



AALBORG UNIVERSITY
DENMARK

Aalborg Universitet

Disclosure of Personal Data in Ubiquitous Social Networking

Sapuppo, Antonio

Publication date:
2013

Document Version
Accepted author manuscript, peer reviewed version

[Link to publication from Aalborg University](#)

Citation for published version (APA):
Sapuppo, A. (2013). *Disclosure of Personal Data in Ubiquitous Social Networking* (1 ed.). Department of Electronic Systems, Aalborg University.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.



Disclosure of Personal Data in Ubiquitous Social Networking

Antonio Sapuppo

Department of Planning and Development

Aalborg University

A thesis submitted for the degree of

Doctor of Philosophy

10 September 2013

To Salvatore and Rosalia

Abstract

Ubiquitous social networking focuses on developing possible advantageous relationships such as friendships, partnerships and business relations in the physical world, by uncovering hidden connections that people share with others nearby. The foundation of these services is based on disclosure of personal information, which can provoke numerous accidental invasions of privacy. This dissertation contributes by addressing two problems, related to support of privacy-aware social networking in ubiquitous computing environments that focus on maximizing potential networking benefits while preserving users' privacy. Firstly, it updates the current privacy guidelines of ubiquitous computing by proposing four drawbacks to be avoided when designing for privacy in ubiquitous social networking environments. Secondly, this dissertation identifies and investigates the determinants that might influence the variation of human data sensitivity under different circumstances, for ensuring accuracy of selective disclosure of personal data. By taking into account the proposed design guidelines and identified influential factors, this dissertation seeks to provide relevant insights into data disclosure for contributing to development of privacy-aware social networking in ubiquitous computing environments.

Resumé

Ubiquitous social networking fokuserer på at skabe nye, sociale relationer mellem folk - venskaber, partnerskaber og endda forretningsforbindelser - i den fysiske verden ved at afdække de skjulte forbindelser og fællestræk blandt folk i det fysiske nærmiljø. Grundlaget for sådanne services er videregivelse af personfølsomme oplysninger, hvilket kan resultere i utilsigtet indtrængen i privatlivet. Denne afhandling berører to problemer ved sikring af persondata i sociale netværk i forbindelse med ubiquitous computing environments, hvor fokus er maksimal social networking med minimal risiko for privatlivet. Først og fremmest opdateres de gældende retningslinjer for ubiquitous computing ved at præsentere fire problemer, der bør håndteres ved design af ubiquitous social networking med personfølsomme oplysninger. Dernæst identificerer og undersøger nærværende afhandling de væsentlige faktorer, der kunne påvirke variationen af sensitiviteten af personfølsomme oplysninger under forskellige omstændigheder, til at sikre en høj præcision i selektiv videregivelse af personfølsomme oplysninger. Ved at tage disse opdaterede retningslinjer og væsentlige faktorer i betragtning søger denne afhandling at bidrage til udviklingen af social networking med sikring af personfølsomme data for øje.

Acknowledgements

I would like to express my gratitude to the Aalborg University for giving me this great opportunity to study for a PhD degree and the Auckland University of Technology for hosting me during my internship in New Zealand. This PhD was also supported by Nokia as a part of the Converged Advanced Media Platforms (CAMMP) project and Otto Mønstedts Fond that partially financed my trips for international conferences and internship in different parts of the world.

During this period, I met and worked with many brilliant researchers with whom I have shared great moments as well as very busy days. I would like to thank Birger Andersen, who was the first person that believed in the topic of this thesis, provided me the first opportunity to research in Copenhagen and suggested me to start the PhD. I am also very grateful to Knud Erik Skouby, without his help and support this PhD would not have taken place.

