex visualisations of users' data in terms of extracting visiting patterns and co-occurrence with friends which can provide more sensitive information and hence may provoke their privacy concerns. It is only offer recommendations for place to visit but it is not clear to the users how the system generated these recommendations. Hence, this tool shows only limited view of the users' data and focuses on having enjoyable experiences for the users where for example it encourage them to share their infographics which can contain some of their private data. Another important aspects of the users' privacy is knowing who can access their data which is also not presented in this tool.

Thus, Foursquare Time Machine is not aiming for privacy awareness where the users have the chance to see how sharing location data can be exploited to infer more of their not explicitly disclosed private information and who can access their data in order to increase their awareness of the privacy implications. Hence, this project do not offer a comprehensive means of exploring how their data can be collected and exploited by different parties.

B.1.2.2 4sqmap.com

4sqmap [6] is a webs application that offers Foursquare users visualised feedbacks about their history of usage and their shared data in more interactive and map-based manner by exploiting Foursquare API as well as Google Maps APIs as demonstrated in Figure B.7. At first, users need to log into this website. Then, it retrieves all of their data shared on Foursquare and provides them back to the user in form of visualisation on Google map. The purpose of this system is to visualise users' data shared on Foursquare mainly on a map in order to make them interesting and easy to browse.

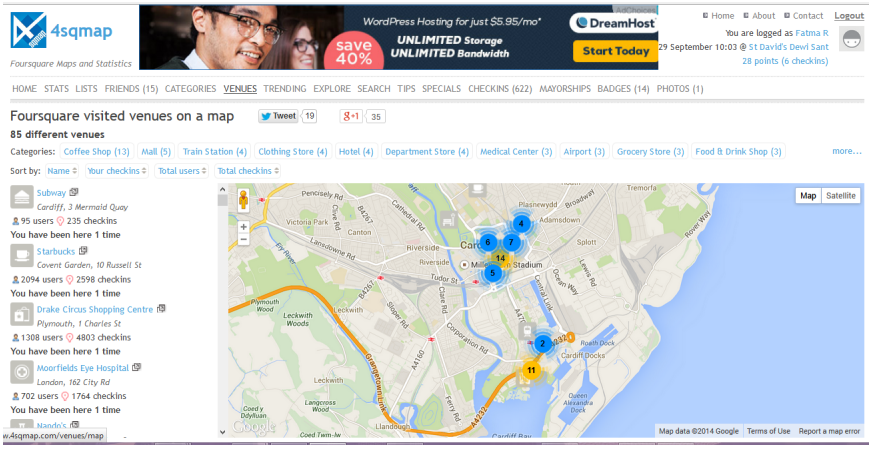


Figure B.7: A snapshot of 4sqmap when viewing the visited venues..

[6] www.4sqmap.com [Accessed: Aug. 2016]

Functions Provided:
The followings are the functions that are provided by this system to the user:

1) The system enables the users to log in with their Foursquare accounts in order to fetch their data.

2) The system provides place recommendations and the special offers available by marking the place on the map and showing the place details under 'Home' tap

3) The system allows the user to generate interactive poster-style stats (infographics) of a user where they are presented with their check-ins on a map, badges as well as top places and categories visited under 'Stats' tap.

4) The system displays the users 'to-do' lists of venues on the map under 'Lists' tap.

5) The system shows the users where their friends at the last 30 minutes, 1 hour, 3 hours, 6 hours, 12 hours, 24 hours or all the time on the map under âĂŸFriends' tap.

6) The system presents the place categories visited by users along with the number of check-ins as well as the number of venues in the category which also be ordered by the number of check-ins or the number of venues under 'Category' tap.

7) The system shows on the maps the venues visited by the users as well as their details under âĂŸVenues' tap where users can sort the venues by the name, the user check-ins, total numbers of visiting users or total number of check-ins.

8) The system displays on the map the trending venues as well as their details in the user current location or any other location entered by the user under 'Trending' tap.

9) The system allows the users to explore venues by searching venue name as well as specifying location and place category where the results are shown on the map under 'Explore' and 'Search' taps.

10) The system enables the user to search tips by their text and displays where the resulted tips at on the map under 'Tips' tag.

11) The system presents to the users the venues that has special offers on the map based on users' current locations and it allows them to search for the special in any other location they enter under 'Specials' tap.

12) The system visualises the users' check-ins on the map where users can specify the time range for these check-ins as well under 'Checkins' tap.

13) The system shows which venues that the users are the Mayor of on the map under 'Mayorships' tap.

14) The system displays where the users badges where unlocked at on the map under 'Badges' tap.

15) The system shows the venues that the users have shared pictures of on the map under 'Photos' tap.

Limitations:
This project seems to provide similar services of Foursquare but in a different approach by focusing on presenting users' data in a more interesting way to view which relies mainly on the map. However, it is not offering any kind of data analyses or Privacy-related information. 4sqmap is not aiming to raise users' awareness of the potential privacy implications of disclosing location information online.

B.1.2.3 Creepy

Creepy [7] is an Open-source intelligence (OSINT) tool for gathering location-related information of users from their shared data on GeoSNs which is available as a desktop application. It collects the users location information form Twitter, Instagram and Flicker by extracting the location from the added location with tweets or geo-tagged pictures in Instagram and Flicker. The derived location tracks these users are then shown on Google Maps. It aims to provide accessible application for showing and retrieving location tracks of any users of Twitter, Instagram or Flicker

Functions Provided:
The followings are the functions that are provided by this system to the user:

1) The system require the user to first configure the connectivity to these three services through a simple wizard in order to be able to use to tool. Hence, users are required to have accounts in the services they want to use and in order to connect with it and authorise this tool for retrieving information.

2) The system enables the user to start using the tool by creating a new project where they can search for other users or their own profiles by email, username, name or id in any of all of these services then select the users who want to retrieve their location information.

[7] `http://ilektrojohn.github.io/creepy/` [Accessed: Oct. 2014]

3) The system allows the user to run analyses on the selected users in order to retrieve their location-related information and plot them on the map which shown in the centre on the tool where the panel on the right side present the text-based spatiotemporal track of them as well as context information attached (usually the tweet or comment text).

4) The system shows a tip stating the source and date of the location as well as the context (if any) associated with it when the user clicks on any pinpointed location on the map as demonstrated in Figure B.8.

5) The system enables the user to show the locations retrieved in form of heatmap instead of points.

6) The system enables the user to show the locations retrieved in form of heatmap instead of points.

7) The system provides the user with filtering mechanism for showing the extracted locations based on distance to a certain point on the map.

8) The system provides the user with filtering mechanism for showing the extracted locations based on a certain period of time.

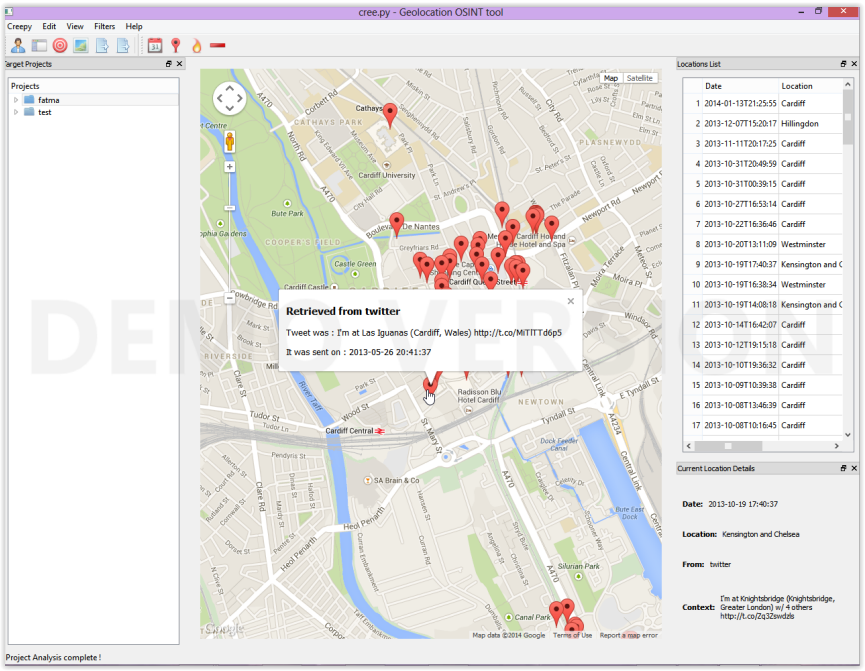


Figure B.8: A snapshot of 4sqmap when viewing the visited venues..

Limitations:

It seems that creepy works fine for extracting and plotting users shared locations on a map. However, it provides no further analyses or utilisation of users' location-related data. It can be

used to extract location tracks of users rather than to show how the shared location as well as information associated with it can be exploited to derive more personal information of users that in turn can compromise their privacy. Moreover, the three services enabled within this tool are LESNs, without incorporating any LBSNs which offer even richer location-based information of users and hence more potential privacy implications.

B.1.3 Discussion

By examining the available tools for information presentation tools of users' location-based data, GeoSNs tend to offer only limited presentation and accessibility options for showing users' shared data. These means revolve around projecting basic location historical tracks. Other external effort focuses on visualising users' shared data in order to facilitate the exploration of it.

Therefore, the privacy awareness tools provided online whether by the service providers or other third parties for the end users that enable them to be better informed about their location data acquisition and utilisation are sparse. They are also not aiming to show the potential privacy implication resulted from location disclosure on GeoSNs. In order to increase users' privacy awareness in a way that allows them to make informed decisions toward their location sharing , they need to be educated about the privacy dimensions involved with a location-sharing activity including who can access their data, what exactly of their data are collected, what can be inferred of their personal information based on it, and what it is utilised for.

B.2 Complete Results of the Task Analysis

Task analysis is simply analysing how a task can be achieved, and includes describing the physical and mental activities involved as well as details about the task's complexity and requirements [129]. Basically, it offers techniques to evaluate Human-Computer Interactions (HCI) in terms of information and processes involved in carrying out a task [130]. Hierarchical task analysis (HTA) is a widely used method for decomposing a task into goals and sub-goals required for accomplishing it in a hierarchical manner [131]. Cognitive Task Analysis (CTA) is another techniques that focuses on the cognitive factor of the task, including modelling human behaviour, performance and skills in relation to a task [132, 131]. Information gained from task analysis is used for diverse purposes such as training, systems and interface design as well as for optimising the human factors and minimising potential errors [133, 130].

The aim of this section is to employ systemic task analysis to investigate location-based tasks

in GeoSNs in order to address the gap in the knowledge of users' awareness and the related privacy risks. The analysis was carried out based on several aspects, namely:

1) Defining the task and establishing an abstract state transition diagram of it
2) Showing the complete steps of performing the task
3) Examining what information and interactions are required to achieve the task whether explicitly entered or collected previously
4) Examining what information presented to the user through performing the task
5) Investigating the potential privacy risks corresponding to the task in terms of what personal information can be derived
6) Studying the gap between users' awareness and the related privacy implication by examining the limitations of information presentation

B.2.1 Place discovery

B.2.1.1 Description

exploring places that is presented and recommended by the application. This task is carried out in the virtual presence mode and the state transition diagram for achieving this task is presented in Figure B.9.

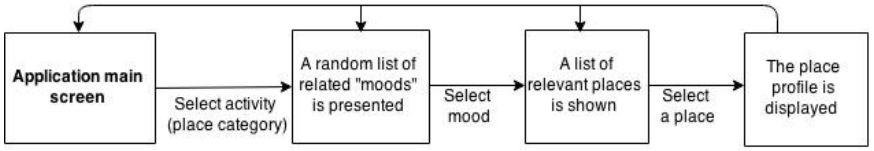


Figure B.9: State transition diagram of the place discovery task..

B.2.1.2 Task Steps

1) **Step one:**
When the user first opens the application, it mainly presents information about discovering nearby places as it the basic service to provide as presented in Figure B.10(a). The user can discover nearby places by selecting a category of place the user want to discover as taps on the top of the screen. The application by default chooses a category to

present probably based on the temporal information such as lunch category in the afternoon. When selecting a category, a list of relevant 'moods' is displayed that represents random options of interests and features for the selected place category.

Other information displayed:

- Names of random nearby places.
- Preview of other nearby place that the user liked or saved previously which is relevant to this category.

Information and interaction required: selecting a place category and a mood (essential).

2) **Step two:**
The results screen is displayed to the user when selecting a "mood" as presented in Figure B.10(b) showing a list of places that matches the selected category and mood along with details about the place including rating and the latest tip.

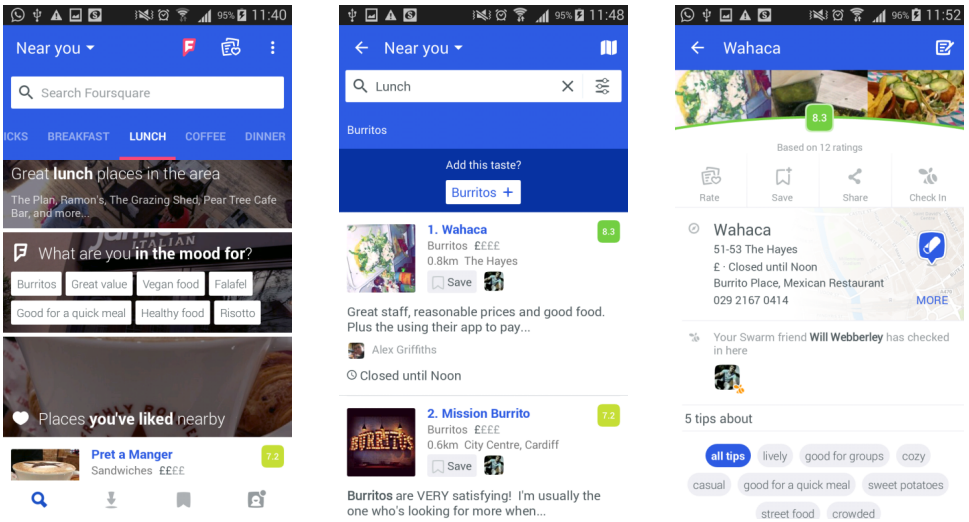
Other information displayed:

- Who of the users' friends have visited it the presented places
- If the user has saved the place before
- The chosen 'mood' is clearly displayed in middle of the screen where the users is asked to add the selected "mood" to their "taste" which is a feature that learns about users' interests and preferences in order to personalise the service for them as well be discussed later

Information and interaction required:

- Selecting a place to browse (essential)
- Adding the 'mood' to the user "tastes" (optional)
- edit the search (optional)

3) **Step three:**
When the user clicks on a place, the application displays back this place profile as illustrated in Figure B.10(c) including the place location on map and address, friends who visited the place, pictures of the place as well as the tips. The user can perform more task in a place that will be discussed later.



(a) (b) (c)

Figure B.10: Screenshots showing steps of place discovery task..

B.2.1.3 Privacy Implications and the Corresponding User Awareness

Users' information and interactions collected when discovering places on Foursquare can lead to the following privacy implications:

- Inferring what kind and place the user is interested in visiting and in what time by associating the selected place category to discover with the time of performing the task
- Extracting accurate users' interests and preferences for a particular place category in the given time that are provided by the user when choosing "moods" which can tell even more about the kind of activity they users would like to perform and their personality
- Revealing users' interests in visiting particular places when browsing their profiles

These consequences seem to be not clear to users as all they want is to find a place that matches their needs. The design of this task revolves around attracting the users to discover venues and encouraging them to state as well as register (as in adding tastes) their interests. Nevertheless, it lacks presenting relevant information that can contribute to raise their awareness of the possible risks including:

- Instantly notifying the users when selecting a category that it is added to their implicit profile in associated with the given time and emphasis on it if it matches previous behaviour of the user

- Instantly notifying the users when selecting a mood that it is added to their implicit profile in associated with the selected category and the given time and emphasis on it if it matches previous behaviour of the user

- Enabling the user accessing their activity log where they can see the details of their recorded interactions in terms of viewing place categories the user has been interested in clustered in temporal manner while showing the users' moods for each category.

Showing nearby place the user liked or saved displays a bit of the user recorded history in a beneficial aspect without showing what such information can be exploited for (which will be discussed later).

B.2.2 Searching for a place

B.2.2.1 Description

Searching for places based on keywords and other filtering options. This task is carried out in the virtual presence mode and the state transition diagram for achieving this task is presented in Figure B.11.

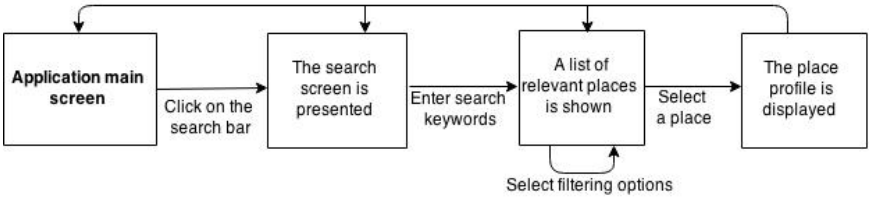


Figure B.11: State transition diagram of the place search task..

B.2.2.2 Task Steps

1) **Step one:**
When the user click on the search bar showing on the top of main screen presented in Figure1, the screen in Figure B.12(a) is displayed allowing the user to type in as well as showing the user recent search history as well as suggestions for place categories and features.

Information and interaction required: entering search keywords or select from suggestions (essential)

2) **Step two:**
Once the user typed the keyword and clicked search, a similar result screen of the place discovery task is shown as demonstrated in Figure B.12(b). In addition, the user can filter the results by selecting the features and category of the place as illustrated in Figure 8 which in turn will update the results accordingly.

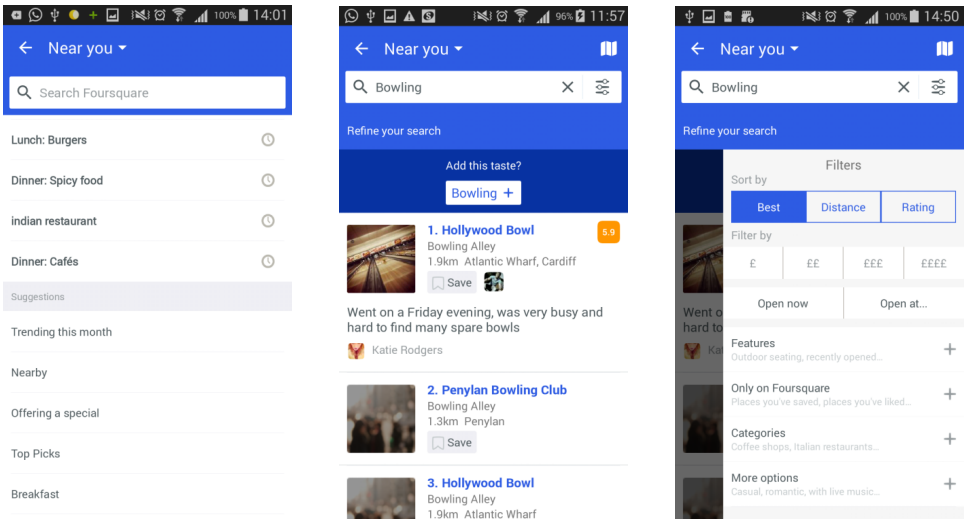
Other information displayed:

- Who of the users' friends have visited it the presented places
- If the user has saved the place before
- The entered keywords can be detected as a 'mood' which displayed in middle of the screen where the users is asked to it to their 'taste'

Information and interaction required:

- Selecting filtering options (optional)
- Adding a 'mood' to the user's 'tastes'
- Selecting a place to browse (essential)

3) **Step three:**
When the user clicks on a place to view, the place details will be presented in a similar manner of Figure B.12(c).



(a) (b) (c)

Figure B.12: Screenshots showing steps of place search task..

B.2.2.3 Privacy Implications and the Corresponding User Awareness

Users' information and interactions collected when searching for places on Foursquare can lead to the following privacy implications:

- Inferring the users' interests in particular times by exploiting the search keywords along with the time
- Extracting the users' preferences when refining the search with the given filtering options

Users focus in finding a venue in the current time without keep track of their past searches and they have no way of see their registered search profile. Hence, they are not aware of it or how it can be utilised for. Therefore, the design of this task is limiting the users' awareness of the privacy implications since it is not showing to them:

- The stored search keywords and their association with time
- The repeated patterns of search that can be used for information inference

B.2.3 Personalise the service

B.2.3.1 Description

The users can explicitly personalise the service according to their interests and preference by adding "tastes" either directly through the personalise task or during performing place discovery and search tasks as demonstrated previously. This task is performed during the virtual presence mode of use and its state transition diagram is presented in Figure B.13.

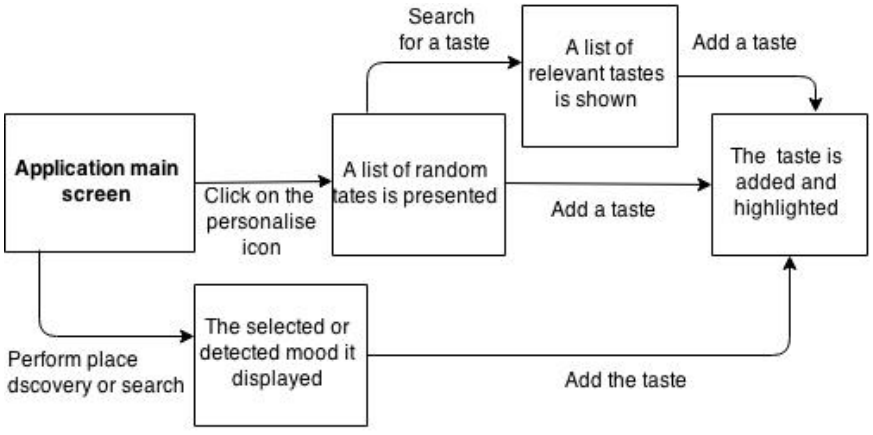


Figure B.13: State transition diagram of the personalisation task..

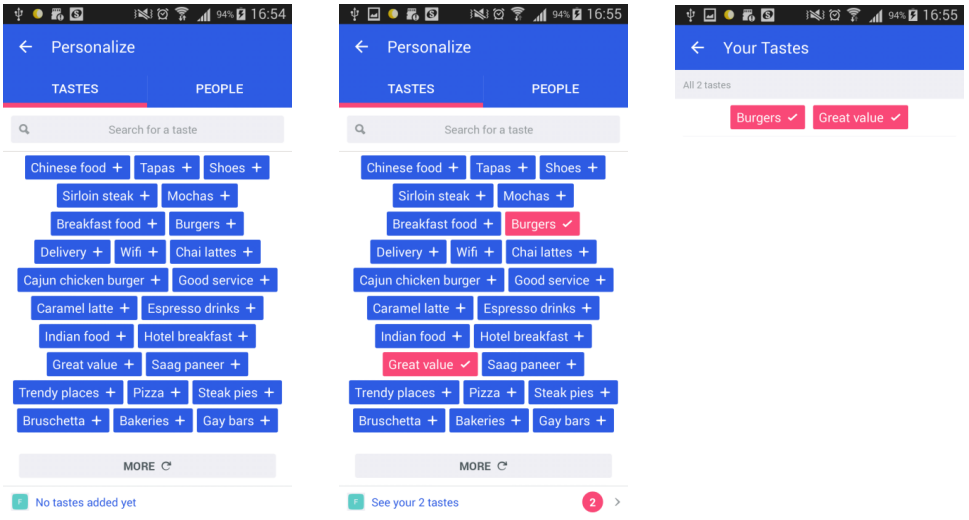
B.2.3.2 Task Steps

1) **Step one:**
When clicking on Foursquare 'F' icon on the top right corner as shown in Figure 2, the user is presented with a random list of various kinds of interest, preferences and features relevant to places generally which is illustrated in Figure B.14(a). The user can search for a taste, load more of them or selected from what presented. The user can also add a taste while performing the place discovery or search as in Figure B.10(b) and B.12(b).

Information and interaction required: selecting a taste to add (essential)

2) **Step two:** The taste is highlighted in red colour when added as shown in Figure B.14(b).

3) **Step three:**
The user can view the selected taste by clicking on the bottom link that navigates to display Figure B.14(c).



(a) (b) (c)

Figure B.14: Screenshots showing steps of place personalisation task..

B.2.3.3 Privacy Implications and the Corresponding User Awareness

Adding 'tastes' by the user can pose the following privacy implications:

- Extracting accurate information that reflects the user real interests and preferences not estimated ones

- Exploiting users' tastes with other spatiotemporal (check-in) information to derive what the user do or like in a place or to predict future movements.