I owe many thanks to my supervisors Lene Tolstrup Sørensen and Reza Tadayoni for their significant comments on the PhD thesis and publications as well as to Emil Heinze, with whom I have enjoyed many brainstorming meetings and who always offered his help in numerous

circumstances during this PhD. Also, I would like to acknowledge all the CMI members for allowing me to work in a friendly environment, as well as Patrick Gyaase, Benjamin Kwofie, Kosta Pandazos and the anonymous reviewers for providing useful comments on my PhD thesis.

There is also another brilliant researcher Boon-Chong Seet, to whom I express gratitude for supervising me during my stay in New Zealand. I will always be grateful to Boon-Chong for advising and guiding me into the field of ubiquitous computing, and providing me useful comments on my publications and PhD thesis as well as co-authoring one of the key papers of this dissertation.

I am very thankful to Gian Paolo Perrucci for always being a source of innovative ideas and for his useful comments on the thesis, and to João Figueiras, with whom I had the pleasure to co-author one of the key papers of this dissertation. Many thanks also to Stavris Solo who always offered his help with recruiting participants for my investigations. I also would like to take the opportunity to acknowledge all my international friends, who become part of my family here in Denmark and supported me all the time during this long trip.

Last, but not least, I would like to thank my family in Sicily for making this study in Denmark possible and Egle Juzokaite. Probably there are no words to express my immense gratitude to Egle for encouraging, supporting, inspiring, advising and helping me all the time during these years. Without her, this work would not have been possible.

List of Publications

The thesis is based on the following five journal papers:

PAPER A Privacy and Technology Challenges for Ubiquitous Social Networking. Sapuppo, A. and Seet B.-C. *Accepted for publication in International Journal of Ad hoc and Ubiquitous Computing.*

PAPER B Ubiquitous Social Networking: Concept and Evaluation. Sapuppo, A. *Sensor Letters*, Vol. 10, No. 8, Pages 1632-1644, 2012.

PAPER C Designing for Privacy in Ubiquitous Social Networking. Sapuppo, A and Figueiras, J. *Accepted for publication in International Journal of Ad hoc and Ubiquitous Computing.*

PAPER D Privacy Analysis in Mobile Social Networks: the Influential Factors for Disclosure of Personal Information. Sapuppo, A. *International Journal of Wireless and Mobile Computing*, Vol. 5, No. 4, Pages 315-326, 2012.

PAPER E The Influential Factors for the Variation of Data Sensitivity in Ubiquitous Social Networking. Sapuppo, A. *International Journal of Wireless and Mobile Computing*, Vol. 6, No. 2, Pages 115-130, 2013.

The following publications have also been carried out by the author of this thesis during his Ph.D. studies.

Journal Papers

- An Empirical Investigation of Disclosure of Personal Information in Ubiquitous Social Computing. Sapuppo, A. and Seet B.C. *International Journal of Computer Theory and Engineering*, Vol.4, No. 3, pp. 373-378, 2012
- User Profiles, Personalization and Privacy: WWRF Outlook. Olesen, H., Noll, J., Hofmann, M., Hammershøj, A., Sapuppo, A., Iqbal, Z., Elahi, N., Chowdhury, M., Heikkinen, S., Sutterer, M., Hornung, G., Schnabel, C., Hornung, G., Drögehorn, O., Thuveson, H. and Sorge, C. *WWRF Outlook Series*, No. 3, 2009.

Book Chapters

- Social Ambient Intelligence for Assisted Living: Social Context, Networking and Privacy. Sapuppo, A. and Seet, B.-C. *Ambient Assisted Living*, Taylor and Francis, 2013.

Conference Papers

- Local Social Networks. Sapuppo, A. and Sørensen, L.T. *In Computer Communication and Management: International Proceedings of Computer Science and Information Technology*, Vol. 5, pp.15-22, IACSIT Press, 2011.
- Challenges for Mobile Application Development. Hammershøi, A., Sapuppo, A. and Tadayoni, R. *In 14th International Conference on Intelligence in Next Generation Networks (ICIN): "Weaving Applications into the Network Fabric"*, IEEE, 2010.
- Spiderweb : A Social Mobile Network. Sapuppo, A. *In: Wireless Conference (EW), 2010 European*, IEEE, pp.475-481, 2010.
- Mobile Platforms: An analysis of Mobile Operating Systems and Software development platforms. Hammershøi, A., Sapuppo, A.

and Tadayoni, R. *In CMI International Conference on Social Net-
working and Communities*, 2009.

Gli uomini passano, le idee restano. Restano le loro tensioni morali e continueranno a camminare sulle gambe di altri uomini.

Giovanni Falcone

Contents

| | |
|--|------------|
| Abstract | iv |
| Resumè | v |
| Acknowledgements | vii |
| List of Publications | xi |
| 1 Introduction | 1 |
| 1.1 Privacy in ubiquitous computing | 8 |
| 1.2 Problem formulation | 13 |
| 1.2.1 Designing for personal privacy in ubiquitous computing | 15 |
| 1.2.2 The variation of human data sensitivity | 19 |
| 1.3 Research design and approach | 24 |
| 1.3.1 The selection of participants | 27 |
| 1.3.2 Research process | 29 |
| 1.3.3 Investigation methodology | 35 |
| 1.4 Research contributions | 43 |

CONTENTS

| | | |
|----------|---|------------|
| 1.5 | Dissertation outline | 45 |
| 2 | Summary of the papers | 47 |
| 2.1 | Paper A | 48 |
| 2.2 | Paper B | 50 |
| 2.3 | Paper C | 52 |
| 2.4 | Paper D | 54 |
| 2.5 | Paper E | 56 |
| 3 | Results | 59 |
| 3.1 | Privacy design guidelines | 62 |
| 3.2 | The influential factors for data disclosure | 78 |
| 3.3 | Investigation limitations | 84 |
| 4 | Conclusions | 87 |
| | References | 93 |
| A | Privacy and technology challenges for ubiquitous social network- | |
| | ing | 119 |
| A.1 | Introduction | 121 |
| A.2 | Context acquisition | 126 |
| A.2.1 | Users' identities and type of relationships | 128 |
| A.2.2 | Users' activities | 131 |
| A.2.3 | Users' online profiles | 133 |
| A.3 | Software architectures | 134 |

| | | |
|----------|---|------------|
| A.3.1 | Centralized architecture | 135 |
| A.3.2 | Decentralized architecture | 137 |
| A.3.3 | Hybrid architecture | 141 |
| A.4 | Privacy design guidelines | 143 |
| A.4.1 | Data protection | 144 |
| A.4.2 | Personal privacy | 149 |
| A.5 | Privacy management models | 153 |
| A.5.1 | Predefined privacy preferences | 154 |
| A.5.2 | Ad hoc privacy control | 156 |
| A.6 | Discussion | 161 |
| B | Ubiquitous social networking: concept and evaluation | 167 |
| B.1 | Introduction | 169 |
| B.2 | Local social networks | 171 |
| B.2.1 | Example scenario | 171 |
| B.2.2 | Definition | 174 |
| B.2.3 | Preliminary architecture | 174 |
| B.2.4 | The Spiderweb prototype | 178 |
| B.2.5 | Application areas | 180 |
| B.3 | Investigation methodology and design | 181 |
| B.3.1 | Background | 182 |
| B.3.2 | Participants | 184 |
| B.4 | Investigation results | 185 |
| B.4.1 | Perceived usefulness | 186 |