The users' main aim of adding the "tastes" is to facilitate finding places matching their preferences. All of what they can view is a list of their tastes. However, the application is not showing to users the other privacy dimensions of giving away this kind of information. Hence, the application interface lack presenting:

- How these tastes are interconnected to form a user's profile of interests
- The behavioural profile in terms of showing how the user' mobility (i.e. check-ins) relates to their stated tastes

B.2.4 Saving a place

B.2.4.1 Description

the users can record any place and add them to their list as a reminder for them of the places they are interested in visiting in the future in an organised manner. The state transition diagram for achieving this task is presented in Figure B.15.

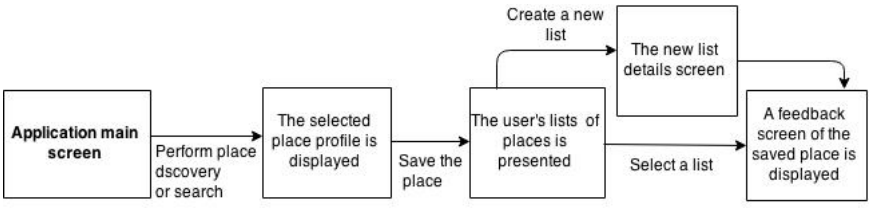


Figure B.15: State transition diagram of the saving a place task..

B.2.4.2 Task Steps

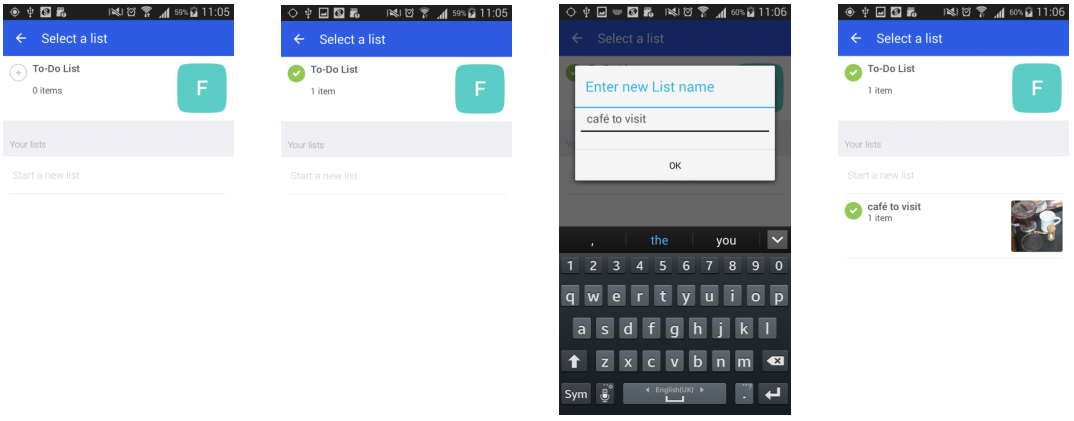
1) **Step one:**
Saving a place starts from step two and three of performing the place discovery or search tasks. To save a place, the user clicks on the save button shown in the result list of places or in the place details screen illustrated in Figure B.10(b) and B.10(c).

Information and interaction required: selecting a place to save (essential)

2) **Step two:**
The users are navigated to the screen shown in Figure B.16(a) where they can save the place in the default "To-Do-List" or create new list to add the place to as demonstrated in Figure B.16(b) and B.16(c).

*Information and interaction required:*selecting a list to save the place into (essential)

3) **Step three:** The list is marked green as demonstrated in Figure B.16(d) to indicate it was successfully saved.



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

Figure B.16: Screenshots showing steps of saving a place task..

B.2.4.3 Privacy Implications and the Corresponding User Awareness

Users' information and interactions collected when saving places on Foursquare can lead to the following privacy implications:

- Inferring interests and preferences
- Predicting the future movement especially the user tend to visit the saved places when examining the check-in information

The use seems to not be conscious such utilisations as they want to save their time for future place discovery. There are gaps in the information presentation can contribute in increasing their privacy awareness including:

- Notifying the user about the interests and preferences extracted and added when saving a place especially if they match the user' ones derived from performing other tasks such as personalising the service

- Notifying the user that this place registered to be visited in the future in an estimated certainty derived from examining the user's visiting behaviour of the saved places

B.2.5 Sharing a place

B.2.5.1 Description

the user can share a place with their social connections or on other social networks. The state transition diagram for accomplishing this task is presented in Figure B.17.

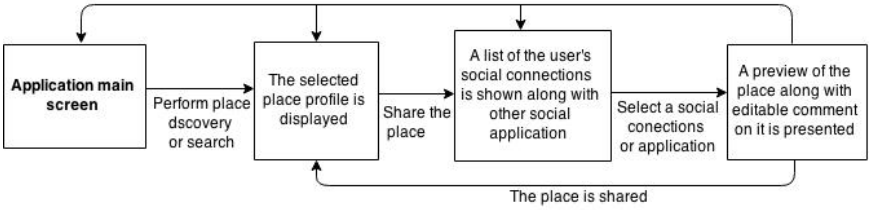


Figure B.17: State transition diagram of the place sharing task..

B.2.5.2 Task Steps

1) **Step one:**
Sharing a place starts when viewing a place details page such as Figure 4 by clicking on the share button in the top.

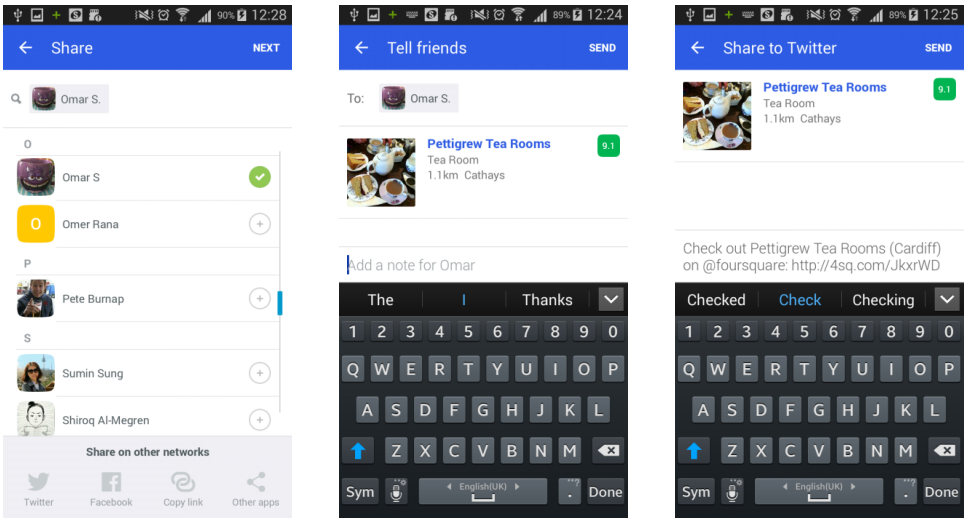
Information and interaction required: selecting a place to share (essential)

2) **Step two:**
The screen in Figure B.18(a) is displayed where the user can select from the social connection to share with or share the place widely on other social networks.

Information and interaction required: selecting social contacts or an application (essential)

3) **Step three:**
When sharing with selected social contacts, the screen in Figure B.18(b) is shown including a preview of the recipients and the place as well as the ability to write a comment, while choosing to share with other social navigates to screening Figure B.18(c) displaying a preview of the place as well as pre-defined text to share including the place details and like that can be edited. Once shared, the user is redirected to the place profile screen showing a temporary bubble tip saying the 'just shared!'.

Information and interaction required: writing a comment about the place (optional)



(a) (b) (c)

Figure B.18: Screenshots showing steps of place sharing task..

B.2.5.3 Privacy Implications and the Corresponding User Awareness

Users' information and interactions collected when sharing places on Foursquare can lead to the following privacy implications:

- Sharing a place with selected friends can imply that these friends are interested in this kind of place and hence can used to infer their interests and preferences as well as the user's.
- Sharing a place with other social networks facilitates and widens the accessibility of the user personal information and hence increase the potentials of jeopardising users' privacy
- Mining the comment shared with the place can reveal user emotional and physical experience associated with the place such as activities and interests as well as information about the friends.

Users share places to recommend them to others and to disclose information related to these places. Nevertheless, it is not shown to them how sharing places can impact their privacy by:

- Displaying what personal information of the users or their friends can be derived
- Presenting who can access this information

B.2.6 Rating a place

B.2.6.1 Description

The user can rate a place based on different aspects. The state transition diagram for achieving this task is presented in Figure B.19.

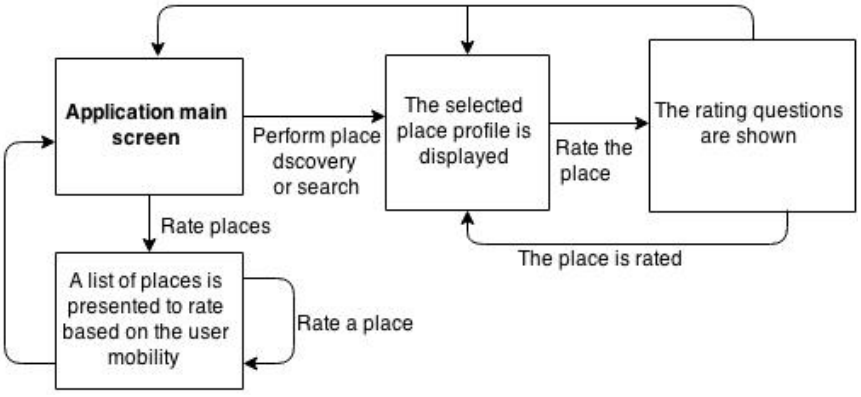


Figure B.19: State transition diagram of the place rating task..

B.2.6.2 Task Steps

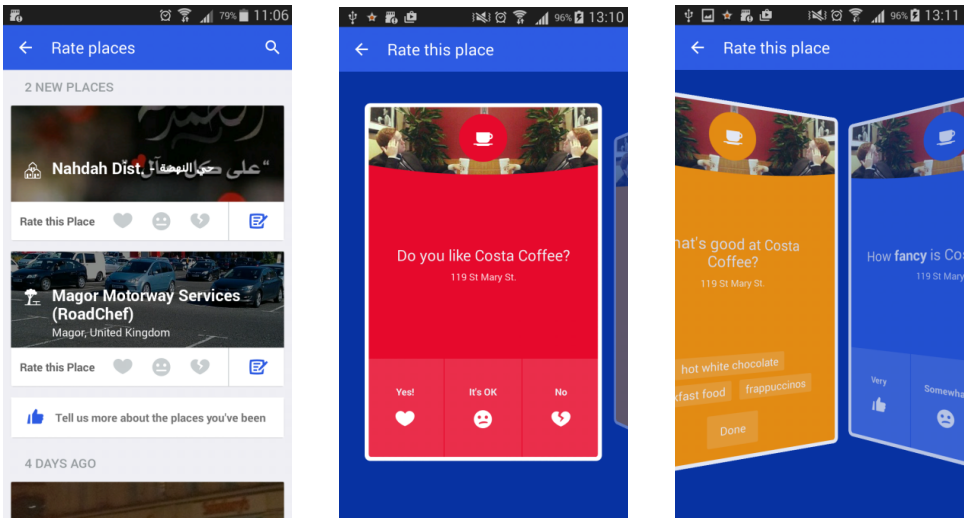
1) **Step one:**
Rating a place starts when viewing a place details page such as Figure B.20(a) by clicking on the rate button in the top. The user can also rate a number of places suggested by the application as demonstrated in Figure 23 via clicking on the rate places 'list' icon on the top right of main application screen shown in Figure B.10(a).

Other information displayed: selecting a place to rate (essential)

Information and interaction required: previews of the places are presented when rating places suggested by the application

2) **Step two:**
The user rates the place by choosing one of the three options as shown in Figure B.20(b). Once one option has been chosen, the user is directly presented with a series of slides asking about this user feedback regarding the place features as demonstrated in Figure B.20(c). The rating icon on the place profile then changes to represent the user' selected rate.

Information and interaction required: select a rate for the place (essential)



(a) (b) (c)

Figure B.20: Screenshots showing steps of place discovery task..

B.2.6.3 Privacy Implications and the Corresponding User Awareness

When users rate places, they explicitly state their opinions. Hence, users' information and interactions collected when rating places on Foursquare in the virtual mode can lead to the following privacy implications:

- Extracting the users' accurate interests, preferences and favourite places
- Inferring the returning probability to a place in the future
- Utilising the rated place for place recommendation for the friends

Furthermore, the functionality of rating a group of places recommended by the application actually shows recent places that the user has been to without checking into them or nearby places of the user's previous locations which suggest that the application is constantly recording the user location even without using the service and hence the user' location mobility is implicitly tracked .

The users rate a place at the given time without being able to see how this information can be linked to their previously shared data and how it can be exploited. The application is not presenting to them relevant information that influence their privacy awareness including:

- A list of the extracted interests and preferences of the user
- The relationship between these interests and particular places implicitly or explicitly visited by the user

- How these interests and preferences are linked to the general network of the user interest profile

B.2.7 Checking into a place

Description

Users can register their presence in a place by checking into it. The check-in task along with the other sub-tasks are illustrated in Figure B.21.

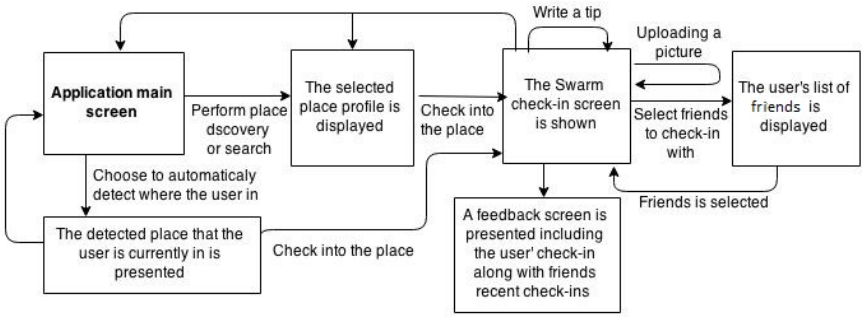


Figure B.21: State transition diagram of the task of check-in to a place..

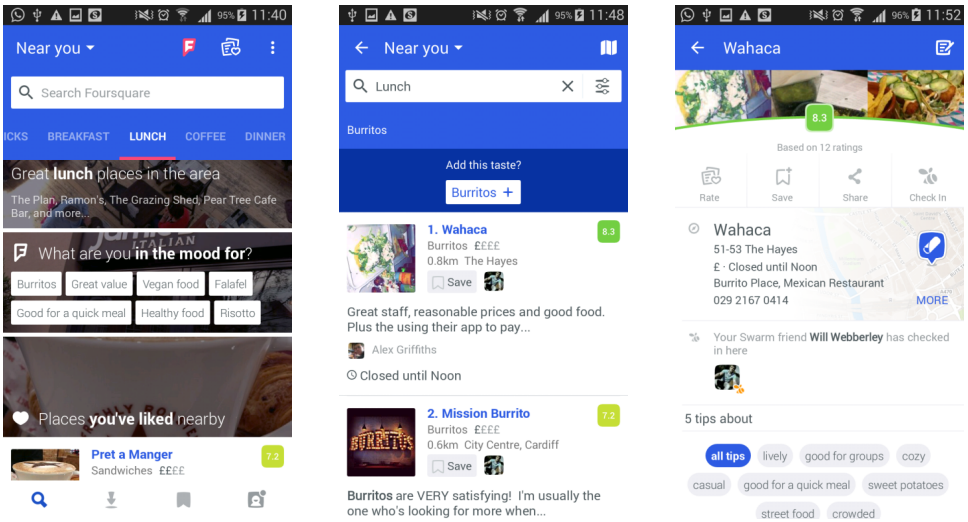
Task Steps

Figures B.22 and B.23 present screen-shots of the step needed for completing this task.

1) Searching for a place from the main application window (Figure B.22(a) and B.22(b))
2) Selecting and viewing a place (Figure B.22(c))
3) clicking on the check-in icon at the top of the place profile page, where the user is directed to Swarm (Figure B.23(a))
4) Optionally commenting on the place, adding a picture, tag friends as displayed in friends, or sharing the check-in on other social networks (Figure B.23(b) and B.23(c))
5) Checked-in to the place (Figure B.23(d)

Information required by the user: a place name or type.

Information displayed by the application: the user recent check-in and where the friends are in correspondence to the users' location, clustered based on distance. In addition, a 'leader-board' can be shown demonstrating the users' rank among their friends in terms of the number of check-ins into the place type, as illustrated in Figure B.23(e).



(a) (b) (c)

Figure B.22: Screenshots showing the steps of searching for a place..

Information that can be inferred:

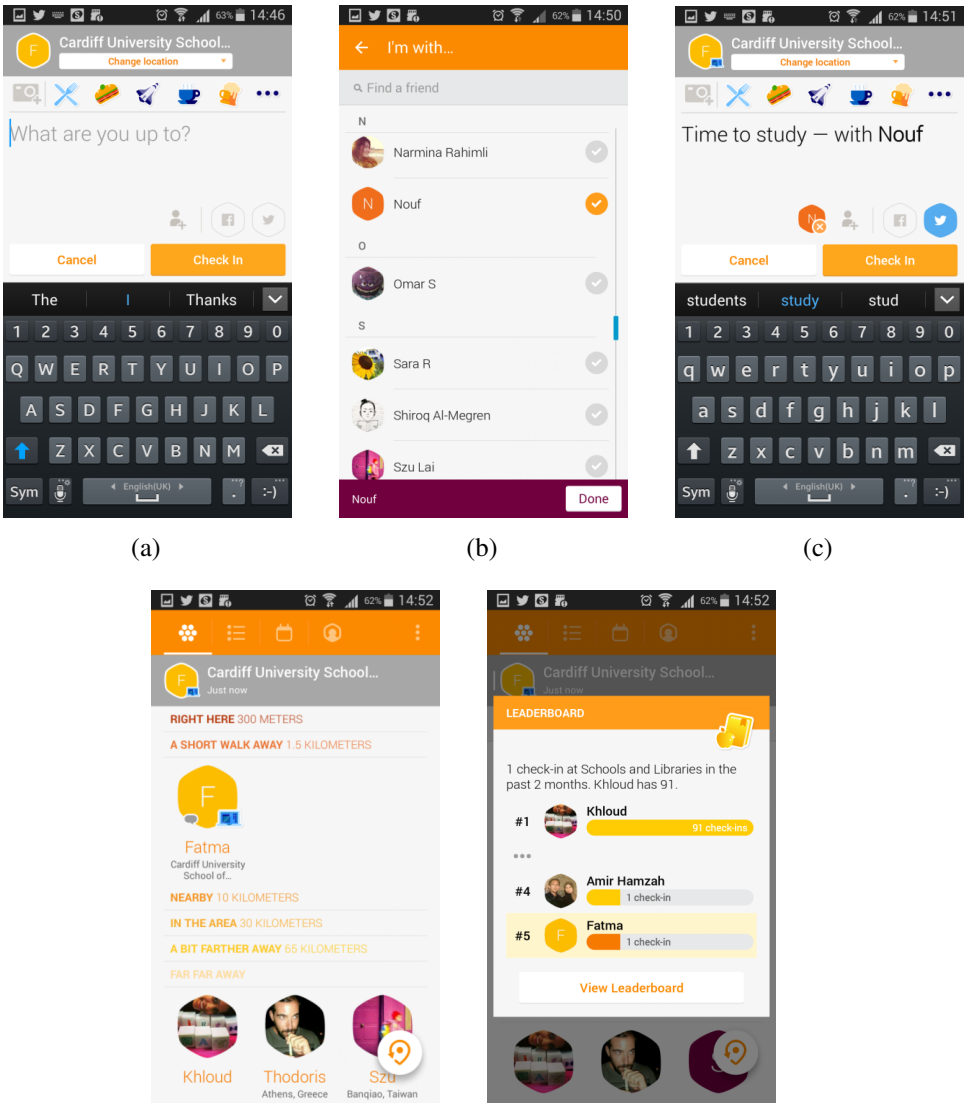
Check-ins provides very precise spatial and temporal data along with other user generated data which can be used to construct a comprehensive location-based profile of users, and thus lead to the following inferences:

- Revealing users' current presence or absence, sensitive places, and top places visited.
- Infer interests, preferences and activities performed.
- Extracting spatiotemporal mobility patterns, which enables:
  - Inferring probability of returning to a place.
  - Predicting the user's future movement (especially if the user has liked or saved the place previously), and transitions from one place to another.
- Extracting co-location with friends their common interests and activities.
- Allowing wider accessibility to the user data when sharing on other social networks, and hence more possible threats.

B.2.8 Writing a tip

Description

Users can leave comments to express their views and experience on places. as a way of expressing what they think of them and what brings them to them. They can write a tip about a place at



(a) (b) (c)

(d) (e)

Figure B.23: Screenshots showing the steps of the check-in task.

any time even if they are not present at this place. The state transition diagram for writing a tip is presented in Figure B.24.

Task Steps

1) Searching for a place from the main application window (Figure B.22(a) and B.22(b))
2) Selecting and viewing a place (Figure B.22(c))
3) Clicking on the writing icon in the top right corner.
4) Writing a tip (Figure B.25(a))

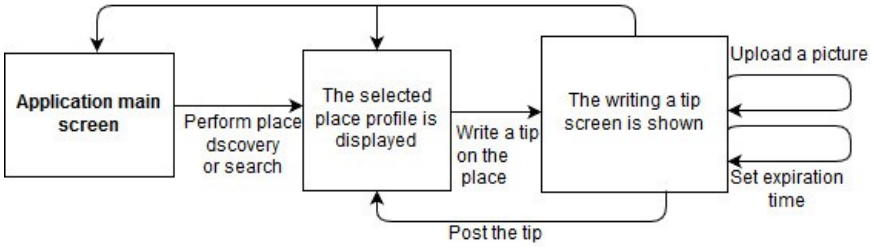
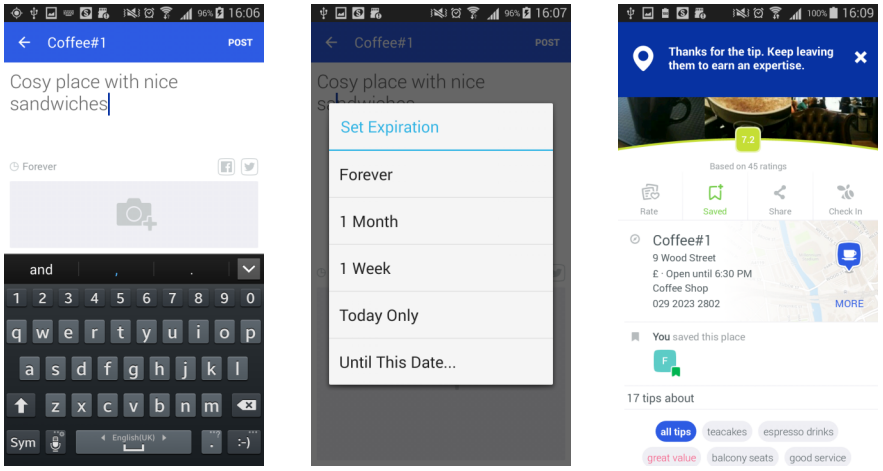


Figure B.24: State transition diagram of the task of writing a tip.

5) Optionally uploading a picture, sharing the tip on other social networks, and set an expiration time for it (Figure B.25(b)).

6) Posting it (Figure B.25(c)).



(a) (b) (c)

Figure B.25: Screenshots showing the steps of the writing a tip task..

Information required by the user: a place name.

Information displayed by the application: the place details and address, as well as other people' comments on it.

Information that can be inferred:
A tip can be a set of words or complete sentences. Text analysed of the tip can be used to extract the following:

- Deriving users' interests in and preferences about a place, as well as the activities they performed.
- Implying that a user has visited a place previously, or is even in a place at a given time. A tip is associated with a time-stamped, which can reveal the user' mobility.

B.2.9 Making a plan

B.2.9.1 Description

The user of Swarm can make plans for visiting places and share them with friends in Swarm. The state transition of this task is displayed in Figure B.26.