CONTENTS

| | | |
|----------|--|------------|
| B.4.2 | Acceptance of prerequisites | 188 |
| B.5 | Conclusions | 201 |
| C | Designing for privacy in ubiquitous social networking | 205 |
| C.1 | Introduction | 207 |
| C.2 | Privacy design guidelines | 210 |
| C.3 | Ubiquitous social networking designs | 214 |
| C.3.1 | Negative case studies for the privacy pitfalls | 214 |
| C.3.2 | Positive case study for the privacy pitfalls | 219 |
| C.3.3 | Additional privacy limitations | 221 |
| C.4 | Privacy drawbacks to be avoided by designers | 222 |
| C.4.1 | Negative case studies for the privacy drawbacks | 225 |
| C.4.2 | Interdependencies among the four drawbacks | 226 |
| C.5 | Privacy-aware platform design | 228 |
| C.5.1 | User profile management | 229 |
| C.5.2 | Communication flow | 232 |
| C.5.3 | Analysis of privacy guidelines | 234 |
| C.6 | Evaluation | 244 |
| C.6.1 | Background | 244 |
| C.6.2 | Participants | 246 |
| C.6.3 | Results | 248 |
| C.6.4 | Investigation limitations | 255 |
| C.7 | Conclusions | 256 |

| | |
|--|------------|
| D Privacy analysis in mobile social networks: the influential factors for disclosure of personal data | 259 |
| D.1 Introduction | 261 |
| D.2 Human data disclosure | 264 |
| D.3 Design of the surveys | 267 |
| D.3.1 Survey I | 268 |
| D.3.2 Survey II | 269 |
| D.4 Participants of the surveys | 272 |
| D.4.1 Respondents of the first survey | 273 |
| D.4.2 Respondents of the second survey | 274 |
| D.5 Survey results and discussion | 275 |
| D.5.1 Results of the survey I | 275 |
| D.5.2 Results of the survey II | 283 |
| D.6 Conclusions | 289 |
| E The influential factors for the variation of data sensitivity in ubiquitous social networking | 293 |
| E.1 Introduction | 295 |
| E.2 Related Work | 298 |
| E.3 The influential factors | 301 |
| E.4 Investigation methodology and design | 305 |
| E.4.1 Phase 1: Quantitative investigation | 306 |
| E.4.2 Phase 2: Qualitative investigation | 313 |
| E.5 Participants | 316 |

CONTENTS

| | |
|--------------------------------------|------------|
| E.6 Investigation results | 318 |
| E.6.1 Quantitative results | 318 |
| E.6.2 Qualitative results | 325 |
| E.7 Discussion | 332 |
| F Legal Documentation | 337 |

1

Introduction

"The first wave of computing, from 1940 to 1980, was dominated by many people serving one computer. The second wave, still peaking, has one person and one computer in uneasy symbiosis, staring at each other across the desktop without really inhabiting each other's worlds. The third wave, just beginning, has many computers serving each person everywhere in the world"

Marc Weiser

Marc Weiser's vision of ubiquitous computing was developed more than 20 years ago and it is still under research. This vision emphasized two key aspects for further development of computing: computers must be both invisible and calm. Weiser introduced the notion of invisible computing by arguing that computers should not be the central focus of users' attention, but instead users should focus on their tasks, rather than tools. The second aspect regards ensuring calmness of such

1. INTRODUCTION

technologies. Weiser discussed that computers should not cause stress or be either distractive or intrusive, but enhance people's lives and make their tasks easier. To better explain these aspects, Weiser referred to eyeglasses as an example of a good tool that enables users to concentrate on their tasks, such as reading a book and not on the tool, i.e. eyeglasses, which becomes invisible to them. Further, such tool as eyeglasses is not intrusive, as users do not feel distracted or slowed down on achieving their goals when utilizing it [192, 193].

In order to ensure invisible and calm technologies, the latest wave of computing - referred to as ubicomp - embeds many seamless highly specialized devices, interconnected between each other, within people's surroundings. These devices are aware of their current environments and users and, consequently, they are able to improve humans' lives and support their everyday tasks [191]. Numerous application areas of ubicomp range from education [34, 52] where academic service learning can be facilitated [7, 53, 110, 203], to other environments, such as museums [23, 35, 86], exhibition visits [36, 181] or smart home infrastructures [70, 117]. Moreover, a majority of ubicomp research focused on the healthcare application domain [47, 162, 174], general entertainment, e.g. gaming [17, 38, 67], shopping assistance [78], tourism [40, 156, 175] and sport technologies [11, 39, 137, 200].

Among this large variety of ubicomp application areas, enhancing human communication in the physical world has been receiving increasing attention during the last years. Such relevant interest is a natural consequence of the tremendous success of online social networks sites that have quickly improved the communication between people by enabling their users to stay in touch with friends from the whole world, share pictures, talk, chat, send messages, look for new acquaintances and

new job opportunities [48, 206]. Nowadays there are numerous online social networks, which serve users with various interests and goals, e.g. professional, music. Such variety of possibilities attracted the attention of millions of people, many of whom have integrated these networks into their daily practice [18]. However, these online services present serious limitations in regard to the enhancement of social networking between users in the physical surroundings. As a result, people with similar interests and professional goals fail to leverage interpersonal affinities with others if they do not have an established connection in one of the available online sites [171].

The great popularity of online social networks has inspired ubicomp researchers and practitioners to investigate possibilities for improving human communication by enhancing social networking and transferring online social networks benefits to the physical world [65, 153, 183]. To achieve this goal, it is essential for ubicomp to embody social intelligence in order to intelligently and naturally support human communication in the physical world. Social intelligence can be defined as the ability of the environment to acquire and apply users' social context in order to foster social interactions among its inhabitants [61, 166, 180, 184]. This can be considered as an evolution of ubicomp, where a social dimension has been introduced to respond to the social nature of the users and increase awareness, knowledge and intelligence of these environments.

This extension of ubicomp can be defined as ubiquitous social computing and the networking services established in such environments as Ubiquitous Social Networking (USN). The term "networking" is preferred instead of "network", as in the case of online social network sites, because this dissertation follows Boyd and

1. INTRODUCTION

Ellison's [18] views about the nature of *networking*. The authors claimed that networking emphasizes relationship initiation, often between strangers. While *networking* is possible in online social networks, it is not their primary goal. In fact, these online sites target at supporting communication within users' existing social networks. On the contrary, the main target of USN is *networking*, as it focuses on developing possible advantageous relationships such as friendships, partnerships and business relations by uncovering hidden connections that people share with others nearby. Particularly, USN facilitates initialization of face-to-face interactions between strangers with similar interests, i.e. *people who do not know each other, but probably should*. As a result, the value of social networking is significantly enhanced and benefits are available immediately upon demand [65, 171]. Potential application areas of USN are numerous and they range from professional, where these services might lead to new opportunities such as connecting employers with potential employees, to big events, such as conferences, company events and exhibitions that usually comprise large amounts of participants who potentially share similar professional or social interests [65, 168].

In order to better explain the concept of USN services, a scenario is presented in Figure 1.1. A user named Bob, who is marked in blue, is located in a public place, such as a canteen of an ordinary work place. Bob is surrounded by people whom he knows, marked in green, and people who are strangers to him, marked in white. Even if Bob does not interact with all people in the canteen, his ubiquitous devices do that for him by exchanging personal information with other people in his proximity, as shown in Figure 1.2-A. Due to automatic exchange of users' personal information, this process does not interfere with the current users' activities, and

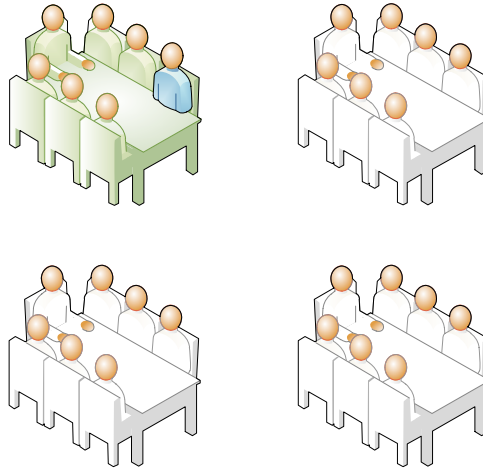


Figure 1.1: Ubiquitous social networking example scenario

it allows USN to develop an understanding about who the people nearby are, as well as their respective preferences. As shown in Figure 1.2-B, these services are capable of identifying users with similar interests, and thus highlighting relevant