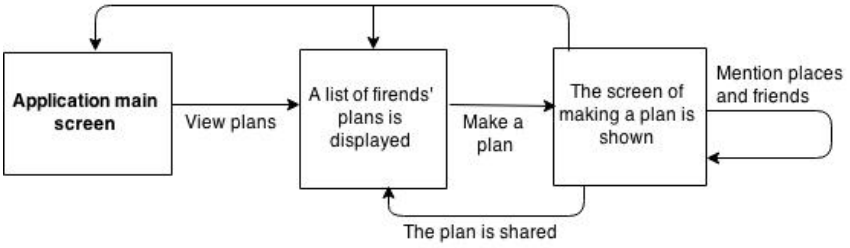


Figure B.26: State transition diagram of making a plan task..

B.2.9.2 Task Steps

1) **Step one:**
To create a plan, the user clicks on the calendar icon in the bottom left of the plans screen shown in Figure B.27(a). The user then can write his or her plans regarding visiting a place where the user can mention friends and places and then send it as demonstrated in Figure B.27(b).

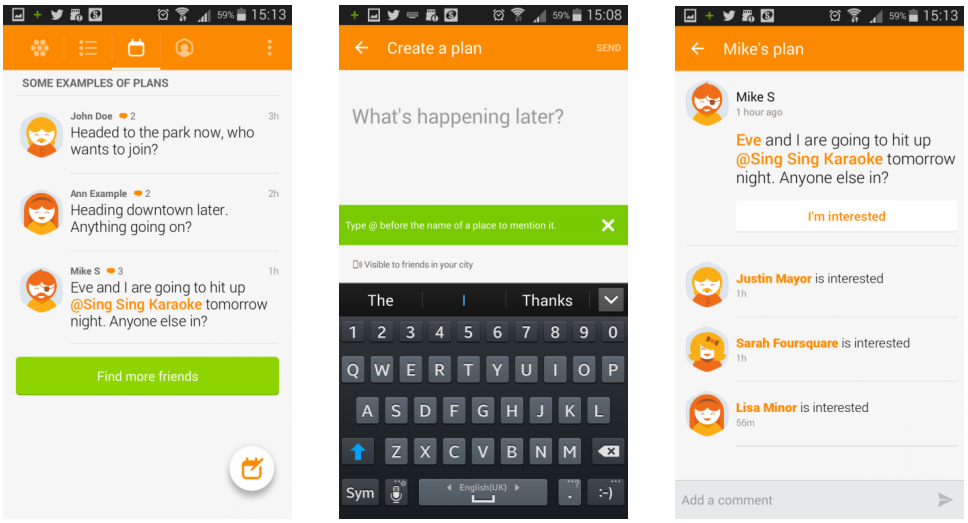
Other information displayed: the plans of friends are displayed including where and when to go and with whom.

Information and interaction required:

- Write a plan (essential)
- Tag places (optional)
- Tag friends (optional)

2) **Step two:**
The use plan is shared with the friends in same city of the user. The friends can interact with this plan be stating their participation by clicking on “I’m interested” and leaving a comment on the plan which is presented in Figure B.27(c).

Information and interaction required: sharing the plan (essential)



(a) (b) (c)

Figure B.27: Screen-shots showing steps of making a plan task..

B.2.9.3 Privacy Implications and the Corresponding User Awareness

The plans in fact add another dimension to the location privacy risks as the user is not only providing current location information but also future one which can result in knowing the user explicitly and accurately provides the future mobility in terms of:

- where the user will visit next
- when
- with whom

The user mainly share plans as an easy way to arrange a visit for a place with friends or to find who can accompany the user to where this user is going. Nevertheless, important information related the user's privacy when knowing the place and time of the future visit is not shown including:

- Tracking future movements of the user
- Knowing what the user is interest in visiting or doing in the future
- Deriving the absence and presence in particular places which can lead physically harass the user or take advantage of not being in particular places of the user
- Involving the companions in the same privacy threats

B.2.10 Background location tracking

This task is carried out by the application rather than the users, where their locations are implicitly and constantly collected without the need for users to interact with the application, hence they are unaware of it. Foursquare and Swarm, once installed, continuously track and store users' device locations, even when the application is not in use. Moreover, Swarm can share the district level of the users' current location with their friends without even checking-in or using the application. This task is included in the privacy settings of the application and is enabled by default without notifying the user about it. The hidden tracking of users' location enables the collection of complete and accurate location trajectories that reflect a user's real-life mobility. Thus, this can highly jeopardise the users' privacy, since it allows the disclosure of sensitive places that users do not want to reveal by checking-in to, as well as extracting the duration of stay and deriving the place semantics for the users' entire location track. Hence, this leads to the inference of more precise personal information and mobility patterns that can include all of the possible location privacy threats.

In order for users to be aware of this hidden task and its privacy consequences, they should have full access to their location profile, including the collected whether while using the application or in the background tracking, as the user's mobility patterns and personal information can be accurately inferred. Users should also be explicitly consented each time their location is collected by the application.

Appendix C

Individual Analysis of the Check-in Scenarios, and Survey Materials and Question of The Study on Towards Real-Time Informed Consent for Location Privacy in GeoSNs

In this appendix, the analysis results for each of the six check-in scenarios are introduced. In addition, the six developed check-in scenarios for the study are presented first along with their screen-shot of the Swarm applications and the privacy feedback tool. Then, the survey used are provided including the instructions and questions.

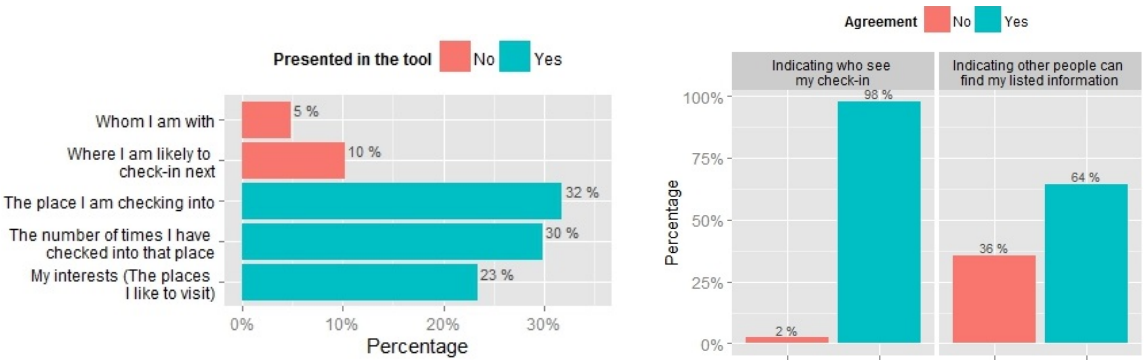
C.1 Individual Analysis of the Check-in Scenarios

In this section, results analysis for each of the six check-in scenarios (with feedback only and with feedback and control) are presented.

C.1.0.1 First Yellow Check-in Scenario

This scenario has a yellow (moderate) privacy level. Most of the participants were able to comprehend the privacy information presented in the tool. In particular, they understood on average three-quarters of the presented information (73%). The inferred information the participates were able to capture easily are "The place I am checking into" (32% of the information options), followed by "The number of times I have checked into that place" (30%) as shown in Figure C.1(a). Moreover, the majority of them reported that they were informed about who

can see their check-in (98%) and were aware that other people can find out all of the listed information about them (64%) as illustrated in Figure C.1(b).



(a) (b)

Figure C.1: (a) What revealed information by the tool the participants were able to capture, (b) The participants' awareness of their check-in data accessibility achieved by the tool..

In terms of the sufficiency of the content provided, 68% of the participants stated that the tool sufficiently indicates the privacy risks associated with this check-in, whereas only 11% disagreed with that as demonstrated in Figure C.2. 54% of them reported that the information provided in the tool makes them concerned about their privacy in this check-in, while only 26% were not. Most of them actually (74%) showed their agreement with the privacy level proposed by the tool (Moderate-Yellow). Only 10% disagreed where some of them thought it should be riskier by saying "This seems like a fairly high degree of access to information – not moderate.", while others thought that "there is no risk.".

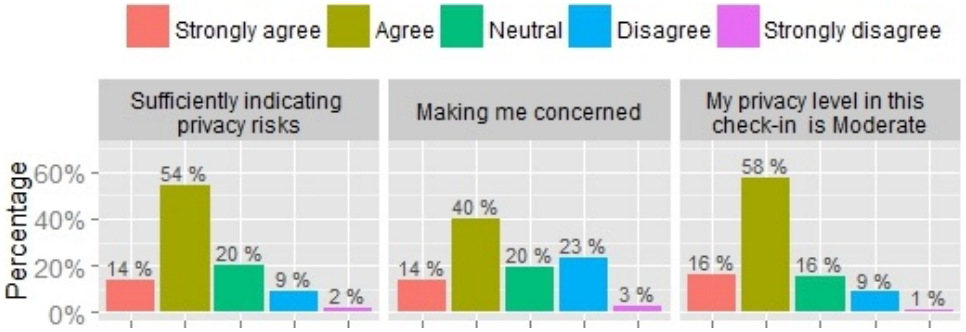
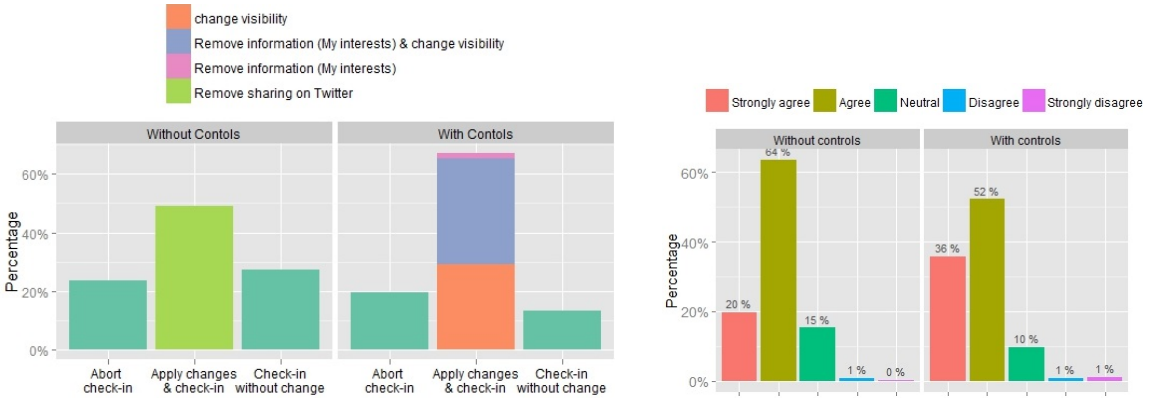


Figure C.2: (The participants' reaction towards the tool effectiveness..

Examining how the privacy feedback tool can impact participants' check-in behaviour in case of feedback only showed that most of them (49%) wanted to change their check-in information and proceed by removing sharing on Twitter option, and 27% wanted to check-in without any changes, whereas 24% of them wanted to abort the check-in completely as presented in Figure C.3(a). However, when the privacy controls are provided by the tool, the participants' tendency

towards changing their check-in information in terms of removing their extracted information and changing the check-in visibility, and proceed raised by 18% (67% of them), where 54% of them wanted to delete their interests and change the check-in visibility, 43% wanted only to change the check-in visibility, and the rest to just remove their interests. Both attitudes of aborting check-in and checking-in without any changes also decreased representing 20% and 13% of them respectively. 77% of those who wanted to change the visibility chose the cheek-in to be viewed by friends only, while the remaining 23% set it as private. In regards to the tool role on making informed decisions, 84% of the participants agreed that the tool helped them with making this decision in case of feedback only, while this agreement even increased in the cased of having the controls representing 88% of them as shown in Figure C.3(b). That donates that offering the users privacy controls over the utilisation and accessibility of their data actually allows them to better manage their privacy and make use of the application in the same time.



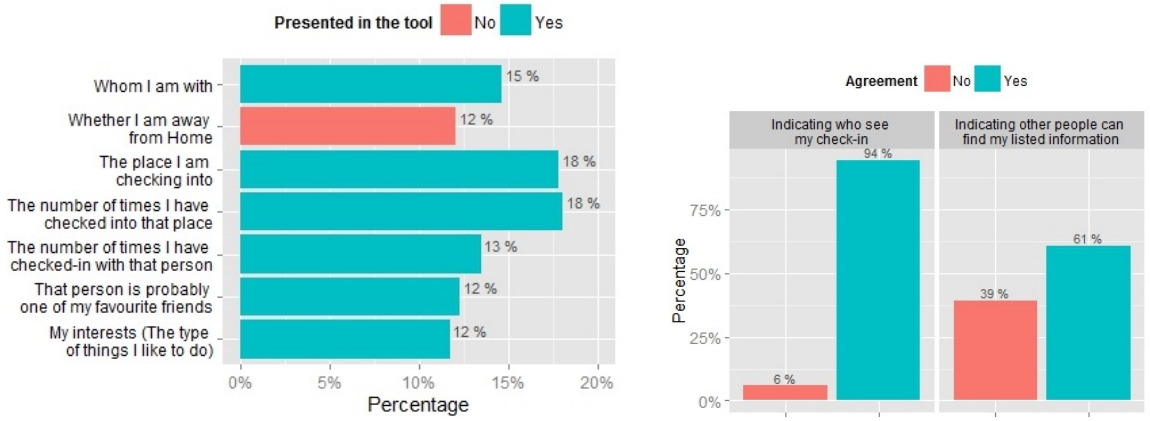
(a) (b)

Figure C.3: (a) The participants' check-in decision with and without the privacy controls, (b) The participants' views about the tool role on making informed decisions..

C.1.0.2 Second Yellow Check-in Scenario

This has also a moderate privacy level. The participants were able to realise on average more than half of the extracted personal information presented in the tool (62%). They mostly captured the information related to "The place I am checking into" and "The number of times I have checked into that place" (18% each), followed by "Who I am with" (15%) as illustrated in Figure C.4(a). In addition, the majority of them reported that the tool clearly indicates information about their information acceptability, particularity who can see their check-in (94%) and that other people can find out all of the listed information about them (61%) as displayed in Figure C.4(b).

As for the tool effectiveness, 68% of the participants stated that the tool sufficiently indicates the privacy risks related to this check-in, while only 12% disagreed with that as shown in Figure C.5.



(a) (b)

Figure C.4: (a) What revealed infromation by the tool the participants were able to capture, (b) The participants' awareness of their check-in data accessibility achieved by the tool..

56% of them reported that the information provided in the tool makes them concerned about their privacy, whereas the concern level of 24% of them was not impacted. Moreover, most of the participants (67%) agreed with the privacy level indicated by the tool which (Moderate). Only 15% had other views on the privacy level proposed. Again, some of them think "it is more than just moderate, the application is profiling me and allowing any random person to know these things about me. That's extremely scary.", while others think "Risk is low as long as you control your friends to people you know"', which suggest that they are not fully understand the associated privacy risks where an adversary can be inside the service provider or its third parties who access their data.

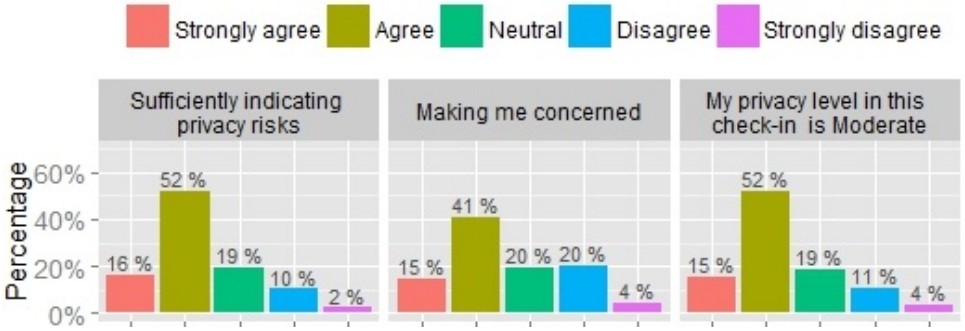
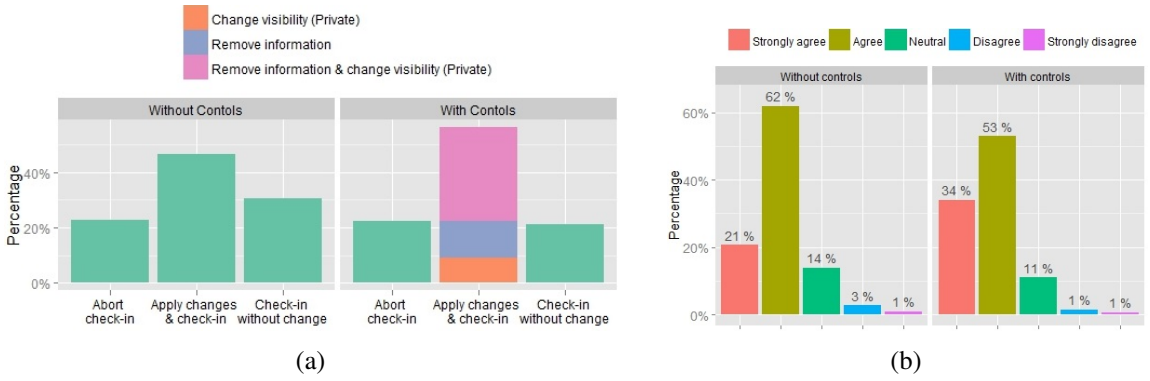


Figure C.5: (The participants' reaction towards the tool effectiveness..

In terms of the participants' behaviour towards this check-in's scenario, most of them (46%) wanted to change their check-in information before proceeding, and 31% wanted to check-in without any changes, whereas 23% of them wanted to abort the check-in completely as in the case of feedback only as presented in Figure C.6(a). For those who wanted to change their check-in information, removing the tip has the highest are the most desired modification

(65%), followed by removing the friend tag (35%). Nevertheless, the participants' willingness towards changing their check-in information (deleting extracted information and changing the visibility) and proceed with the check-in raised by 11% when the privacy controls are provided by the tool (57% of them) where 60% of them wanted to remove their derived interests and change the visibility to private, 25% wanted just to remove the inferred information and the remaining to just set the visibility to private. Checking-in without changes is dropped by 10% (22% of them), while there is no much difference in their attitude towards aborting the check-in. Participants who chose to change their check-in information wanted mostly to remove the favourite friend, followed by removing the interests representing receptively 55% and 45% of the selected changes. Most of the participants found the tool to be helpful in making informed decisions (83%) in the case of feedback only, whereas this influence even increased when the controls are provided (87%) as shown in Figure C.6(b).



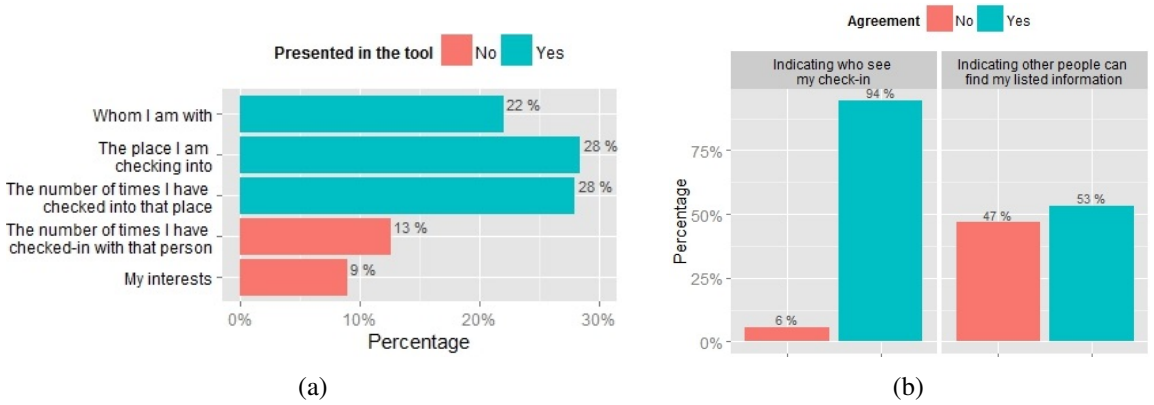
(a) (b)

Figure C.6: (a) The participants' check-in decision with and without the privacy controls, (b) The participants' views about the tool role on making informed decisions..

C.1.0.3 Green Check-in Scenario

The is the only scenario tagged as green (safe) privacy level. Similarly, the participants were able to pin point most of the revealed personal information presented in the privacy feedback (80%), particularly "The place I am checking into"and "The number of times I have checked into that place" (28% each of the information options) as presented in Figure C.7(a). Furthermore, the majority of them agreed that the tool shows information about their data accessibility in terms of check-in visibility (94%) and personal information inferences (53%) as demonstrated in Figure C.7(b).

In regards to how the privacy tool influence participants' attitude , 64% of them were in favour of the sufficiency of the provided content in terms of related privacy implications as presented in Figure **??**. Surprisingly, 45% of them stated that the information provided in the tool impact their privacy concerns giving that it is a safe scenario. 79% showed their agreement with the



(a) (b)

Figure C.7: (a) What revealed infromation by the tool the participants were able to capture, (b) The participants' awareness of their check-in data accessibility achieved by the tool..

privacy level proposed by the tool (Green-Safe). The 5% who disagreed explained by saying things like "because 3rd party's still have access to my information". This group of participants can be categorised as privacy fundamentalists based on Westin/ Harris privacy segmentation index since they show higher levels of concern towards their privacy [86].

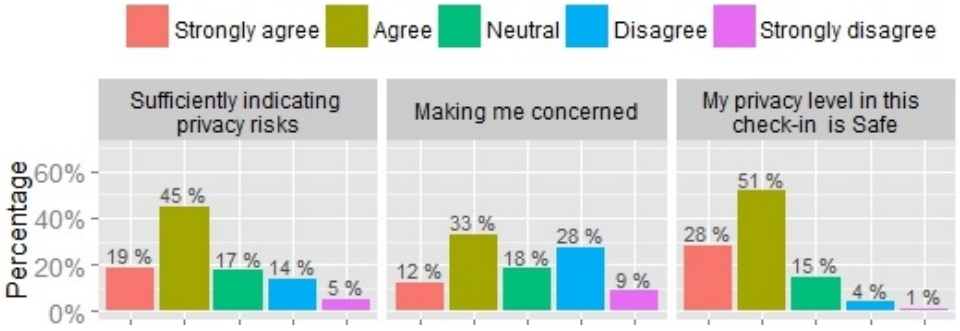
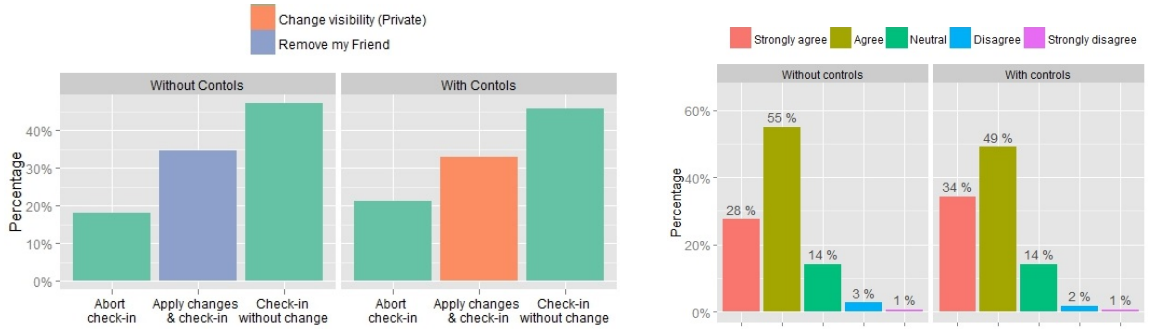


Figure C.8: (The participants' reaction towards the tool effectiveness..

Participants' check-in behaviour in this scenario is also impacted by the privacy tool. Most of them (46%) wanted to check-in without any changes as expected since it relatively not causing a privacy risk, 34.5% wanted to change their check-in information and proceed (by removing their friend tag in the case of feedback only and setting the visibility to private when the privacy controls is provided, whereas only 19.5% of them wanted to abort the check-in completely as presented in Figure C.9(a). In this scenario, there is no noticeable variation between participants check-in reported behaviour with or without having privacy controls which might be due to the fact there is not much to control. Privacy feedback showed to influence decision-making process for location disclosure where 83% of the participants found it easier to decide upon their check-in action whether the privacy controls are offered or not as shown in Figure 39. Specifically, 34% strongly agreed about his matter in the case of having the privacy control option as demonstrated in Figure C.9(b).