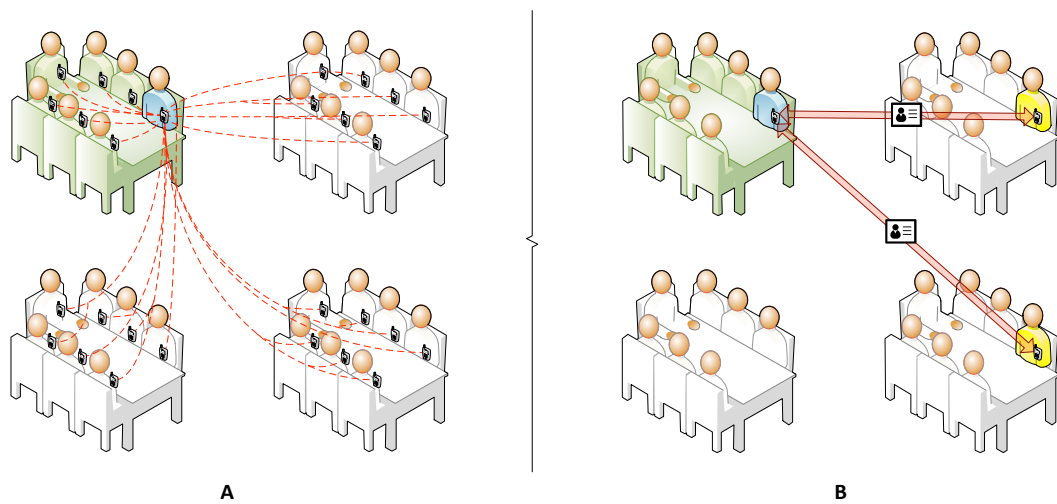


Figure 1.2: Automated personal information exchange and user profile matching in ubiquitous social networking

1. INTRODUCTION

social paths between users that would remain hidden otherwise. When USN services find profile similarities between Bob and other users, who are highlighted in yellow in Figure 1.2-B, users are notified about each others' presence and, therefore, have the opportunity to immediately initiate a face-to-face communication.

To translate the USN vision into reality, there is an indisputable need for social sensors, as sensing and inferring of the relevant environmental and social information of the users are crucial for reaching the primary goal to foster social interactions. Sensor networks have been deployed in ubicomp environments for obtaining context information. However, these sensors cannot be considered optimal due to their limited resources, such as restricted energy, networking and storage capabilities and they are also incapable of gathering users' social information, such as users' preferences [93, 129]. To overcome these limitations, latest research identifies smartphones as potential social sensors [64]. These devices are already equipped with GPS, accelerometer and other sensing components, enabling acquisition of environmental context information [21, 143]. They are also ideally suited to provide an insight into social behavior patterns, because they can be seen as wearable and inconspicuous sensors, constantly carrying users' personal information [21]. Finally, smartphones are also considered to be gateways to social networks, which are significant sources of users' personal information such as interests, preferences and list of friends [21, 64].

The vast amount of relevant context information, acquired by such social sensors, enables more personalized USN services, where inhabitants of USN environments are the primary targets of sensing. These services can be considered to be based on people-centric sensing approaches, because they focus mainly on social components,

such as user preferences, identities and relationships between the users, rather than just machines and devices [32, 33, 73]. This collection and further dissemination of users' personal information provides a crucial foundation for USN services to better empower people in their social conduct and enhance their sense of social connectedness, which in turn will play an important role in their physical and cognitive well-being [43, 164, 173].