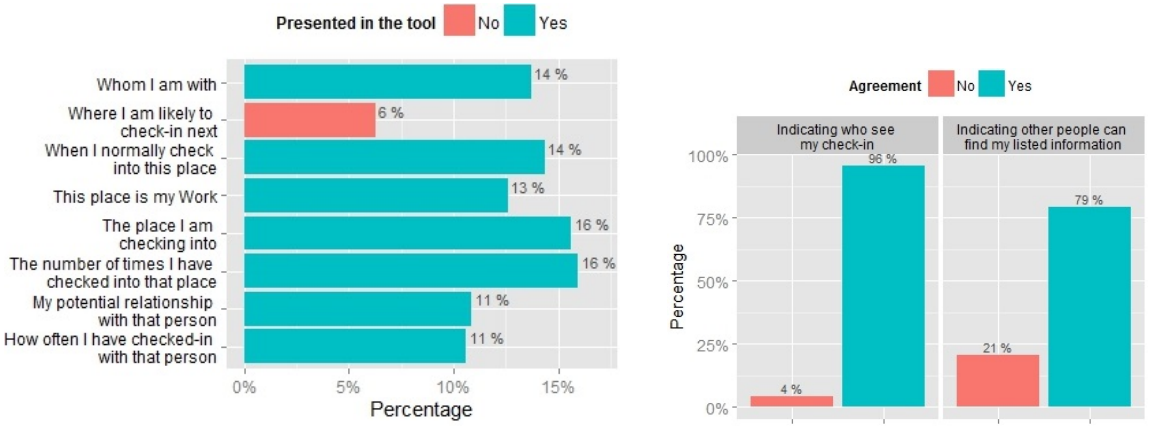


(a) (b)

Figure C.9: (a) The participants' check-in decision with and without the privacy controls, (b) The participants' views about the tool role on making informed decisions..

C.1.0.4 First Red Check-in Scenario

This scenario has a red (risky) privacy level. Most of the participants were able to comprehend the privacy information presented in the tool (62.2%). In particular, they mostly were able to recall "The place I am checking into" and "The number of times I have checked into that place" (16% each), followed by "When I normally check into this place" and "Who I am with" (14% each) as illustrated in Figure C.10(a). Similarly to previous scenarios, the vast majority of them agreed that the tool indicates who can see their check-in (96%) and that other people can find out all of the listed information about them (79%) as displayed in Figure C.10(b).



(a) (b)

Figure C.10: (a) What revealed infromation by the tool the participants were able to capture, (b) The participants' awareness of their check-in data accessibility achieved by the tool..

Moreover, 80% of the participants agreed that the tool sufficiently indicates the privacy risks related to this check-in, whereas only 6% disagreed as shown in Figure C.11. Most of them (74%) stated that the information provided in the tool makes them concerned about their privacy which is so far the highest concern level among the previous scenarios since this scenario lead

to more serious privacy threats. In addition, most of the participants' opinions (79%) about the privacy level of this check-in were consistent with the ones indicated by the tool (Risky-Red). Only 10% were not where their explanations revealed that they still have a limited perception about the relationship between privacy implications and location disclosure since the majority of them said there is no risk because "only my friends will see my location" and "because of the privacy setting I am provided with.".

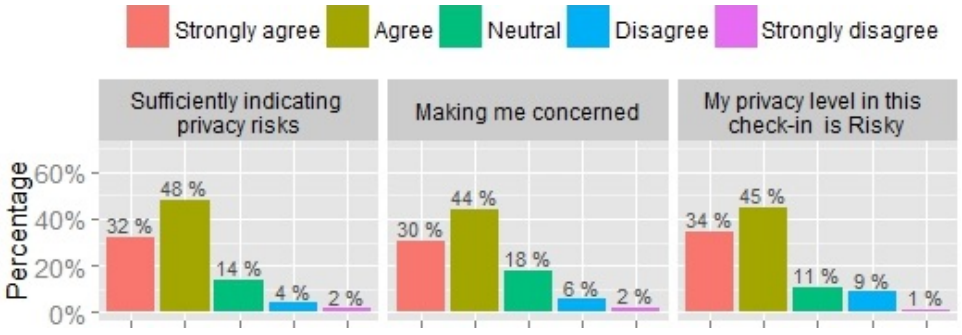
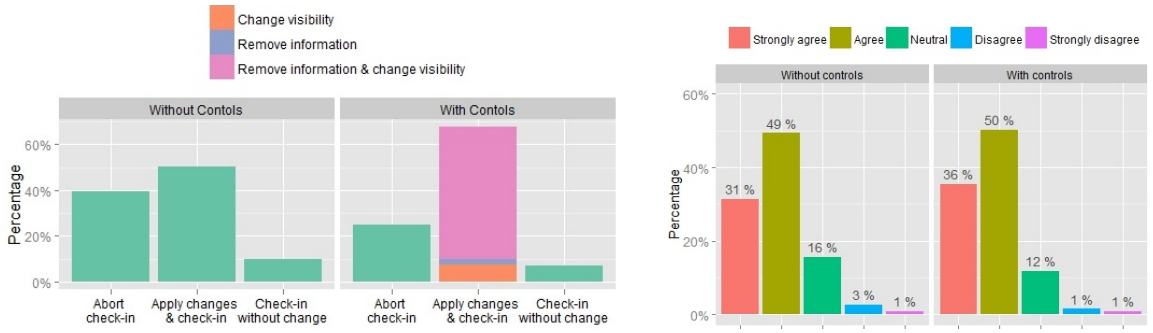


Figure C.11: (The participants' reaction towards the tool effectiveness..

Examining how the privacy feedback tool can influence participants' check-in behaviour in case of feedback only showed that most of them (50%) wanted to change their check-in information and proceed, and 40% to abort the check-in completely, whereas only 10% wanted to check-in without any changes as presented in Figure C.12(a). Opting out from sharing the check-in on Twitter and Facebook was the most preferred change by those who wanted to edit their check-in information, followed by removing the friend tag representing 57% and 43% respectively.

However, the participants' willingness towards changing their check-in information using the control options and proceed with the check-in increased by 18% (68% of them) where 85% of them wanted to both remove their inferred information and change the visibility, 10% wanted just to change the visibility, and the remaining to only remove the inferred information. Aborting the check-in is dropped by 15%, and checking-in without changes is dropped by only 3% (representing 25% and 7% of the participants respectively). Thus, providing the privacy controls showed to impact the users' attitude by allowing them to make use of the application while protecting their privacy.

The participants who were tempted to change their check-in information mostly selected deleting the visit pattern(38%), followed by removing the inferred place type (35%) and their relationships (27%). Moreover, 72% of them selected to set the check-in visibility to friends only, whereas the rest chose to make it private. When capturing the tool impact on making informed decisions by the users, as shown in Figure C.12(b), 80% of the participants were in favour that the tool enabled them to make this decision, while this influence increased when the controls are provided (86%).

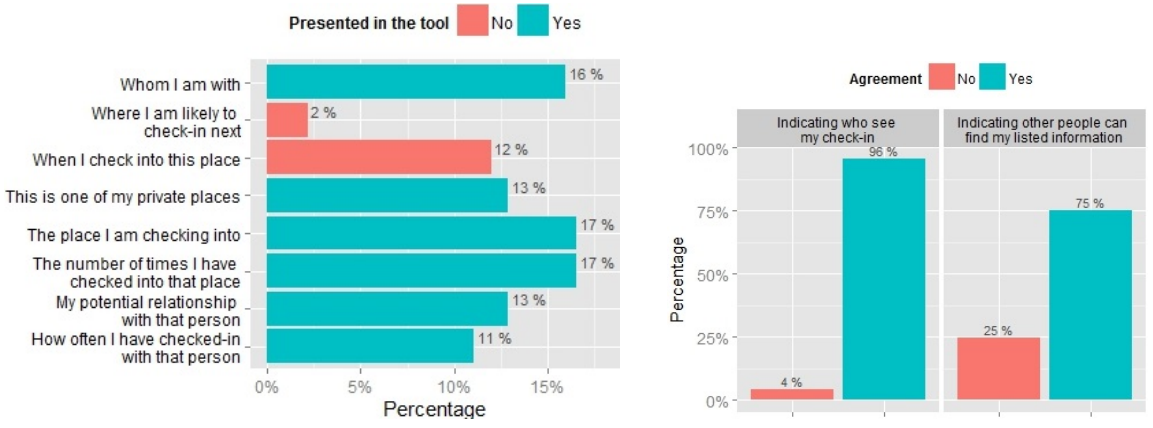


(a) (b)

Figure C.12: (a) The participants' check-in decision with and without the privacy controls, (b) The participants' views about the tool role on making informed decisions..

C.1.0.5 Second Red Check-in Scenario

This scenario has also a risky privacy level. The participants were able to realise on average more than half of the extracted personal information presented in the tool (64%). They mostly captured the information related to "The place I am checking into" and "The number of times I have checked into that place" (17% each of the information options), followed by "Who I am with" (16%) as presented in Figure C.13(a). Moreover, the majority of them reported that they were informed about who can see their check-in (96%) and were aware that other people can find out all of the listed information about them (75%) as illustrated in Figure C.13(b).



(a) (b)

Figure C.13: (a) What revealed infromation by the tool the participants were able to capture, (b) The participants' awareness of their check-in data accessibility achieved by the tool..

In terms of the sufficiency and effectiveness of the content provided,, 76% of the participants reported that the content is enough for showing the privacy implication associated with this check-in as presented in Figure C.14. Most of them (72%) stated that they are concerned about their privacy after being presented with the privacy feedback in this check-in. In addition, 77%

of the participants' views on this check-in privacy level were aligned the proposed level by the privacy tool (Risky-Red).

When exploring the justifications of those 9% who disagreed with the indicated privacy level, some of them actually show their agreement such as "Dangerous, I change my mind on what I think about SWARM" and "I think it should be risky". Others think that this check-in is not risky show just because "There is nothing to worry about" or because "only friends can see" while this check-in is set to be shared on Twitter, and "I am protected by my privacy settings". This portion of participant can either have a limited awareness of how their locations sharing is linked to potential privacy threats, or can be categories as privacy unconcerned based on Westin/ Harris privacy segmentation index since they seems to care less about their privacy [86].

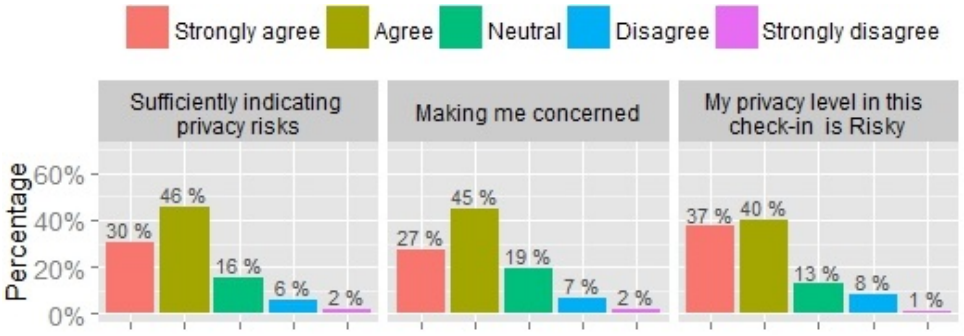


Figure C.14: (The participants' reaction towards the tool effectiveness..

When examining how the participants' behaviour towards this check-in's scenario are impacted by the tool, most of them (47%) wanted to change their check-in information and proceed, followed by 43% who wanted to abort the check-in completely which is highest aborting rate so far, whereas only 10% wanted to check-in without any changes in case of feedback only as presented in Figure C.15(a). Disabling check-in sharing on Twitter is the most desired change to apply by those who wanted to modify their check-in information, followed by removing the friend tag representing 59% and 41% respectively.

Nevertheless, the participants' tendency towards changing their check-in information by using the privacy controls and proceed with the check-in increased by 14% (61% of them) where 89% of them wanted to remove their inferred information and change the visibility, 8% wanted just to change the visibility, and the rest to just remove the inferred information. Aborting the check-in is dropped by 14%, (representing 29%), whereas their willingness to checking-in without changes remained unchanged. For those participants who chose to change their check-in information, removing their private place is the most preferred choice for them, followed by removing the inferred relationships representing 60% and 40% of the selected options. Moreover, 62% of them wanted to set the visibility to friends only, while the rest chose to make it private. The privacy feedback tool also enabled the participants to make informed check-in

decision in case of feedback only (82% of them) as shown in Figure C.15(b), while they the tool was slightly more helpful to decision-making process in the case of having the privacy control options (84%).

(a) (b)

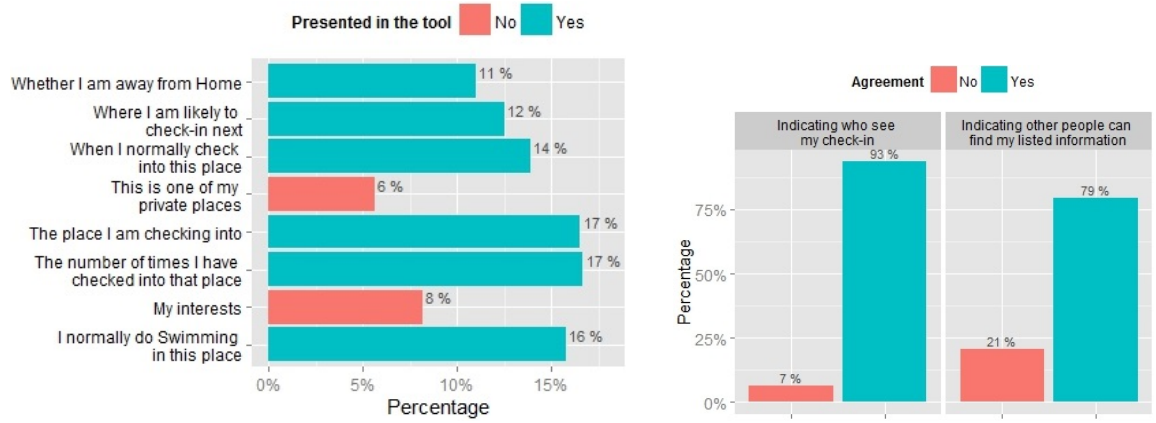
Figure C.15: (a) The participants' check-in decision with and without the privacy controls, (b) The participants' views about the tool role on making informed decisions..

C.1.0.6 Third Red Check-in Scenario

This is the last scenario in the red (risky) privacy level category. The participants were able to pin point most of the revealed personal information presented in the privacy feedback (63.2%), particularly "The place I am checking into" and "The number of times I have checked into that place" (17% each), followed by "I normally do swimming in this place" (16%) and "When I normally check into this place" (14%) as illustrated in Figure C.16(a). Similarly, the majority of them agreed that the tool shows information about their data accessibility in terms of check-in visibility (93%) and potential information inferences by others (79%) as displayed in Figure C.16(b).

Observing how the privacy tool influence participants' attitude revealed that 75% of the participants agreed that the tool sufficiently indicates the privacy risks related to this check-in as shown in Figure C.17. Most of them (69%) stated that the information provided in the tool triggers their privacy concerns. In addition, most of the participants' opinions (72%) on the privacy level of this check-in were consistent with the ones proposed by the tool (Risky-Red). Only 17% were not where most of their explanations again indicate their limited awareness of how privacy threats can occur since most of them said there is no risk because "Only my friends can see my details. So it's safe." or because "It's a public place.", while the rest just thought that "There is no risk".

The privacy feedback also impacted how the participants' behaved towards this check-in's sharing action where in case of feedback only, most of them (40%) wanted to abort the check-in completely which is the only scenario that scored this level of concern. 37% of them wanted to



(a) (b)

Figure C.16: (a) What revealed infromation by the tool the participants were able to capture, (b) The participants' awareness of their check-in data accessibility achieved by the tool..

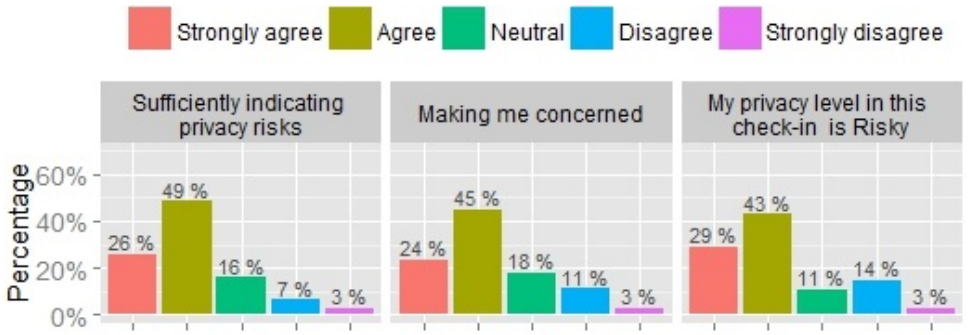


Figure C.17: (The participants' reaction towards the tool effectiveness..

change their check-in information and proceed by removing the tip, whereas only 23% wanted to check-in without any changes as presented in Figure C.18(a). However, the participants' willingness towards changing their check-in information and proceed with the check-in increased by 22% when the privacy controls are provided by the tool (59% of them) where 61% of them wanted to remove their inferred information and change the visibility to private, 36% wanted just remove the inferred information, while the remaining to just change their check-in visibility to private. On the other hand, aborting the check-in is dropped by 12%, and checking-in without changes is also dropped by 10% (28% and 13% of the participants respectively). For those participants who tempted to change their check-in information, deleting their current absence from their private places is the most desirable choice for them (35%), followed by removing where they are likely to check-in next (33%) and their pattern of activity (32%). Privacy feedback showed to influence decision-making process for location disclosure where 81% of the participants reported that the tool enabled them making informed decision comparing to 86% when the privacy controls are offered by the tool as demonstrated in Figure C.18(b).

(a) (b)

Figure C.18: (a) The participants' check-in decision with and without the privacy controls, (b) The participants' views about the tool role on making informed decisions..

C.2 Scenarios and Screen-shots

1) *You are about to check into "Mill Lane Post Office"- a Post Office. You check-in with a friend of yours.*

The screen-shots of this scenario is presented in Figure C.19

(a) (b) (c)

Figure C.19: (The green-level scenario with the location privacy awareness tool shown in (a) with the privacy notice shown in the (b) feedback-only condition, (c) feedback and control condition..

2) *You are about to check into "Chimichanga"- a mexican restaurant. You choose to share this check-in on Twitter account.*

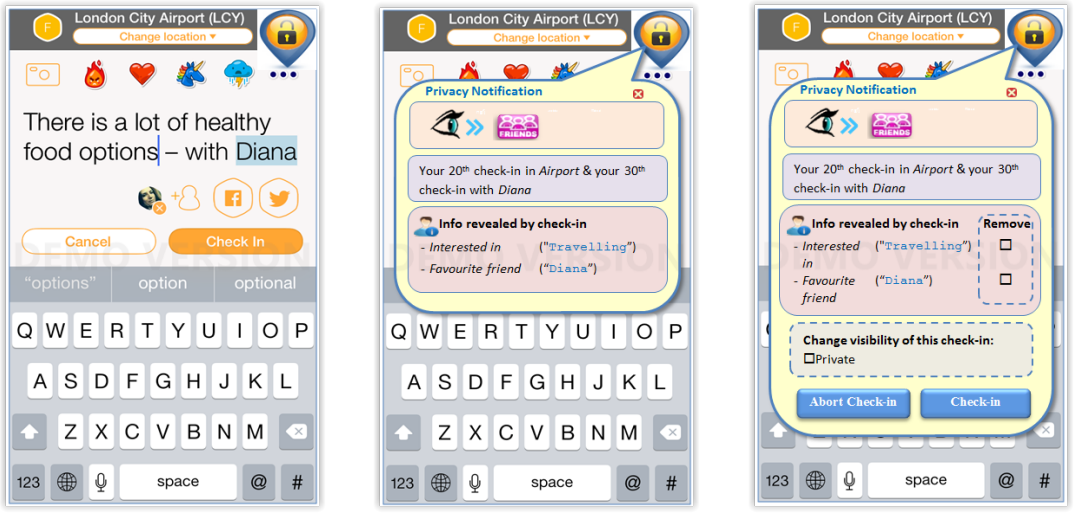
The screen-shots of this scenario is presented in Figure C.20

(a) (b) (c)

Figure C.20: (The first amber-level scenario with the location privacy awareness tool shown in (a) with the privacy notice shown in the (b) feedback-only condition, (c) feedback and control condition..

3) *You are about to check into "London City Airport"- an Airport. You write a tip and check-in with a friend of yours.*

The screen-shots of this scenario is presented in Figure C.21

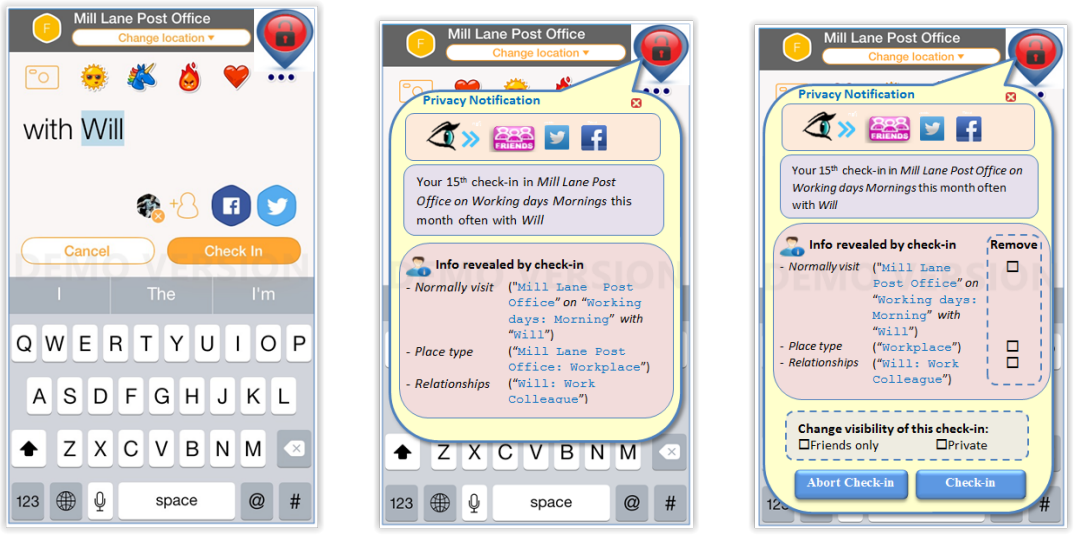


(a) (b) (c)

Figure C.21: (The second amber-level scenario with the location privacy awareness tool shown in (a) with the privacy notice shown in the (b) feedback-only condition, (c) feedback and control condition..

4) *You are about to check into "Mill Lane Post Office"- a Post Office. You check-in with a friend of yours and choose to share this check-in on Facebook and Twitter accounts.*

The screen-shots of this scenario is presented in Figure C.22

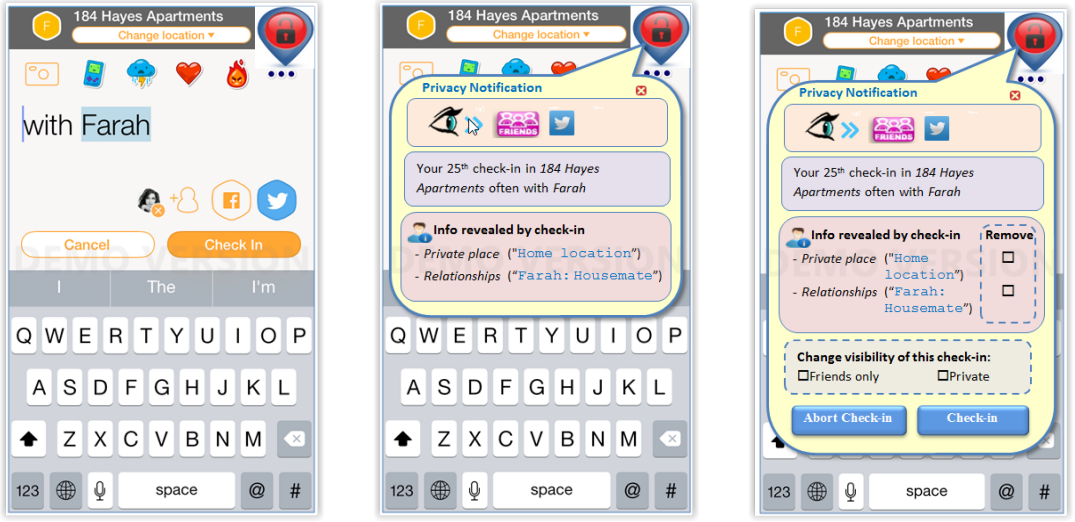


(a) (b) (c)

Figure C.22: (The first red-level scenario with the location privacy awareness tool shown in (a) with the privacy notice shown in the (b) feedback-only condition, (c) feedback and control condition..

5) *You are about to check into "184 Hayes Apartments"- a Home . You check-in with a friend of yours. You choose to share this check-in on Twitter account.*

The screen-shots of this scenario is presented in Figure C.23



(a) (b) (c)

Figure C.23: (The first red-level scenario with the location privacy awareness tool shown in (a) with the privacy notice shown in the (b) feedback-only condition, (c) feedback and control condition..

6) *You are about to check into "Cardiff International Pool"- a Pool. You write a tip.*

The screen-shots of this scenario is presented in Figure C.24

(a) (b) (c)

Figure C.24: (The first red-level scenario with the location privacy awareness tool shown in (a) with the privacy notice shown in the (b) feedback-only condition, (c) feedback and control condition..

C.3 The Survey

This is a user study for examining Location Privacy concerns, awareness and attitude on social networks. Specifically, it presents realistic scenarios of performing check-ins on the Foursquare/Swarm mobile application and tries to assess your awareness and understanding of information collected by the application.

It is important that you consider the questions carefully and to try to provide an honest answer. The study has four main sections and takes about 15 minutes to complete. To ensure that we get some sensible information for our study, a participant needs to have the following qualities:

- Be a user of Swarm.
- Check into venues frequently, and not less than three times a week on average.

C.3.1 Qualities Test

To assure that you meet the requirements of this study, please provide the following information

- I am a user of Swarm application?
  - ○ Yes *[the participant proceed in the survey]*
  - ○ No *[the participant exist the survey]*
- I check-in into venues at least three times a week on average?
  - ○ Yes *[the participant proceed in the survey]*
  - ○ No *[the participant exist the survey]*

C.3.2 Pre-study

In this section, I am hoping to gauge some basic understanding about yourself and your activities with social networking applications.

C.3.2.1 Demographics

- How old are you?
- What is your gender?
  - ○ Male
  - ○ Female
- Where are you from
  - ○ North America
  - ○ South America
  - ○ Europe
  - ○ Africa
  - ○ Asia
  - ○ Australia

C.3.2.2 Social networks Experience

- How often do you check-in on Foursquare
  - ○ Frequently (once or more a day)
  - ○ Moderately (several times a week)
- How often do you enable “Location Services” or other similar location features on your mobile applications?
  - ○ Frequently (it is always on)
  - ○ Moderately (when required by an application)
  - ○ Not sure how to enable this feature
  - ○ I always disable this feature
- How often do you use other social networks; Facebook, Twitter, Instagram, others?
  - ○ Daily
  - ○ Occasionally
  - ○ Rarely

C.3.2.3 Privacy Perceptions

- How likely are you to read through websites "Terms of Services" or "Terms and Conditions" before agreeing to them?
  - ○ Every-time
  - ○ Sometimes
  - ○ Never
- I currently feel safe using Swarm application
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- Do you know who your Friends are on Swarm?
  - ○ All of them
  - ○ Most of them
  - ○ Some of them
  - ○ None of them
- I Check-in on Swarm to let my friends know:
  - □ Where I am now
  - □ A place I like to visit
  - □ What I like about a place
  - □ What I do in a place
  - □ Whom I am with in a place
  - □ Other – please specify [ ]
- I think I know who can see my Check-in data on Swarm
  - 5-point Likert Scale of Strongly agree Strongly disagree
- I know that FourSquare/Swarm shares my personal data with third-party agencies for targeted-marketing/advertising purposes
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- I know that FourSquare/Swarm shares my personal data with third-party agencies to be used for other purposes (other than marketing)

  - *5-point Likert Scale of Strongly agree Strongly disagree*

- Check-ins are

  □ Useful: useful services and recommendations provided to me are relevant.

  □ Enjoyable: I like advertising my locations

  □ Other []

- Check-ins may be harmful or dangerous

  ○ I agree

  ○ I disagree

- Please explain your answer []

- It would be useful to control who can see my Check-in data on Swarm

  - *5-point Likert Scale of Strongly agree Strongly disagree*

- I think my Check-in data are valuable

  - *5-point Likert Scale of Strongly agree Strongly disagree*

- I would sell my Check-in data for:

  ○ $ 5

  ○ $ 20

  ○ $ 50

  ○ Not at all.

- On Swarm, I updated my privacy settings related to:

  □ Who can see my contact information

  □ Visibility of my check-ins to the venue managers

  □ Enabling my friends to check me in and including my name on their social accounts

  □ Sharing my neighbourhood âĂŞlevel location with my friends

  □ Allowing Swarm to use my background location even when the app closed

  □ Allowing Foursquare to serve behavioural targeted ads for outsiders

  □ None

- The provided settings on SWARM are enough to protect my privacy

  - *5-point Likert Scale of Strongly agree Strongly disagree*

- I checked my privacy settings on SWARM

  ○ Recently

  ○ A while back

  ○ Never

- I can remember the following about my Check-ins

  □ Place and time of my last Check-in

  □ Places I visited last month

  □ My top visited places

  □ My Friend that I mostly check-in with

- I am able to retrieve my Check-in history

  ○ Yes, with ease

  ○ Yes, but need to search for the relevant tool

  ○ No

- I know that Foursquare/Swarm collects my location data even if I have not made a Check-in or used the application

  ○ Yes

  ○ No

C.3.3 Swarm Check-in Scenarios with Privacy Notification Tool

The Swarm check-in screen has been modified to include a privacy alert tool. The tool consist of a privacy indicator that is intended to indicate your location privacy level, which you can explore by clicking on it (the "lock" icon) where the privacy notifications are shown. The indicator colour changes to reflect the threat level of the check-in.

The first screen in each scenario is the normal check-in screen with the privacy alert tool and the second screen shows what you get if you click on it to explore its content.

Each scenario will be followed by a set of questions to get your feedback on the location privacy tool.

Figure C.25 shows the layout of the privacy notification tool.

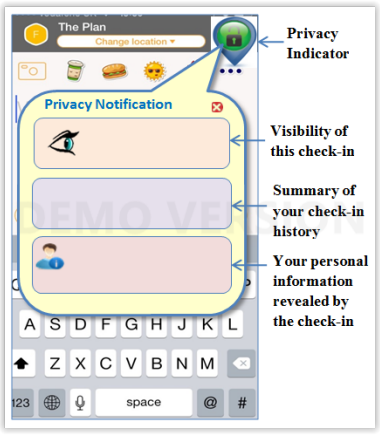


Figure C.25: (The layout of the privacy notification tool..

Each scenario is presented to the participants along with the fist two screen-shots of the scenario (figures (a) and (b)) showing the Swarm screen with the privacy indicator first, then when feedback is displayed (but without privacy controls). Each scenario followed by these set of questions:

1) The tool shows that with my check-in I gave Swarm the following information:

   - *A multiple choice of a list including information presented and others not presented in the tool. For example:*
     - ☐ The place I am checking into
     - ☐ Where I am likely to check-in next
     - ☐ The places I like to visit

2) The tool indicates that my Friends on *Swarm* and on *Twitter* can see my check-in

   - ○ Yes
   - ○ No

3) The tool indicates that other people can find out ALL of the listed information about me

   - ○ Yes
   - ○ No

4) The information provided sufficiently indicates possible privacy risks to myself

   - *5-point Likert Scale of Strongly agree Strongly disagree*

5) If you think the information is not sufficient, please explain: []

6) The information provided by the tool makes me concerned about my privacy

   - *5-point Likert Scale of Strongly agree Strongly disagree*

7) Based on information other people may find about you from this check-in, the privacy risk level is set to MODERATE. Do you agree?

   - *5-point Likert Scale of Strongly agree Strongly disagree*

8) Please explain your answer []

9) Based on the information provided by the tool, what would you do with this check-in?

   ○ Change my check-in information and proceed (Go to Q10)

   ○ Abort the check-in (Go to Q11)

   ○ Proceed with the check-in without change (Go to Q11)

10) What would you change in the check-in?

    - *A multiple choice of a list including options that a user can manually manage. For example:*

      □ Remove sharing on Twitter

      □ Remove my friend tag

11) The tool helps me make an informed decision when checking-in

    - *5-point Likert Scale of Strongly agree Strongly disagree*

C.3.4 Privacy Notification Tool with Control Options

In this section, our privacy alert notification is modified to allow you to control your check-in in different ways.

Again, the scenarios are shown and then followed by a set of questions to get your feedback on the location privacy tool.

The same scenarios are presented here to the participants along with the third screen-shot of the scenarios (figure (c)) showing the Swarm screen when feedback is displayed with privacy controls). Each scenario followed by these set of questions:

1) What would you remove of your revealed information?

- *A multiple choice of a list including the personal information revealed by this check-in and presented in the tool. For example:*
  - □ My interests
  - □ My favourite friend
  - □ Nothing (no change)

2) What would you change the visibility of this check-in to:
   - ○ Friends only
   - ○ Private
   - ○ No change

3) What would you do in this check-in?
   - ○ Apply my previous changes to the control options and check-in
   - ○ Abort check-in
   - ○ Check-in without change

4) The Control Options provided will help me make an informed decision when checking-in
   - *5-point Likert Scale of Strongly agree Strongly disagree*

C.3.5 Post-study

In this section, I am hoping to capture your opinion about the previously presented privacy alert tool and your privacy perceptions.

C.3.5.1 Your Privacy Attitude Based on the Scenarios

- Before this study, I was NOT fully aware of the personal information given by my check-ins
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- I under-estimated the privacy risk resulting from my check-in activity
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- I am now more concerned about my location privacy
  - *5-point Likert Scale of Strongly agree Strongly disagree*

- I would be cautious when sharing my location information
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- As you can see from the scenarios, Foursquare/Swarm can create a rich profile based on your location information and share the information with third parties, would you like to know:

	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
What of your information are collected	○	○	○	○	○
Who has access to your data	○	○	○	○	○
What your data are used for	○	○	○	○	○

- Would you like to control:

	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
What of your information are collected	○	○	○	○	○
Who has access to your data	○	○	○	○	○
What your data are used for	○	○	○	○	○

C.3.5.2 Reaction Towards the Tool

- I would like to use such a tool
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- The tool is informative
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- The tool increases my Awareness about my privacy
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- The tool allows me to Control my privacy
  - *5-point Likert Scale of Strongly agree Strongly disagree*

- The tool would increase my TRUST in the application
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- The tool will make me use the application more frequently
  - *5-point Likert Scale of Strongly agree Strongly disagree*

C.3.5.3 The Tool Design Implications

- The Privacy indicator on the screen is noticeable
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- The Lock icon is helpful for indicating my privacy status
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- The colour scheme used for the privacy Lock is logical
  (Green-Safe — Yellow-Moderate — Red-Risky)
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- Knowing that the privacy notification is only shown when clicking on the indicator, the tool would not disturb my use of Swarm when checking in
  (Lock icon opens up the privacy notification panel only when clicked)
  - *5-point Likert Scale of Strongly agree Strongly disagree*

Appendix D

The Survey Scenarios and Questions for Modelling Location Privacy Perceptions in GeoSNs

This section provides the survey scenarios and questions of 6 treatment out of the 12 which are basically the three data dimensions with Friends visibility and Realistic awareness and then with Public visibility and Attackers' view awareness. These 6 examples can be used to replicated the other 6 treatments which involves Friends visibility and Attackers' view awareness and then with Public visibility and Realistic awareness since the scenarios are the same with just changes visibility scoop of the sharing question with friends or other users (public).

1) Spatial/ Friends/ Realistic

You are using a social network that enables you to share your location and you have set the Visibility of your profile to Friends, so only your friends can see all your shared locations. Imagine that you are using the application in the different scenarios below and think about what you will do in these situations.

1) You are in the Central Shopping Mall in Town. You are looking for a new pair of shoes. You may have been here Occasionally in the past. [**insensitive**]

   Would you share your location now with friends?

   ○ Yes

   ○ Maybe

   ○ No

2) You are in Cineworld cinema Complex in Town. You are there to watch a movie. You have been here Once a couple of months ago. [**insensitive**]

   Would you share your location now with friends?

   - *Same options as in 1.*

3) You are in a Mexican Restaurant in Town. You are having dinner with a friend. This is your First time in this place. [**insensitive**]

   Would you share your location now with friends?

   - *Same options as in 1.*

4) You are in the VFit fitness Centre near your home. You are there for your yoga class. You may have been here a Few times before. [**insensitive**]

   Would you share your location now with friends?

   - *Same options as in 1.*

5) You are in the Main Hospital in Town. You are there for your routine check-up. You visited this hospital only Once in last year. [**sensitive**]

   Would you share your location now with friends?

   - *Same options as in 1.*

6) You are in the Main Hospital in Town. You are there with a friend who has an appointment. You visited this hospital yourself only Once last year. [**sensitive**]

   Would you share your location now with friends?

   - *Same options as in 1.*

7) You are in a Religious Centre (such as a Church, a Temple, or a Mosque ...) that you belong to. You have been here a Few times in the past. [**sensitive**]

Would you share your location now with friends?

- *Same options as in 1.*

8) You are attending a meeting in your local political party headquarters (conservative/ democratic, liberal, ...). You have been here Twice lately. [**sensitive**]

   Would you share your location now with friends?

   - *Same options as in 1.*

9) You are visiting a local community centre belonging to your ethnic group (Greek/Asian/Chinese ...). You are there attending a religious festival celebration. You visited this place Once in the past. [**sensitive**]

   Would you share your location now with friends?

   - *Same options as in 1.*

10) You are now visiting your friend at 16 Park Place (an apartment building in town). You have not been here previously. [**personal**]

    Would you share your location now with friends?

    - *Same options as in 1.*

If you answered "Maybe" or "No" to any of the previous questions, then please explain your answer (why you selected maybe or not to share location) [open-ended]

2) Spatial/ Public/ Attackers' View

You are using a social network application that enables you to track the places you visit. Over time, it will record your Location Track including the places you check-in/visit, when you checked-in these places and with whom. You are now able to share your location tracks built while you are using the application.

Imagine that you are using the application in the different scenarios below. You will be asked whether you would share your location tracks with Other Users of the application.

Please consider your decision of whether to share your location tracks based only on the situation in every scenario separately.

1) You are now in the Central Shopping Mall in Town. You are looking for a new pair of shoes. You have been here Occasionally this place in the past. If you Share your location track, Other users on the social network will be able to See that you have checked in this shopping mall occasionally in the past. [**insensitive**]

   Would you share your location track now with other users?

   ○ Yes
   ○ Maybe
   ○ No

2) You are in Cineworld Cinema Complex in Town. You are there to watch a movie. You have been here Once a couple of months ago. If you Share your location track, Other users on the social network will be able to See that you have checked in this cinema in the past. [**insensitive**]

   Would you share your location track now with other users?

   - *Same options as in 1.*

3) You are in a Mexican Restaurant in Town. You are having dinner with a friend. This is your First time in this place. This restaurant is not shared on your profile yet. [**insensitive**]

   Would you share your location track now with other users?

   - *Same options as in 1.*

4) You are in the VFit fitness Centre near your home. You are there for your yoga class. You may have been here a Few times before. If you Share your location track, Other users on the social network will be able to See that you have checked in this Fitness Centre in the past. [**insensitive**]

Would you share your location track now with other users?

- *Same options as in 1.*

5) You are now in the Main Hospital in Town. You are there for your routine check-up. You have visited this hospital only Once last year. If you Share your location track, Other users on the social network will be able to See that you have checked in this hospital last year. [**sensitive**]

Would you share your location track now with other users?

- *Same options as in 1.*

6) You are now in the Main Hospital in Town. You are there with a friend for his appointment. You have visited hospital only Once last year. If you Share your location track, Other users on the social network will be able to See that you have checked in this hospital last year. [**sensitive**]

Would you share your location track now with other users?

- *Same options as in 1.*

7) You are in a Religious Centre (such as a Church, a Temple, a Mosque ...) that you belong to. You have been here a Few times in the past. If you Share your location track, Other users on the social network will be able to See that you have checked in this religious centre few times in the past. [**sensitive**]

Would you share your location track now with other users?

- *Same options as in 1.*

8) You are attending a meeting in your local political party headquarters (conservative/ democratic/liberal ...). You have been here Twice lately. If you Share your location track, Other users on the social network will be able to See that you have checked in this political party's headquarter twice in the past. [**sensitive**]

Would you share your location track now with other users?

- *Same options as in 1.*

9) You are visiting a local Community Centre belonging to your ethnic group (Greek/Asian/Chinese ...). You are there attending a religious festival celebration. You have visited this place Once in the past. If you Share your location track, Other users on the social network will be able to See from your profile that you have checked in this community centre once in the past. [**sensitive**]

   Would you share your location track now with other users?

   - *Same options as in 1.*

10) You are now visiting your friend at 16 Park Place (an apartment building in town). You have not been here previously. This location is not yet in your profile. [**personal**]

    Would you share your location track now with other users?

    - *Same options as in 1.*

If you answered "Maybe" or "No" to any of the previous questions, then please explain your answer (why you selected maybe or not to share location) [open-ended]

3) Spatial-Social/ Friends/ Realistic

You are using a social network that enables you to share your location and you have set the Visibility of your profile to Friends, so only your friends can see all your shared locations.

Alex is a close friend in real life and on the social network application. *Jack* is not a close friend in real life, but is one of your friend connections on the social network application.

Imagine that you are using the application in the different scenarios below. You will be asked whether you would Share your location in different scenarios and whether you would Tag other people as well (By tagging other people we mean that you identify them when your share your location, by for example, mentioning their names or sharing their photo).

1) You are now in Costa's Coffee Shop with Alex. You have been here Occasionally with Alex in the past. [**insensitive/ close friend**]

Would you share your location now with friends and tag Alex as well?

- ○ Yes
- ○ Maybe
- ○ No

2) You are now in Cineworld Cinema Complex in Town to watch a movie. You have Frequently visited this place in the past. [**insensitive**]

Would you share your location now with friends?

- *Same options as in 1.*

3) You are now having a dinner in a Mexican Restaurant in Town. You have been here Frequently in the past. [**insensitive**]

Would you share your location now with friends?

- *Same options as in 1.*

4) You are in the Good Life Pub in Town and have met Jack there. You have been here Several times in the recent past. [**insensitive/ not close friend**]

Would you share your location now with friends and tag Jack as well?

- *Same options as in 1.*

5) You are now in the Main Hospital in Town with Alex. You are both visiting a friend. You have visited this hospital only Once last year. [**sensitive/ close friend**]

Would you share your location now with friends and tag Alex as well?

- *Same options as in 1.*

6) You are in a Religious Centre (such as a Church, a Temple, or a Mosque ...) that you belong to. You have visited this place Frequently in the past. [**sensitive**]

Would you share your location now with friends?

- *Same options as in 1.*

7) You are attending a meeting in your local political party headquarters (conservative/ democratic/liberal ...). You have been here Frequently in the recent past. [**sensitive**]

   Would you share your location now with friends?

   - *Same options as in 1.*

8) You are visiting a local community centre belonging to your ethnic group (Greek/Asian/Chinese ...). Jack is also there attending a religious festival celebration. You have visited this place Occasionally in the past. [**sensitive/ not close friend**]

   Would you share your location now with friends and tag Jack as well?

   - *Same options as in 1.*

9) You are with Alex at your home at 16 Park Place (an apartment building). You have been here Frequently with Alex in the past. [**personal/ close friend**]

   Would you share your location now with friends and tag Alex as well?

   - *Same options as in 1.*

If you answered "Maybe" or "No" to any of the previous questions, then please explain your answer (why you selected maybe or not to share location) [open-ended]

4) Spatial-Social/ Public/ Attackers' view

You are using a social network application that enables you to track the places you visit. Over time, it will record your Location Track including the places you check-in/visit, when you checked-in these places and with whom. You are now able to share your location track built while you are using the application.

Alex is a close friend in real life and on the social network application. *Jack* is not a close friend in real life, but is one of your friend connections on the social network application.

Imagine that you are using the application in the different scenarios below.

You will be asked whether you would Share your location tracks with Other Users of the application in different scenarios and whether you would Tag other people as well (By tagging other people we mean that you identify them when your share your location, by for example, mentioning their names or sharing their photo).

Please consider your decision of whether to share your location tracks based only on the situation in every scenario separately.

1) You are now in Costa's Coffee shop with Alex. You have been here Occasionally with Alex in the past. If you Share your location track, Other users on the social network will be able to See that you have occasionally both checked-in this place in the past with Alex. [**insensitive/ close friend**]

   Would you share your location track now with other users and tag Alex as well?

   ○ Yes

   ○ Maybe

   ○ No

2) You are now in Cineworld Cinema Complex in Town to watch a movie. You have Frequently visited this place in the past. If you Share your location track, Other users on the social network will be able to See you have frequently checked-in this cinema in the past.[**insensitive**]

   Would you share your location track now with other users?

   - *Same options as in 1.*

3) You are now having a dinner in a Mexican Restaurant in town. You have been here Frequently in the past. If you Share your location track, Other users on the social network will be able to See that you have frequently checked-in this place in the past. [**insensitive**]

   Would you share your location track now with other users?

- *Same options as in 1.*

4) You are in the Good Life Pub in Town and have met Jack there. You have been here Several times in the recent past. If you Share your location track, Other users on the social network will be able to See your association with this pub. [**insensitive/ not close friend**]

   Would you share your location track now with other users and tag Jack as well?

   - *Same options as in 1.*

5) You are now in the Main Hospital in Town with Alex. You are both visiting a friend. You have visited this hospital only Once last year. If you Share your location track, Other users on the social network will be able to See that you have been in this hospital last year. [**sensitive/ close friend**]

   Would you share your location track now with other users and tag Alex as well?

   - *Same options as in 1.*

6) You are in a Religious Centre (such as a Church, a Temple, or a Mosque ...) that you belong to. You have visited this place Frequently in the past. If you Share your location track, Other users on the social network will be able to See that you have checked-in this religious centre frequently in the past. [**sensitive**]

   Would you share your location track now with other users?

   - *Same options as in 1.*

7) You are attending a meeting in your local political party headquarters (conservative/ democratic/liberal...). You have been here Frequently in the recent past. If you Share your location track, Other users on the social network will be able to See you have been visiting this political centre frequently in the recent past. [**sensitive**]

   Would you share your location track now with other users?

   - *Same options as in 1.*

8) You are visiting a local community centre belonging to your ethnic group (Greek/Asian/Chinese..). Jack is also there attending a religious festival celebration. You have visited this place Occasionally in the past. If you Share your location track, Other users on the social network will be able to See your association with this community centre. [**sensitive/ not close friend**]

   Would you share your location track now with other users and tag Jack as well?

   - *Same options as in 1.*

9) You are with Alex at your home at 16 Park Place (an apartment building). You have been here Frequently with Alex in the past. If you Share your location track, Other users can probably guess that this address is your home address and that Alex is your close friend. [**personal/ close friend**]

   Would you share your location track now with other users and tag Alex as well?

   - *Same options as in 1.*

If you answered "Maybe" or "No" to any of the previous questions, then please explain your answer (why you selected maybe or not to share location) [open-ended]

5) Spatial-Social-Temporal/ Friends/ Realistic

You are using a social network that enables you to share your location and you have set the Visibility of your profile to Friends, so only your friends can see all your shared locations.

Alex is a close friend in real life and on the social network application.

Imagine that you are using the application in the different scenarios below. You will be asked whether you would Share your location in different scenarios and whether you would Tag other people as well (By tagging other people we mean that you identify them when your share your location, by for example, mentioning their names or sharing their photo).

1) It is now Wednesday and you are at the Authentic Lebanese restaurant with Alex. You go there Most Wednesdays. [**insensitive/ close friend**]

   Would you share your location now with friends and tag Alex as well?

○ Yes

○ Maybe

○ No

2) You are now in the Weight Watching Clinic on Oxford Street. You Regularly go there with Alex on Thursdays over the past couple of months. [**sensitive/ close friend**]

Would you share your location now with friends and tag Alex as well?

- *Same options as in 1.*

3) It is now Friday night and you are in the Village hall in your neighbourhood. You attend a drama group Every Friday night in Spring. [**insensitive**]

Would you share your location now with friends?

- *Same options as in 1.*

4) It is now Saturday evening and you are in the Good Life Pub in Town. You Regularly go there on Weekends. [**insensitive**]

Would you share your location now with friends?

- *Same options as in 1.*

5) You have been attending meetings in your local political party headquarters (conservative/ democratic/liberal ...) on Weekday Mornings over the past few months in the Village Hall. You are Now at the Village Hall for another group meeting. [**sensitive**]

Would you share your location now with friends?

- *Same options as in 1.*

6) You Regularly go to Cineworld cinema Complex in Town most Tuesday Evenings. It is Now Tuesday Evening and you are in a Mexican Restaurant in town with Alex. [**insensitive/ close friend**]

Would you share your location now with friends and tag Alex as well?

- *Same options as in 1.*

7) You are now participating in a charity event in a Religious Centre (such as Church, Temple, Mosque ...) that you belong to. You have Regularly visited this place on Saturday Afternoons in the last three months. [**sensitive**]

   Would you share your location now with friends?

   - *Same options as in 1.*

8) You are visiting a local community centre belonging to your ethnic group (Greek/Asian/Chinese ...) with your friend Alex. You are there attending a religious festival celebration. You Regularly visit this centre in the Weekends. [**sensitive/ close friend**]

   Would you share your location now with friends and tag Alex as well?

   - *Same options as in 1.*

9) You are with Alex at your home at 16 Park Place (an apartment building). He Normally visit on Sunday Evenings. [**personal/ close friend**]

   Would you share your location now with friends and tag Alex as well?

   - *Same options as in 1.*

10) You are now on a short trip (camping/leisure/work/..) away from home. [**personal**]

    Would you share your location now with friends?

    - *Same options as in 1.*

If you answered "Maybe" or "No" to any of the previous questions, then please explain your answer (why you selected maybe or not to share location) [open-ended]

6) Spatial-Social/ Public/ Attackers' view

You are using a social network application that enables you to track the places you visit. Over time, it will record your Location Track including the places you check-in/visit, when you checked-in these places and with whom. You are now able to share your location track built while you are using the application.

Alex is a close friend in real life and on the social network application.

Imagine that you are using the application in the different scenarios below.

You will be asked whether you would share your location tracks with your Other Users (that you do not know) of the application in different scenarios and whether you would tag other people as well (By tagging other people we mean that you identify them when your share your location, by for example, mentioning their names or sharing their photo).

Please consider your decision of whether to share your location tracks based only on the situation in every scenario separately.

1) It is now Wednesday and you are at the Authentic Lebanese restaurant with Alex. You go there Most Wednesdays. If you Share your location track, Other users of the application will be able to See that you attend this restaurant regularly with Alex most Wednesdays. [**insensitive/ close friend**]

   Would you share your location track now with other users and tag Alex as well?

   ○ Yes

   ○ Maybe

   ○ No

2) You are now in the Weight Watching Clinic on Oxford Street. You Regularly go there with Alex on Thursdays over the past couple of months. If you Share your location track, Other users of the application will be able to See that you attended this clinic regularly with Alex on Thursdays over the past couple of months. [**sensitive/ close friend**]

   Would you share your location track now with other users and tag Alex as well?

   - *Same options as in 1.*

3) It is now Friday night and you are in the Village hall in your neighbourhood. You attend a drama group Every Friday night in Spring. If you Share your location track, Other users of the application will be able to See that you have been at the village hall on Friday nights recently. [**insensitive**]

   Would you share your location track now with other users?

   - *Same options as in 1.*

4) It is now Saturday evening and you are in the Good Life Pub in Town. You Regularly go there on Weekends. If you Share your location track, Other users of the application will be able to See that you regularly go to this pub on Weekends. [**insensitive**]

   Would you share your location track now with other users?

   - *Same options as in 1.*

5) You have been attending meetings in your local political party headquarters (conservative/ democratic/liberal ...) on Weekday Mornings over the past few months in the Village Hall. You are now at the Village Hall for another group meeting. If you Share your location track, Other users of the application will be able to See that you have been attending this political party group meeting at the village hall on Weekday mornings over the past few months. [**sensitive**]

   Would you share your location track now with other users?

   - *Same options as in 1.*

6) You Regularly go to Cineworld Cinema Complex in Town most Tuesday evenings. It is now Tuesday evening and you are in a Mexican Restaurant in town with Alex. If you Share your location track, Other users of the application may recognise that you have been going to this cinema on Tuesday evenings, and that this Tuesday evening you are in a restaurant with Alex. [**insensitive/ close friend**]

   Would you share your location track now with other users and tag Alex as well?

   - *Same options as in 1.*

7) You are now participating in a charity event in a Religious Centre (such as Church, Temple, Mosque ...) that you belong to. You have Regularly visited this place on Saturday Afternoons in the last three months. If you Share your location track, Other users of the application can See that you visited this religious centre frequently on Saturday afternoons over the last three months. [**sensitive**]

   Would you share your location track now with other users?

   - *Same options as in 1.*

8) You are visiting a local community centre belonging to your ethnic group (Greek/Asian/Chinese..) with your friend Alex. You are there attending a religious festival celebration. You have Regularly visited to this centre in the Weekends. If you Share your location track, Other users of the application can See that you have checked in this community centre frequently in the Weekends. [**sensitive/ close friend**]

   Would you share your location track now with other users and tag Alex as well?

   - *Same options as in 1.*

9) You are with Alex at your home at 16 Park Place (an apartment building). He Normally visits on Sunday Evenings. If you Share your location track, Other users of the application may recognise that this address is probably your home address and that you are there with Alex on Sundays. [**personal/ close friend**]

   Would you share your location track now with other users and tag Alex as well?

   - *Same options as in 1.*

10) You are now on a short trip (camping/leisure/work/..) away from home. From your profile, Other users of the application will recognise your home address- where you check-in regularly in the evenings If you Share your location track, they may also realise that you are now Away from home. [**personal**]

    Would you share your location track now with other users?

    - *Same options as in 1.*

If you answered "Maybe" or "No" to any of the previous questions, then please explain your answer (why you selected maybe or not to share location) [open-ended]

Appendix E

Supporting Materials for The Study on Towards on-demand Geo-Profile Visualiser for Privacy Awareness on GeoSNs

This appendix firstly presents the results for the Cognitive Walkthrough usability evaluation for the initial prototype of Geo-Profile visualiser. Then, screen-shots of selected participants' geo-profiles are presented as examples, followed by the interviews' questions of the user-based study conducted to evaluate the prototype.

E.1 The Walk Through

E.1.1 Tasks to Be Tested

The aim in this test is to evaluate the core function of the system which is enhancing the users' privacy awareness in terms of personal information collection, inferences and accessibility. Therefore, it uses the initial system prototype that shows the base design in terms of the interactive visualisation of the privacy-oriented geo-profile and the privacy awareness interfaces presenting the user profile labelling, the direct privacy warning message and information attributes of each of the warning. However, the prototype represents the skeleton of the system with no actual user profile information shown.

E.1.2 System Users and Structure

Prior to conducting the evaluation, it is essential to define the targeted users and the actual structure to complete any provided task. Basically, any user who shares location and location-related information on their social network accounts can use this proposed system. The action

Table E.1: The action sequence for finding and reaching a task.

Step 1: Node	Step 2: Sub-node	Step 3: Tasks Completed
My Places	All visited places	Showing privacy information of user's visited places and when
	Favourite	Showing privacy information of user's favourite places
	Routine Visits	Showing privacy information of user's routine visit to places and,when
My Interests/ activities	All Interests	Showing privacy information of user's overall interests and if they,are in common with friends
	Favourite	Showing privacy information of user's favourite interests and the,related places
	Routine Interests	Showing privacy information of user's routine interests and the,related place and time
My Friendships	All Friends	Showing privacy information of user's friends on the social networks, and their in common interests and activities as well as co-locations
	Favourite	Showing privacy information of user's favourite friends
	Routine Meetings	Showing privacy information of user's routine meetings and the,related place, time, interest, and activity

sequence for finding and reaching a task revolves around nodes and sub-nodes sequence in the base design of privacy-oriented user profile which is demonstrated in Table E.1.

E.1.3 Evaluation Procedure

To run the evaluation, a series of relevant guideline criteria in form of questions that describe the links of how the user can progress with achieving a task need to be answer by the analyst with each step of the task using the proposed interfaces which are [141]:

1) Will the user try to achieve the right effect? (I.e. will the user understand the in-between task needed to achieve the main task?)

2) Will the user notice that the correct action is available? (I.e. is the action visible)

3) Will the user associate the correct action with the effect they are trying to achieve? (I.e. will the user understand that this action is related to the task the user want to accomplish)

4) If the correct action is performed, will the user see that progress is being mad toward solution of their task? (I.e. will the user be presented with appropriate feedback after performing an action indicating that the user is progressing successfully?)

The evaluation is carried out on the task related to âĂŸMy Places' node as a sample for two reasons; the first is all other node has similar pattern of the tasks offered in this node and the second is the tasks are more complex than tasks in the other nodes (some tasks has the same steps and interface). âĂČ

E.1.4 The Walkthrough

In this section, the Walkthrough analysis is presented as a story for each of the selected tasks. The first step is the same among all of these tasks and hence it is presented here to prevent repetition and the following steps for each task are presented next. The answers for all previous

questions will be within the steps. Each evaluated task will be labelled as 'Failure story' if one of the criteria is fail, and 'Success story' otherwise.

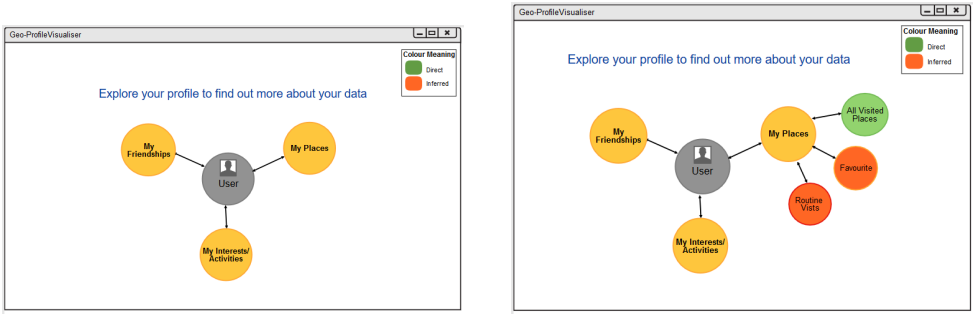
Step 1: Finding the representative node

1) User is trying to find the representative node (My Places) and related to visited places because the system asks for exploring the profile and because it is part of task as shown in Figure E.1(a) .

2) User uses the interactive user profile visualisation by clicking on the node because they can see the visible representation of the user profile and the clickable nodes.

3) User knows that clicking on 'My Places' node it the right action because the provided label is related to what they are trying to do and all other nodes seems wrong.

4) User recognises that they are progressing in the right direction because the relevant sub-nodes are shown.



(a) (b)

Figure E.1: The initial prototype design including (a) The main window and (b) when 'My Places' node is clicked..

E.1.4.1 Task 1: Finding the privacy information of user's visited places

Success story

Step 2: Finding the representative sub-node

1) User is trying to find the representative sub-node (All Visited Places) related to visited places because the system asks for exploring the profile and because it is part of task.

2) User uses the interactive user profile visualisation by clicking on the sub-node (All Visited Places) because they can see the visible and clickable representation of the sub- categories of 'My Places' node as demonstrated in Figure E.1(b) .

3) User knows that clicking on ‘All Visited Places’ sub-node it the right action because the provided label is related to what they are trying to do and all other sub-nodes seems wrong.

4) User recognises that they are progressing in the right direction because the privacy warning interface showing the privacy information of user’s visited places and when is presented where they can see direct privacy warning message as well as explore their relevant information as illustrated in Figure E.2.

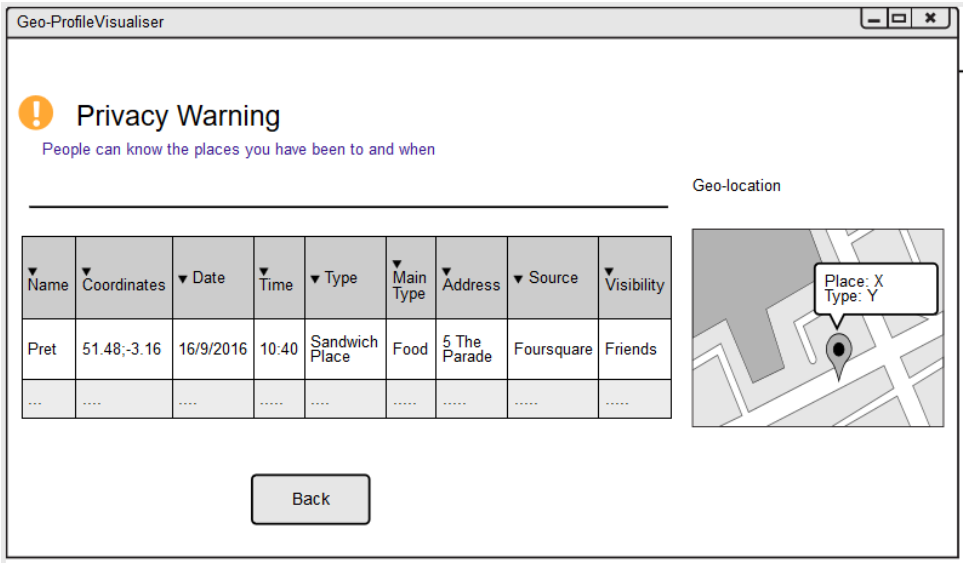


Figure E.2: (The prototype interface shown when clicking on ’All visited places’ sub-node..

E.1.4.2 Task 2: Finding the privacy information of user’s favourite places

Success story

Step 2: Finding the representative sub-node

1) User is trying to find the representative sub-node (Favourite) related to favourite places because the system asks for exploring the profile and because it is part of task.

2) User uses the interactive user profile visualisation by clicking on the sub-node (Favourite) because they can see the visible and clickable representation of the sub-categories of ‘My Places’ node as demonstrated in Figure E.1(b).

3) User knows that clicking on “All Visited Places” sub-node it the right action because the provided label is related to what they are trying to do and all other sub-nodes seems wrong.

4) User recognises that they are progressing in the right direction because the privacy warning interface showing the privacy information of user's favourite places is presented where they can see direct privacy warning message as well as explore their relevant information as illustrated in Figure E.3.

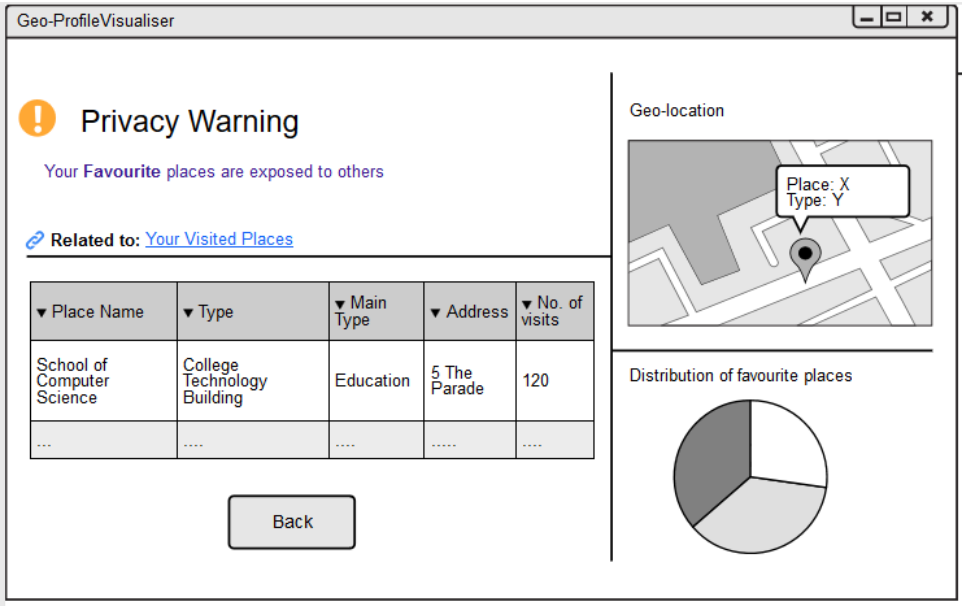


Figure E.3: (The prototype interface shown when clicking on 'Favourite' sub-node..

E.1.4.3 Task 3: Finding the privacy information of user's routine visit to places

Success story

Step 2: Finding the representative sub-node

1) User is trying to find the representative sub-node (Routine Visits) related to favourite places because the system asks for exploring the profile and because it is part of task.

2) User uses the interactive user profile visualisation by clicking on the sub-node (Routine Visits) because they can see the visible and clickable representation of the sub-categories of "My Places" node as demonstrated in Figure E.1(b).

3) User knows that clicking on "Routine Visits" sub-node it the right action because the provided label is related to what they are trying to do and all other sub-nodes seems wrong.

4) User recognises that they are progressing in the right direction because the privacy warning interface showing the privacy information of user's routine visits is presented where

they can see direct privacy warning message as well as explore their relevant information as illustrated in Figure E.4.

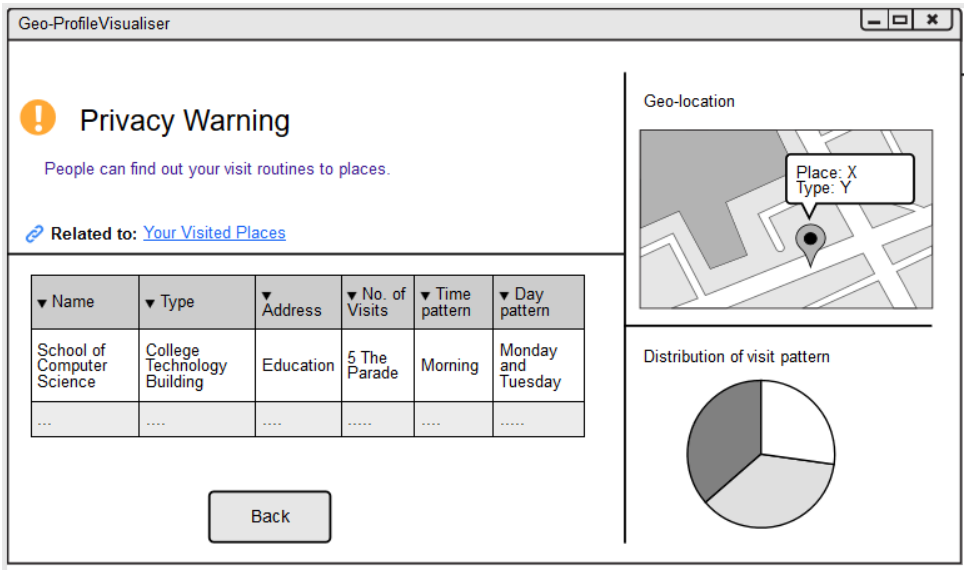


Figure E.4: (The prototype interface shown when clicking on 'Routine Visits' sub-node..

E.2 Examples of Participants' Extracted Geo-profiles

Here, two participants' retrieved geo-profiles using the Geo-profile Visualisation prototype are presented as examples of the rang of personal information that can be extracted.

E.2.1 Example 1: Small Geo-profile

This user has 127 check-in in 66 places and 45 categories. The extracted geo-profile has no visits patterns (in 'Routine Visit' sub-node), and no co-locations with friends (in 'My Friendships' node). The following screen-shots show the retrieved profile.

E.2.2 Example 1: Large Geo-profile

This user has 849 check-in in 165 places and 86 categories. The extracted geo-profile has all information to populate the main node and their sub-nodes as well.

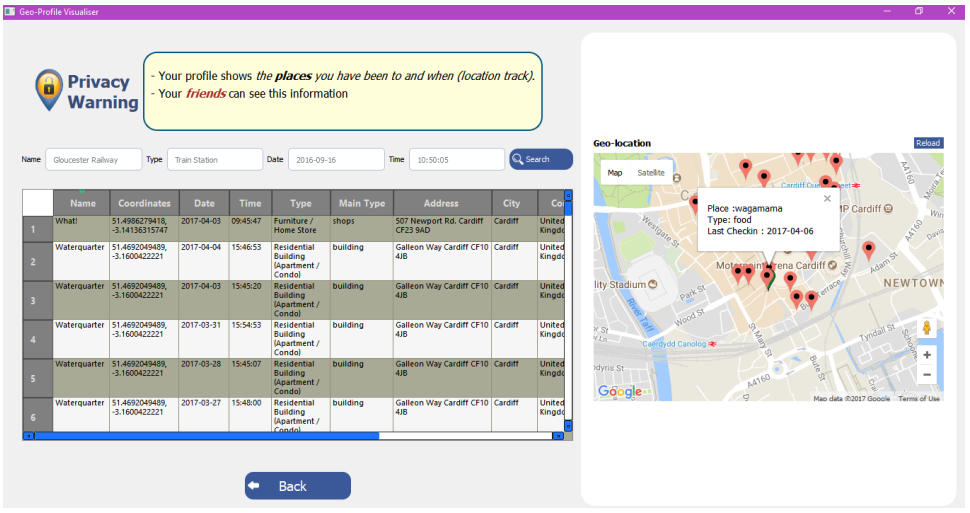


Figure E.5: The user's visited places shown in 'All visited places' sub-node. .

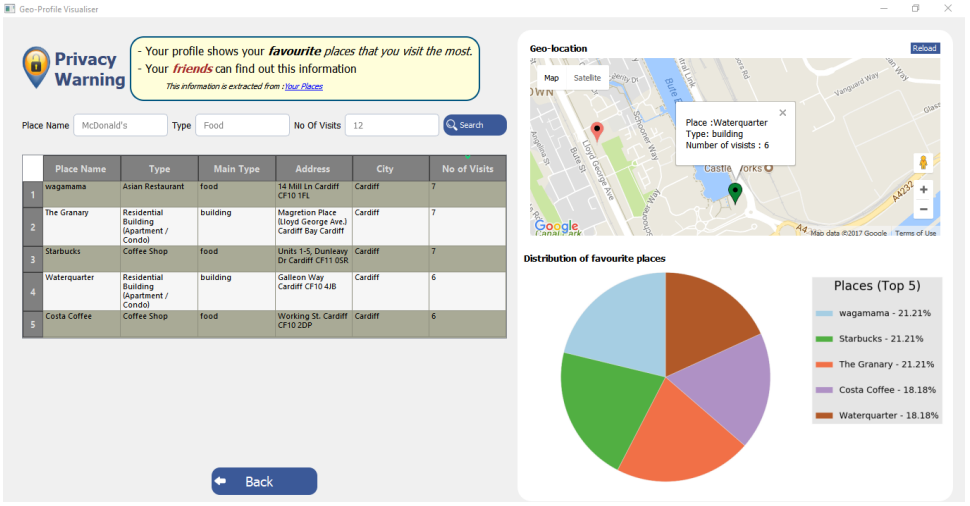


Figure E.6: The user's top places shown in 'Favourite places' sub-node..

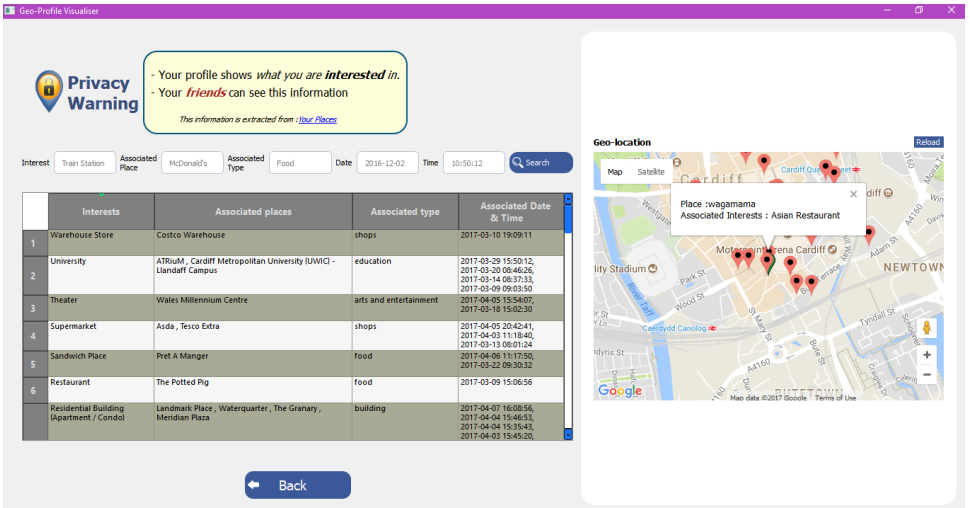


Figure E.7: The user's interests and activities shown in 'All Interests' sub-node..

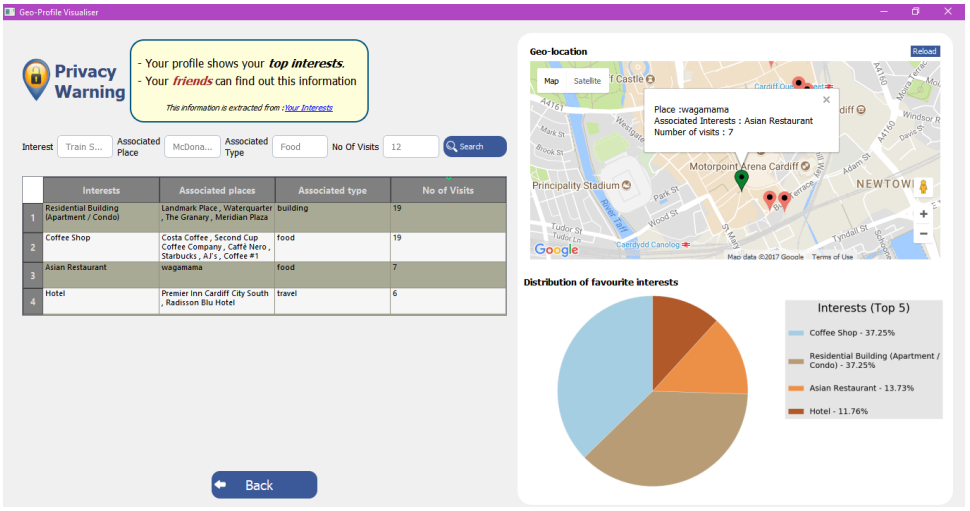


Figure E.8: The user's top interests and activities shown in 'Favourite Interests' sub-node..

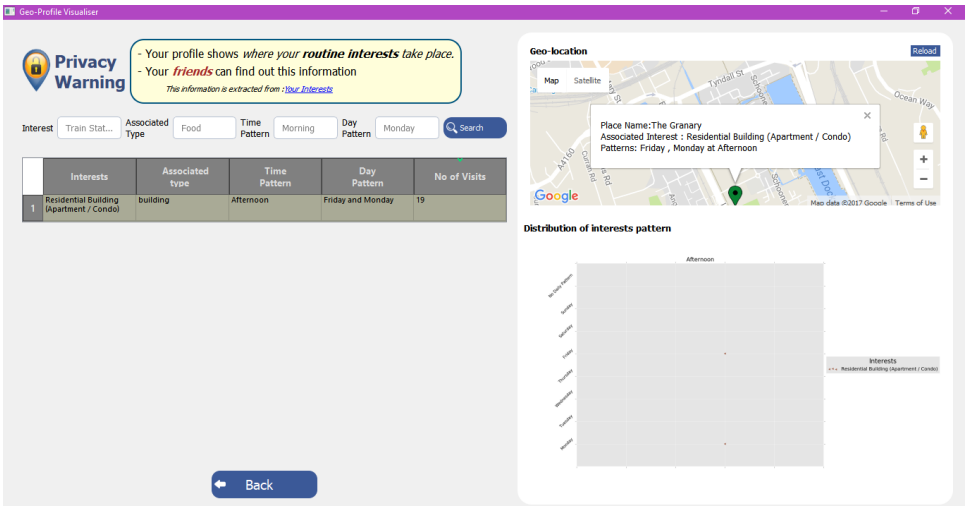


Figure E.9: The user's interests patterns shown in 'Routine Interests' sub-node..

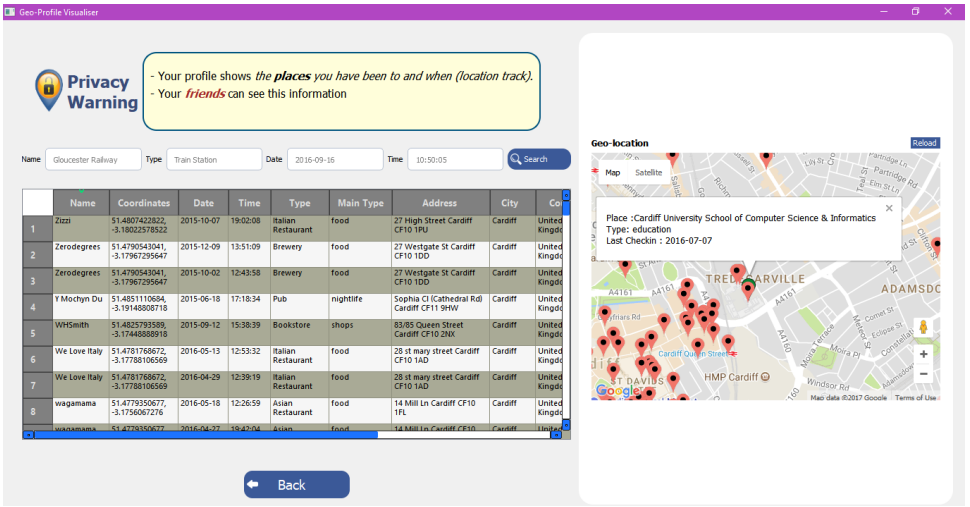


Figure E.10: The user's visited places shown in 'All visited places' sub-node. .

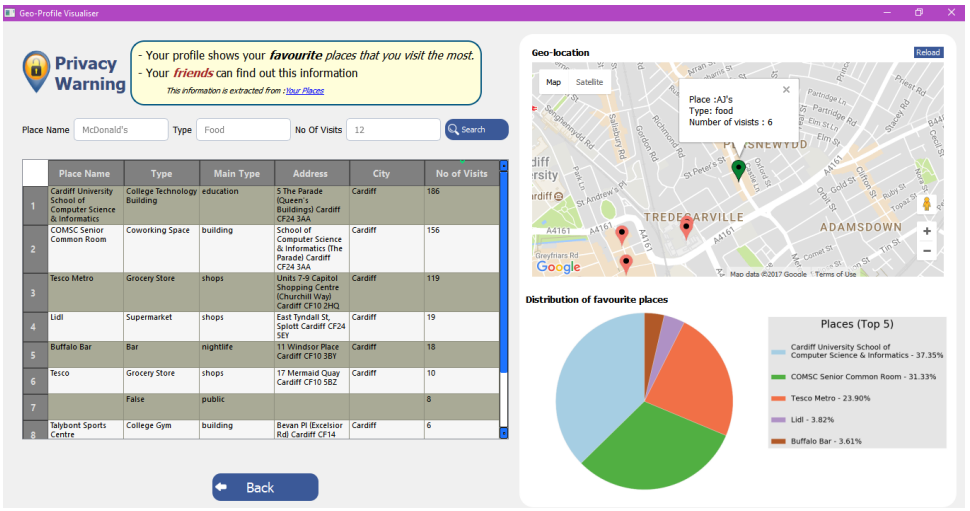


Figure E.11: The user's top places shown in 'Favourite places' sub-node..

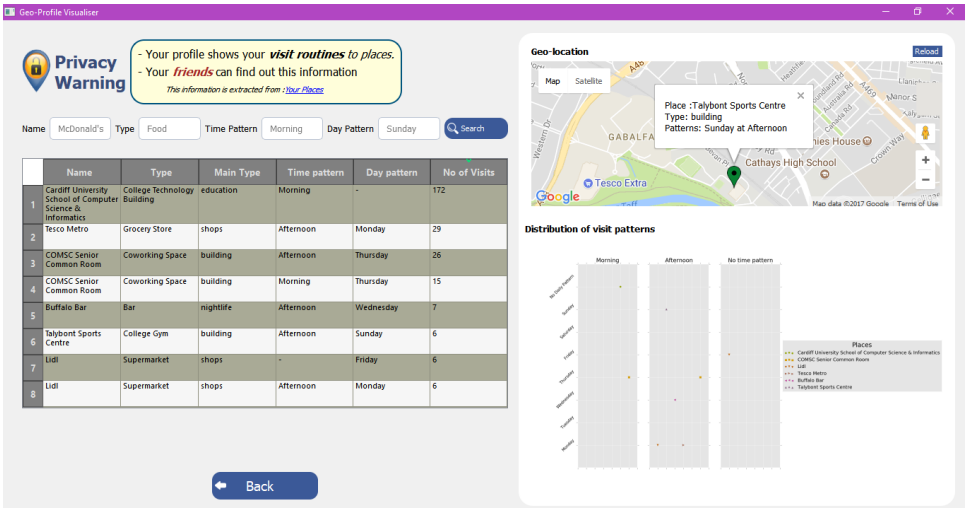


Figure E.12: The user's visits patterns to places shown in 'Routine Visits' sub-node..

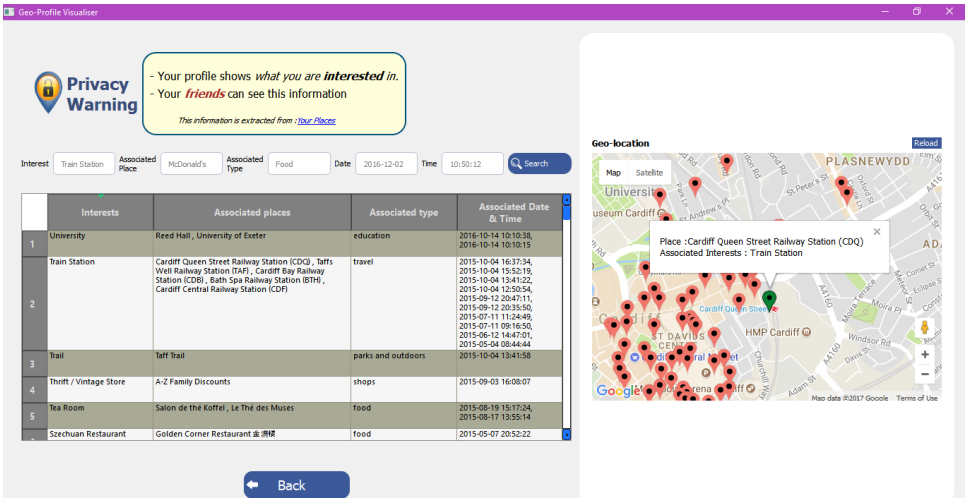


Figure E.13: The user's interests and activities shown in 'All Interests' sub-node..

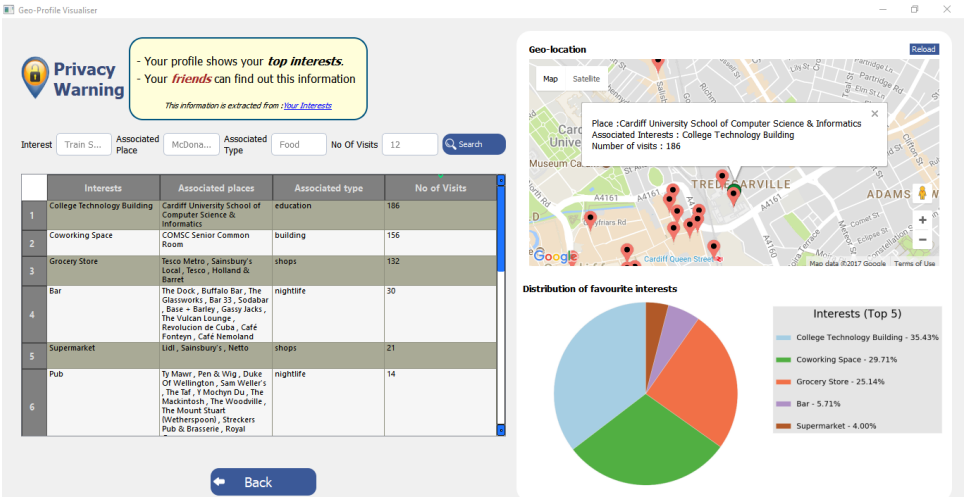


Figure E.14: The user's top interests shown in 'Favourite Interests' sub-node..

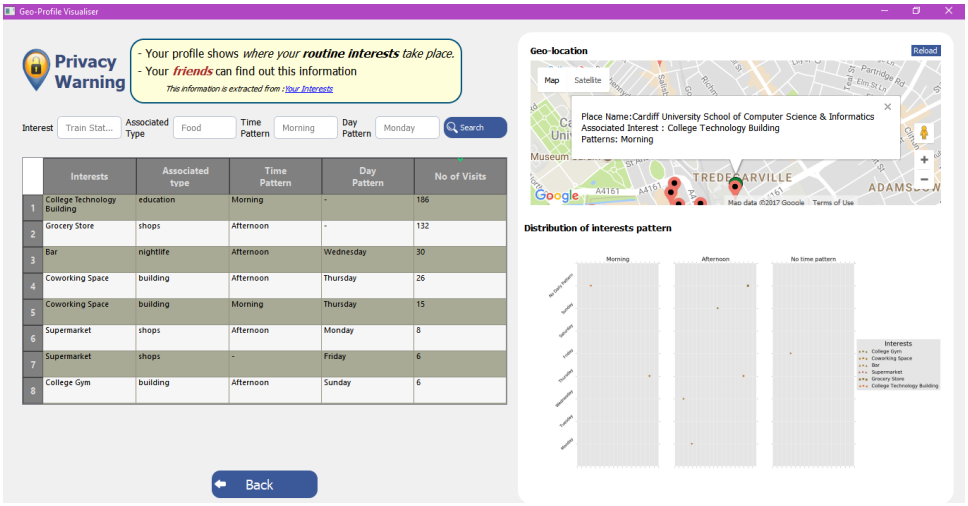


Figure E.15: The user's interests patterns shown in 'Routine Interests' sub-node..

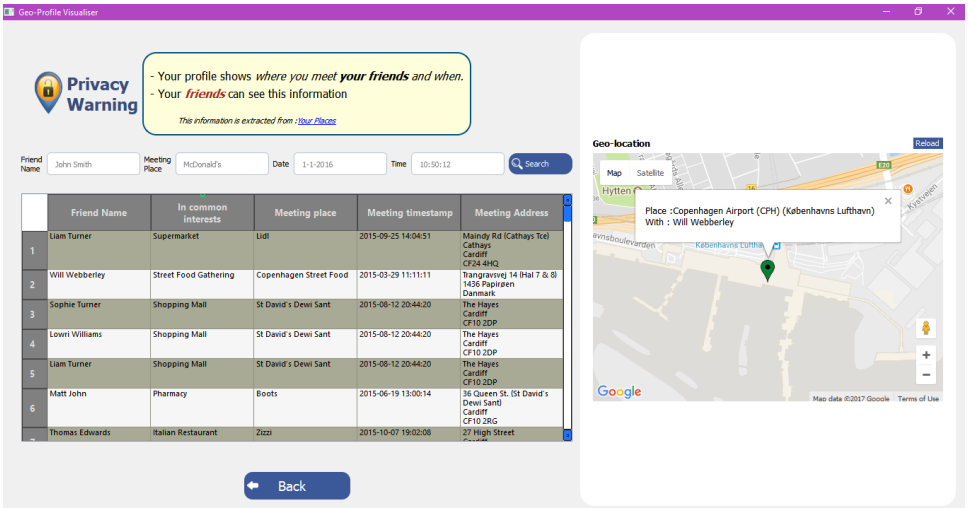


Figure E.16: The user's co-locations with friends shown in 'All Meetings with Friends' sub-node..

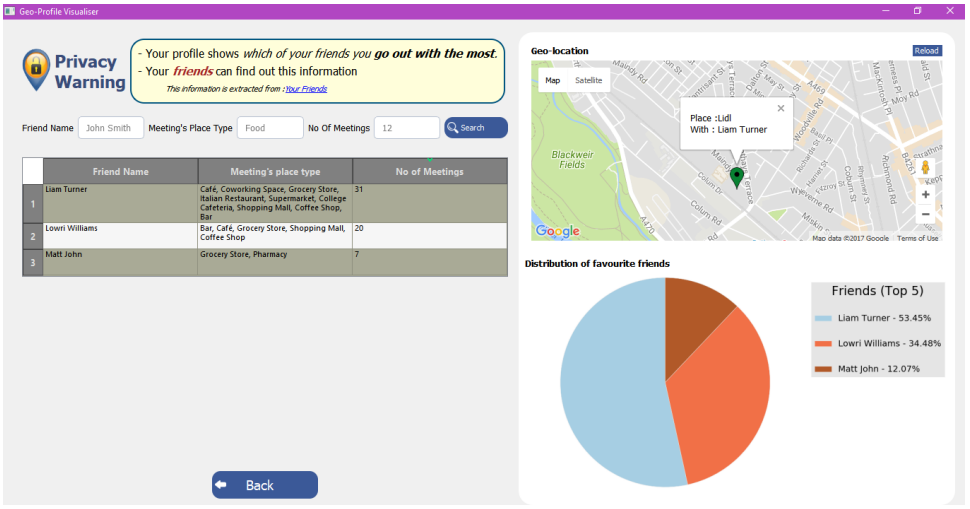


Figure E.17: The user's favourite friends whom had the most co-locations with shown in 'Favourite Friends' sub-node..

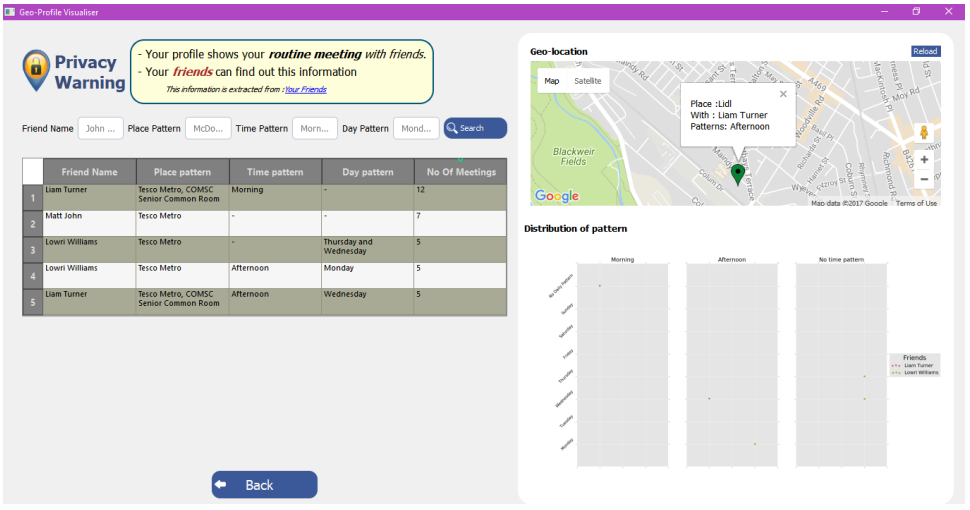


Figure E.18: The user's patterns of co-location with friends shown in 'Routine Meetings' sub-node..

E.3 The Interviews' Questions

E.3.1 Pre-study

E.3.1.1 Demographics

- How old are you?
- What is your gender?
  - ○ Male
  - ○ Female
- What do you work/study?
- Where are you from?
  - ○ North America
  - ○ South America
  - ○ Europe
  - ○ Africa
  - ○ Asia
  - ○ Australia

E.3.1.2 Web and Social Networking Background

- What is your experience in web applications and technologies?
  - ○ Not too experienced
  - ○ Somewhat experienced
  - ○ Experienced
  - ○ Very experienced
- I feel safe using Foursquare/Swarm
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- I am concerned about my online privacy

  - *5-point Likert Scale of Strongly agree Strongly disagree*

- Have you ever regretted sharing certain information online?

  ○ Yes

  ○ No

- Have you ever deleted a post (comment, picture location,..) due to privacy concern?

  ○ Yes

  ○ No

- Have you ever requested to delete your data from a service before (e.g. request to delete the background location in Foursquare or all location information in Twitter)?

  ○ Yes

  ○ No

- Have you ever checked your shared checkins using Foursquare History?

  ○ Yes

  ○ No

- Have you ever used tools/applications that help your mange your online privacy? (e.g. browser add-ons)

  ○ Yes

  ○ No

- If yes, like what?

- Can Foursquare/Swarm collect your location even if you are not using the application

  ○ Yes

  ○ No

  ○ I do not know

- Can Foursquare/Swarm share your data with third-party agencies for targeted-marketing/advertising purposes

  ○ Yes

  ○ No

- ○ I do not know

- Can Foursquare/Swarm shares your data with third-party agencies to be used for other purposes (other than marketing)
  - ○ Yes
  - ○ No
  - ○ I do not know
- If yes, what do you think the other purposes are?
- Who do you think can access you check-in data ?
- Who do you think can access your Swarm check-in data ?
  - □ No one (Private)
  - □ The application (Foursquare/ Swarm)
  - □ My friends on the application
  - □ Other users of the application
  - □ Third parties
  - □ I do not know
- When you share your location on Social Networks, what sort of information can be known about you?
- Have you ever checked your privacy settings in Foursquare/Swarm?
  - ○ Yes
  - ○ No
- How often do you update your privacy setting?
  - ○ Rarely
  - ○ Often
  - ○ Always
- What aspects of your data can you control on Foursquare/Swarm application?
- What of these listed options can you control on Foursquare/Swarm application?
  - □ Who can see my contact information

- □ Visibility of my check-ins to the place managers
- □ Enabling my friends to check me in and including my name on their social media accounts
- □ Whether the application can collect my location when I am not using it (application is closed)
- □ Whether the application can collect my location while I am using it
- □ Check into a place privately (not seen by my friends)
- □ Getting behavioural targeted ads outside the application
- □ Getting behavioural targeted ads inside the application
- □ Deleting all of your check-ins
- □ Deleting your profile
- □ None

E.3.2 The Actual Study

This section involved the use of a location-data access tool specified for each group: Swarm History for the no-awareness group and a Geo-Profile Visualiser tool for the awareness group. For each group, the interviewer started with a brief description of what the tool shows or provides and for a few minutes allowed the participant to explore his/her profile using the tool. Then the participants were asked to carry out pre-defined tasks that were personalised for them on the basis of their generated geo-profile (7 tasks on average), as discussed in Chapter 8.

E.3.2.1 Information Awareness and Privacy Attitude

- This tool helps me understand the information I share when I check-in
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- This tool allows me to view the information I share when I check-in
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- This tool helps me understand the possible information that can be extracted about me when I check-in (e.g. patterns of visits, top interests)
  - *5-point Likert Scale of Strongly agree Strongly disagree*

- This tool allows me to view the possible information that can be extracted about me when I check-in (e.g. patterns of visits, top interests)
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- This tool allows me to know who can access my data
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- This tool motivates me to be more in control of my online data (e.g. deleting posts, updating privacy settings)
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- I am satisfied with the way Foursquare/Swarm collects and stores my data
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- This tool makes me more concerned about my privacy
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- This tool encourages me to alter the way I share my location information to protect my privacy
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- Can the tool help you see/access your data collected/stored? If yes, to what extent?
- Can the tool help you find out who has access to your data? If yes, to what extent? in which ways?
- Can the tool help you find out the kind of information that can be extracted from your data? If yes, like what?
- After carrying out the tasks, are you concerned about your privacy? If yes, what triggers your concern?

E.3.2.2 Usability of Geo-Profile Visualiser Prototype

This section is assigned only for the awareness-group to assess the usability of the Geo-Profile Visualiser prototype.

- It was simple to use the tool

  - *5-point Likert Scale of Strongly agree Strongly disagree*

- This tool provides me with easy access to my location profile (e.g. favourite places, co-locations with friends and my visit routines).
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- This tool organises and presents my location profile in an effective way.
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- The graphic representation of my data in this tool helps me to understand the content easily.
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- The map representation of my data in this tool helps me to understand the content easily.
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- I would use such a tool to learn about and manage my online data.
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- I think that I would like to use this system frequently.
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- I found the system unnecessarily complex.
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- I thought the system was easy to use.
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- I think that I would need the support of a technical person to be able to use this system.
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- I found the various functions in this system were well integrated.
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- I thought there was too much inconsistency in this system.
  - *5-point Likert Scale of Strongly agree Strongly disagree*

- I would imagine that most people would learn to use this system very quickly.
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- I found the system very cumbersome to use.
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- I felt very confident using the system.
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- I needed to learn a lot of things before I could get going with this system.
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- What do you like about the application (design)?
- What do you dislike about the application (design)?

E.3.3 post-study

- I feel safe using Foursquare/Swarm
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- I am concerned about my online privacy
  - *5-point Likert Scale of Strongly agree Strongly disagree*
- I do NOT mind sharing my geo-profiles with others
  - *5-point Likert Scale of Strongly agree Strongly disagree*

The members of the awareness group were asked three further questions:

- Do you think that your initial understanding of you location data collection was limited? How?
- Do you think that your initial understanding of possible utilisation and privacy implications of your shared location data was limited? How?
- Would you change the way you share location data after using this tool? How?
- Would you use such an application? Why?

Bibliography

[1] E. R. T. Digital, "The Infinite Dial 2017 ," tech. rep., 2017.

[2] S. Gambs, O. Heen, and C. Potin, "A comparative privacy analysis of geosocial networks," in *Proceedings of the 4th ACM SIGSPATIAL International Workshop on Security and Privacy in GIS and LBS*, pp. 33–40, ACM, 2011.

[3] D. Freni, C. Ruiz Vicente, S. Mascetti, C. Bettini, and C. S. Jensen, "Preserving location and absence privacy in geo-social networks," in *Proceedings of the 19th ACM international conference on Information and knowledge management*, pp. 309–318, ACM, 2010.

[4] P. Coppens, L. Claeys, C. Veeckman, and J. Pierson, "Privacy in location-based social networks: Researching the interrelatedness of scripts and usage," in *Proceedings of the Symposium on Usable Privacy and Security*, 2014.

[5] M. Duckham and L. Kulik, "Location privacy and location-aware computing," *Dynamic & mobile GIS: investigating change in space and time*, vol. 3, pp. 35–51, 2006.

[6] L. Rossi and M. Musolesi, "It's the way you check-in: identifying users in location-based social networks," in *Proceedings of the second edition of the ACM conference on Online social networks*, pp. 215–226, ACM, 2014.

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

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

[9] Z. Cheng, J. Caverlee, K. Lee, and D. Sui, "Exploring millions of footprints in location sharing services," in *ICWSM*, vol. 2010, pp. 81–88, 2011.

[10] J. Vosecky, D. Jiang, and W. Ng, "Limosa: A system for geographic user interest analysis in twitter," in *Proceedings of the 16th International Conference on Extending Database Technology*, pp. 709–712, ACM, 2013.

[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*, pp. 23–32, 2010.

[12] S. Scellato, A. Noulas, and C. Mascolo, "Exploiting place features in link prediction on location-based social networks categories and subject descriptors," in *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 1046–1054, 2011.

[13] A. R. Beresford and F. Stajano, "Location privacy in pervasive computing," *IEEE Pervasive computing*, vol. 2, no. 1, pp. 46–55, 2003.

[14] L. Barkhuus, "Privacy in location-based services, concern vs. coolness," in *Workshop on Location System Privacy and Control at MobileHCI*, vol. 4, Citeseer, 2004.

[15] K. P. Tang, J. Lin, J. I. Hong, D. P. Siewiorek, and N. Sadeh, "Rethinking location sharing: exploring the implications of social-driven vs. purpose-driven location sharing," in *Proceedings of the 12th ACM international conference on Ubiquitous computing*, pp. 85–94, ACM, 2010.

[16] L. Jin, X. Long, and J. Joshi, "Towards understanding residential privacy by analyzing users' activities in Foursquare," in *Proceedings of the 2012 ACM Workshop on Building analysis datasets and gathering experience returns for security - BADGERS '12*, pp. 25–32, 2012.

[17] P. G. Kelley, J. Bresee, L. F. Cranor, and R. W. Reeder, "A nutrition label for privacy," in *Proceedings of the 5th Symposium on Usable Privacy and Security*, p. 4, ACM, 2009.

[18] 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*, pp. 261–272, ACM, 2013.

[19] M. Madden, "Privacy management on social media sites," tech. rep., 2012.

[20] A. Acquisti and R. Gross, "Imagined communities: Awareness, information sharing, and privacy on the facebook," in *International workshop on privacy enhancing technologies*, pp. 36–58, Springer, 2006.

[21] 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*, p. 12, ACM, 2013.

[22] M. Netter, S. Herbst, and G. Pernul, *Interdisciplinary Impact Analysis of Privacy in Social Networks*. Springer, 2013.

[23] F. Schaub, R. Balebako, A. L. Durity, and L. F. Cranor, "A design space for effective privacy notices," in *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*, pp. 1–17, USENIX Association, 2015.

[24] M. Duckham and L. Kulik, "A formal model of obfuscation and negotiation for location privacy," in *International Conference on Pervasive Computing*, pp. 152–170, Springer, 2005.

[25] J. Krumm, "Inference attacks on location tracks," in *International Conference on Pervasive Computing*, pp. 127–143, Springer, 2007.

[26] J. Lindqvist, J. Cranshaw, J. Wiese, J. Hong, and J. Zimmerman, "I'm the mayor of my house: examining why people use foursquare-a social-driven location sharing application," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '11*, pp. 2409–2418, 2011.

[27] B. Krishnamurthy and C. E. Wills, "Characterizing privacy in online social networks," in *Proceedings of the first workshop on Online social networks*, pp. 37–42, ACM, 2008.

[28] R. Kang, L. Dabbish, N. Fruchter, and S. Kiesler, "my data just goes everywhere:Ť user mental models of the internet and implications for privacy and security," in *Symposium on Usable Privacy and Security (SOUPS)*, pp. 39–52, USENIX Association Berkeley, CA, 2015.

[29] 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.

[30] 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.

[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*, pp. 787–796, ACM, 2015.

[32] 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*, pp. 46–50, 2012.

[33] 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*, pp. 1443–1450, ACM, 2012.

[34] 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*, p. 445, ACM, 2013.

[35] 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*, pp. 51–56, ACM, 2012.

[36] S. Patil, R. Schlegel, A. Kapadia, and A. J. Lee, "Reflection or action?: how patil2014reflectionfeedback and control affect location sharing decisions," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 101–110, ACM, 2014.

[37] 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*, pp. 2003–2012, ACM, 2009.

[38] 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.

[39] J. Krumm, "A survey of computational location privacy," *Personal and Ubiquitous Computing*, vol. 13, no. 6, pp. 391–399, 2009.

[40] F. Stutzman, R. Gross, and A. Acquisti, "Silent listeners: The evolution of privacy and disclosure on facebook," *Journal of privacy and confidentiality*, vol. 4, no. 2, p. 2, 2013.

[41] R. Gross and A. Acquisti, "Information revelation and privacy in online social networks," in *Proceedings of the 2005 ACM workshop on Privacy in the electronic society*, pp. 71–80, ACM, 2005.

[42] M. Balduzzi, C. Platzer, T. Holz, E. Kirda, D. Balzarotti, and C. Kruegel, "Abusing social networks for automated user profiling," in *Recent Advances in Intrusion Detection*, pp. 422–441, Springer, 2010.

[43] M. Madden and K. Zickuhr, "Pew Internet American Life Project," tech. rep., 2011.

[44] R. Shokri, G. Theodorakopoulos, J.-Y. Le Boudec, and J.-P. Hubaux, "Quantifying location privacy," in *Security and privacy (sp), 2011 ieee symposium on*, pp. 247–262, IEEE, 2011.

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

[46] L. Sweeney, "Achieving k-anonymity privacy protection using generalization and suppression," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 05, pp. 571–588, 2002.

[47] M. Gruteser and D. Grunwald, "Anonymous usage of location-based services through spatial and temporal cloaking," in *Proceedings of the 1st international conference on Mobile systems, applications and services*, pp. 31–42, ACM, 2003.

[48] R. Shokri, G. Theodorakopoulos, G. Danezis, J.-P. Hubaux, and J.-Y. Le Boudec, "Quantifying location privacy: the case of sporadic location exposure," in *International Symposium on Privacy Enhancing Technologies Symposium*, pp. 57–76, Springer, 2011.

[49] K. P. Puttaswamy and B. Y. Zhao, "Preserving privacy in location-based mobile social applications," in *Proceedings of the Eleventh Workshop on Mobile Computing Systems & Applications*, pp. 1–6, ACM, 2010.

[50] W. Wei, F. Xu, and Q. Li, "Mobishare: Flexible privacy-preserving location sharing in mobile online social networks," in *INFOCOM, 2012 Proceedings IEEE*, pp. 2616–2620, IEEE, 2012.

[51] X. Zhao, L. Li, and G. Xue, "Checking in without worries: Location privacy in location based social networks," in *INFOCOM, 2013 Proceedings IEEE*, pp. 3003–3011, IEEE, 2013.

[52] A. P. Felt, E. Ha, S. Egelman, A. Haney, E. Chin, and D. Wagner, "Android permissions: User attention, comprehension, and behavior," in *Proceedings of the Eighth Symposium on Usable Privacy and Security*, p. 3, ACM, 2012.

[53] M. L. Damiani, "Third party geolocation services in lbs: privacy requirements and research issues.," *Trans. Data Privacy*, vol. 4, no. 2, pp. 55–72, 2011.

[54] C. E. Wills and C. Tatar, "Understanding what they do with what they know," in *Proceedings of the 2012 ACM Workshop on Privacy in the Electronic Society*, pp. 13–18, ACM, 2012.

[55] B. Schneier, *Carry on: Sound advice from Schneier on security*. John Wiley & Sons, 2013.

[56] Y. Liu, K. P. Gummadi, B. Krishnamurthy, and A. Mislove, "Analyzing facebook privacy settings: user expectations vs. reality," in *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*, pp. 61–70, ACM, 2011.

[57] A. Brush, J. Krumm, and J. Scott, "Exploring end user preferences for location obfuscation, location-based services, and the value of location," in *Proceedings of thekelley2011users 12th ACM international conference on Ubiquitous computing - Ubicomp '10*, pp. 95–104, 2010.

[58] H. Fu, Y. Yang, N. Shingte, J. Lindqvist, and M. Gruteser, "A field study of run-time location access disclosures on android smartphones," *Proc. USEC*, vol. 14, 2014.

[59] B. Debatin, J. P. Lovejoy, A.-K. Horn, and B. N. Hughes, "Facebook and online privacy: Attitudes, behaviors, and unintended consequences," *Journal of Computer-Mediated Communication*, vol. 15, no. 1, pp. 83–108, 2009.

[60] S. Vihavainen, A. Lampinen, A. Oulasvirta, S. Silfverberg, and A. Lehmuskallio, "The clash between privacy and automation in social media," *IEEE Pervasive Computing*, vol. 13, no. 1, pp. 56–63, 2014.

[61] L. Palen and P. Dourish, "Unpacking privacy for a networked world," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 129–136, ACM, 2003.

[62] L. Barkhuus, B. Brown, M. Bell, S. Sherwood, M. Hall, and M. Chalmers, "From awareness to repartee: sharing location within social groups," in *proceedings of the SIGCHI conference on human factors in computing systems*, pp. 497–506, ACM, 2008.

[63] N. Poolsappasit and I. Ray, "Towards a scalable model for location privacy," in *Proceedings of the SIGSPATIAL ACM GIS 2008 International Workshop on Security and Privacy in GIS and LBS*, pp. 46–51, ACM, 2008.

[64] S. Patil and J. Lai, "Who gets to know what when: configuring privacy permissions in an awareness application," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 101–110, ACM, 2005.

[65] P. G. Kelley, M. Benisch, L. F. Cranor, and N. Sadeh, "When are users comfortable sharing locations with advertisers?," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2449–2452, ACM, 2011.

[66] P. Jagtap, A. Joshi, T. Finin, and L. Zavala, "Privacy preservation in context aware geo-social networking applications," *organization*, 2011.

[67] M. Li, H. Zhu, Z. Gao, S. Chen, L. Yu, S. Hu, and K. Ren, "All your location are belong to us: Breaking mobile social networks for automated user location tracking," in *Proceedings of the 15th ACM international symposium on Mobile ad hoc networking and computing*, pp. 43–52, ACM, 2014.

[68] T. Fechner and C. Kray, "Attacking location privacy: exploring human strategies," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pp. 95–98, ACM, 2012.

[69] L. Hickman, "How i became a foursquare cyberstalker," 2010.

[70] Z. Cheng, J. Caverlee, and K. Lee, "You are where you tweet: a content-based approach to geo-locating Twitter users," in *Proceedings of the 19th ACM international conference on Information and Knowledge Management CIKM '10*, pp. 759–768, 2010.

[71] 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*, pp. 898–905, 2012.

[72] 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*, pp. 723–732, 2012.

[73] R. Li, S. Wang, and K. C.-C. Chang, "Multiple location profiling for users and relationships from social network and content," *Proceedings of the VLDB Endowment*, vol. 5, no. 11, pp. 1603–1614, 2012.

[74] T. Qin, R. Xiao, L. Fang, X. Xie, and L. Zhang, "An efficient location extraction algorithm by leveraging web contextual information," in *proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems*, pp. 53–60, ACM, 2010.

[75] Z. Huo, X. Meng, and R. Zhang, "Feel free to check-in: Privacy alert against hidden location inference attacks in geosns," in *International Conference on Database Systems for Advanced Applications*, pp. 377–391, Springer, 2013.

[76] H. Gao, J. Tang, and H. Liu, "Exploring social-historical ties on location-based social networks," in *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media*, pp. 1140–121, 2012.

[77] 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*, pp. 1582–1586, 2012.

[78] D. J. Crandall, L. Backstrom, D. Cosley, S. Suri, D. Huttenlocher, and J. Kleinberg, "Inferring social ties from geographic coincidences," *Proceedings of the National Academy of Sciences*, vol. 107, no. 52, pp. 22436–22441, 2010.

[79] S. Scellato, A. Noulas, R. Lambiotte, and C. Mascolo, "Socio-spatial properties of online location-based social networks," in *ICWSM*, pp. 329–336, 2011.

[80] G. B. Colombo, M. J. Chorley, M. J. Williams, S. M. Allen, and R. M. Whitaker, "You are where you eat: Foursquare checkins as indicators of human mobility and behaviour," in *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2012 IEEE International Conference on*, pp. 217–222, IEEE, 2012.

[81] T. Kurashima, T. Iwata, G. Irie, and K. Fujimura, "Travel route recommendation using geotags in photo sharing sites," in *Proceedings of the 19th ACM international conference on Information and knowledge management*, pp. 579–588, ACM, 2010.

[82] T. Kurashima, T. Iwata, T. Hoshide, N. Takaya, and K. Fujimura, "Geo topic model: joint modeling of user's activity area and interests for location recommendation," in *Proceedings of the sixth ACM international conference on Web search and data mining*, pp. 375–384, ACM, 2013.

[83] L. B. Marinho, I. Nunes, T. Sandholm, C. Nóbrega, J. Araújo, and C. E. S. Pires, "Improving location recommendations with temporal pattern extraction," in *Proceedings of the 18th Brazilian symposium on Multimedia and the web*, pp. 293–296, ACM, 2012.

[84] 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*, pp. 295–304, ACM, 2015.

[85] M. J. Chorley, G. B. Colombo, S. M. Allen, and R. M. Whitaker, "Visiting patterns and personality of foursquare users," in *Cloud and Green Computing (CGC), 2013 Third International Conference on*, pp. 271–276, IEEE, 2013.

[86] H. Taylor, "Most people are privacy pragmatists who, while concerned about privacy, will sometimes trade it off for other benefits," *The Harris Poll*, vol. 17, no. 19, p. 44, 2003.

[87] Y. Wang and A. Kobsa, "Privacy-enhancing technologies," *Social and Organizational Liabilities in Information Security*, pp. 203–227, 2008.

[88] B. Henne, C. Szongott, and M. Smith, "Snapme if you can: privacy threats of other peoples' geo-tagged media and what we can do about it," in *Proceedings of the sixth ACM conference on Security and privacy in wireless and mobile networks*, pp. 95–106, ACM, 2013.

[89] L. Barkhuus and A. K. Dey, "Location-based services for mobile telephony: a study of users' privacy concerns.," in *Interact*, vol. 3, pp. 702–712, 2003.

[90] S. Lederer, J. Mankoff, and A. K. Dey, "Who wants to know what when? privacy preference determinants in ubiquitous computing," in *CHI'03 extended abstracts on Human factors in computing systems*, pp. 724–725, ACM, 2003.

[91] S. Consolvo, I. E. Smith, T. Matthews, A. LaMarca, J. Tabert, and P. Powledge, "Location disclosure to social relations: why, when, & what people want to share," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 81–90, ACM, 2005.

[92] 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*, pp. 207–216, ACM, 2011.

[93] P. A. Norberg, D. R. Horne, and D. A. Horne, "The privacy paradox: Personal information disclosure intentions versus behaviors," *Journal of Consumer Affairs*, vol. 41, no. 1, pp. 100–126, 2007.

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

[95] 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 ECSCWâĂŹ93*, pp. 77–92, Springer, 1993.

[96] 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, pp. 214–221, 1999.

[97] M. Langheinrich, "Privacy by design: principles of privacy-aware ubiquitous systems," in *Ubicomp 2001: Ubiquitous Computing*, pp. 273–291, Springer, 2001.

[98] 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.

[99] R. Shokri, J. Freudiger, and J.-P. Hubaux, "A unified framework for location privacy," tech. rep., 2010.

[100] A. M. Zafeiropoulou, K. O'Hara, D. Millard, and C. Webber, "Location data and privacy: a framework for analysis," 2012.

[101] R. De Paula, X. Ding, P. Dourish, K. Nies, B. Pillet, D. Redmiles, J. Ren, J. Rode, *et al.*, "Two experiences designing for effective security," in *Proceedings of the 2005 symposium on Usable privacy and security*, pp. 25–34, ACM, 2005.

[102] L. Emanuel, C. Bevan, and D. Hodges, "What does your profile really say about you?: privacy warning systems and self-disclosure in online social network spaces," in *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, pp. 799–804, ACM, 2013.

[103] K. S. Bordens and B. B. Abbott, *Research design and methods: A process approach*. McGraw-Hill, 2002.

[104] Y. Wang, P. G. Leon, X. Chen, and S. Komanduri, "From facebook regrets to facebook privacy nudges," *Ohio St. LJ*, vol. 74, p. 1307, 2013.

[105] A. Acquisti, "Nudging privacy: The behavioral economics of personal information," *IEEE security & privacy*, vol. 7, no. 6, 2009.

[106] M. S. Bargh, R. Meijer, S. Choenni, and P. Conradie, "Privacy protection in data sharing: towards feedback based solutions," in *Proceedings of the 8th International Conference on Theory and Practice of Electronic Governance*, pp. 28–36, ACM, 2014.

[107] L. Jedrzejczyk, B. A. Price, A. K. Bandara, and B. Nuseibeh, "On the impact of real-time feedback on users' behaviour in mobile location-sharing applications," in *Proceedings of the Sixth Symposium on Usable Privacy and Security*, p. 14, ACM, 2010.

[108] 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*, pp. 140–150, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2014.

[109] P. Shi, H. Xu, and X. L. Zhang, "Informing security indicator design in web browsers," in *Proceedings of the 2011 iConference*, pp. 569–575, ACM, 2011.

[110] C. Bravo-Lillo, S. Komanduri, L. F. Cranor, R. W. Reeder, M. Sleeper, J. Downs, and S. Schechter, "Your attention please: designing security-decision uis to make genuine risks harder to ignore," in *Proceedings of the Ninth Symposium on Usable Privacy and Security*, p. 6, ACM, 2013.

[111] 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*, p. 2, ACM, 2011.

[112] T. Hughes-Roberts and E. Kani-Zabihi, "On-line privacy behavior: Using user interfaces for salient factors," *Journal of Computer and Communications*, vol. 2, no. 04, p. 220, 2014.

[113] C. Bravo-Lillo, L. F. Cranor, J. Downs, and S. Komanduri, "Bridging the gap in computer security warnings: A mental model approach," *IEEE Security & Privacy*, no. 2, pp. 18–26, 2010.

[114] N. Valkanova, S. Jorda, M. Tomitsch, and A. Vande Moere, "Reveal-it!: the impact of a social visualization projection on public awareness and discourse," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 3461–3470, ACM, 2013.

[115] S. E. Middleton, N. R. Shadbolt, and D. C. De Roure, "Capturing interest through inference and visualization: Ontological user profiling in recommender systems," in *Proceedings of the 2nd international conference on Knowledge capture*, pp. 62–69, ACM, 2003.

[116] J. Heer and D. Boyd, "Vizster: Visualizing online social networks," in *Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on*, pp. 32–39, IEEE, 2005.

[117] D. Tchuente, M.-F. Canut, N. B. Jessel, A. Péninou, and A. El Haddadi, "Visualizing the evolution of users' profiles from online social networks," in *Advances in Social Networks Analysis and Mining (ASONAM), 2010 International Conference on*, pp. 370–374, IEEE, 2010.

[118] T. Plumbaum, K. Schulz, M. Kurze, and S. Albayrak, "My personal user interface: A semantic user-centric approach to manage and share user information," *Human Interface and the Management of Information. Interacting with Information*, pp. 585–593, 2011.

[119] K. Church, J. Neumann, M. Cherubini, and N. Oliver, "The map trap?: an evaluation of map versus text-based interfaces for location-based mobile search services," in *Pro-*

ceedings of the 19th international conference on World wide web, pp. 261–270, ACM, 2010.

[120] A. Cuttone, S. Lehmann, and J. E. Larsen, "A mobile personal informatics system with interactive visualizations of mobility and social interactions," in *Proceedings of the 1st ACM international workshop on Personal data meets distributed multimedia*, pp. 27–30, ACM, 2013.

[121] J. Rode, C. Johansson, P. DiGioia, K. Nies, D. H. Nguyen, J. Ren, P. Dourish, D. Redmiles, *et al.*, "Seeing further: extending visualization as a basis for usable security," in *Proceedings of the second symposium on Usable privacy and security*, pp. 145–155, ACM, 2006.

[122] Y. Wang, L. Gou, A. Xu, M. X. Zhou, H. Yang, and H. Badenes, "Veilme: An interactive visualization tool for privacy configuration of using personality traits," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 817–826, ACM, 2015.

[123] E. Kani-Zabihi and M. Helmhout, "Increasing service usersŠ privacy awareness by introducing on-line interactive privacy features," in *Nordic Conference on Secure IT Systems*, pp. 131–148, Springer, 2011.

[124] J. Angulo, S. Fischer-Hübner, T. Pulls, and E. Wästlund, "Usable transparency with the data track: a tool for visualizing data disclosures," in *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, pp. 1803–1808, ACM, 2015.

[125] A. Popescu, G. Grefenstette, and P.-A. Moëllic, "Gazetiki: automatic creation of a geographical gazetteer," in *ACM/IEEE Joint Conference on Digital Libraries*, pp. 85–93, ACM, 2008.

[126] 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.

[127] C. Roda, "Human attention and its implications for human–computer interaction," *Human Attention in Digital Environments*, p. 11, 2011.

[128] M. I. Posner, C. R. Snyder, and B. J. Davidson, "Attention and the detection of signals.," *Journal of experimental psychology: General*, vol. 109, no. 2, p. 160, 1980.

[129] L. K. Ainsworth and B. Kirwan, *A guide to task analysis*. Taylor & Francis, 1992.

[130] B. Kirwan and L. K. Ainsworth, *A guide to task analysis: the task analysis working group*. CRC press, 1992.

[131] J. Annett and N. A. Stanton, *Task analysis*. CRC Press, 2000.

[132] J. M. Schraagen, S. F. Chipman, and V. L. Shalin, *Cognitive task analysis*. Psychology Press, 2000.

[133] J. T. Hackos and J. Redish, "User and task analysis for interface design," 1998.

[134] N. Bevan, "International standards for hci and usability," *International journal of human-computer studies*, vol. 55, no. 4, pp. 533–552, 2001.

[135] V. Hinze-Hoare, "Four principles fundamental to design practice for human centred systems," *arXiv preprint cs/0409041*, 2004.

[136] K. Krol, J. M. Spring, S. Parkin, and M. A. Sasse, "Towards robust experimental design for user studies in security and privacy," 2016.

[137] 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.

[138] J. Tidwell, *Designing interfaces: Patterns for effective interaction design*. " O'Reilly Media, Inc.", 2010.

[139] J. Nielsen, "Usability inspection methods," in *Conference companion on Human factors in computing systems*, pp. 413–414, ACM, 1994.

[140] J. Nielsen and R. Molich, "Heuristic evaluation of user interfaces," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 249–256, ACM, 1990.

[141] C. Wharton, J. Rieman, C. Lewis, and P. Polson, "The cognitive walkthrough method: A practitioner's guide," in *Usability inspection methods*, pp. 105–140, John Wiley & Sons, Inc., 1994.

[142] D. L. Morgan, "Focus groups," *Annual review of sociology*, pp. 129–152, 1996.

[143] J. Kitzinger, "Qualitative research. introducing focus groups.," *BMJ: British medical journal*, vol. 311, no. 7000, p. 299, 1995.

[144] J. Cohen, "Statistical power analysis for the behavioural sciences. hillside," *NJ: Lawrence Earlbaum Associates*, 1988.