On the other hand, collection and processing of users' personal information might lead to numerous potential invasions of users' privacy, which can be both accidental and intentional. For instance, the former might regard employers receiving employees' sensitive personal information that they did not even ask for, but suddenly found it on their "desks" [120]. Further, intentional intrusions could include violations of privacy by curious marketers or overprotective parents. Others might also use the collected information for stalking, bullying, intentional damage of reputation or even breach of civil rights by major corporations or governments [98]. Due to these numerous risks, potential users and researchers have already raised serious worries about the adoption of ubiquitous technologies and expressed need for better protection of users' privacy [2, 9, 89, 109, 194]. If not addressed responsibly, these concerns might discourage people from disclosing their personal information and thus threaten the further development of ubiquitous social computing technologies.

1.1 Privacy in ubiquitous computing

Even when ubicomp was just a vision, privacy threats were already considered as the greatest barriers to its long-term success [191]. As earlier mentioned, sensors are nowadays capable of acquiring not only environmental data, but also obtaining users' personal information. The technological development is moving towards people-centric era, where humans are the main focus of sensing. In people-centric sensing, users are parts of mobile sensor networks, where mobile devices are conceptually tied to individuals. Smartphones, always carried by the users, are capable of acquiring not only environmental data, but obtaining users' personal information as well, thanks to their sensing components. Therefore, these devices are considered to be the key elements for the development of ubiquitous social computing, because they are ideally suited to provide insight into social behaviour patterns [139]. However, the introduction of these social sensors in ubicomp leads to new crucial challenges for ensuring users' privacy, which include safe collecting, storing, processing and dissemination of users' personal information [106].

When addressing privacy in ubicomp, the majority of previous studies were inspired by Brandeis, who defined privacy as "*The right to be let alone*" and argued that individuals should always be able to protect their confidential data from undesired visibility [19]. These studies aimed at ensuring secrecy of users' personal information by applying for instance cryptography techniques, such as [16, 77, 112]. Other research works, instead, suggested to secure users' personal information by introducing different anonymity strategies - e.g. [14, 92, 133, 140] - that enabled users to be unidentifiable within a set of subjects, due to removal of connections between

the data owner and information [120]. Such solutions, which applied anonymity techniques, were inspired by the legal definition of privacy, interpreted by Smith as "*A condition of limited access to identifiable information about individuals*" [177].

The privacy definitions, provided by Brandeis and Smith, as well as the introduced solutions, i.e. cryptography and anonymity, refer to the first aspect within the field of privacy in ubicomp, i.e. the concept of *data protection* as reviewed by Iachello and Hong [100]. This concept, also known as informational self-determination, specifically regards the protection of users' confidential personal information from unwanted publicity. It is typically related to the management of identifiable personal information by third party entities and consists of rules that regulate how, when and for what reasons data can be collected, processed and disseminated [196, 197]. Data protection closely intertwines privacy with security, as it deals with many aspects commonly discussed in security research works, such as defending users' personal information from being revealed to, or modified by, an unauthorized entity or individual [179].

In order to control such personal information management, the fair information practices were created to guide data protection policies worldwide [100, 120]. The fair information practices are one of the earliest guidelines that indisputably influenced all subsequent data protection legislations, by including notions, such as collection and use limitation, openness and transparency, individual participation, accountability and reasonable security. Later on, these guidelines were also adapted to ubicomp by Langheinrich [120], who identified several main areas of innovation and system design for data protection in such environments: notifying the user appropriately; taking into account user's choice and seeking for consent;

1. INTRODUCTION

enforcing limitation of scope within the concepts of proximity and locality; enabling anonymity and pseudonymity when necessary; providing adequate security and appropriate data access.

The protection of data against various intentional threats, such as access by unauthorized users, represents only a part of the issues arising within the field of privacy in ubicomp. There are many other situations when users want to disclose their personal information to other people, in order to gain benefits in exchange to their information sharing [98, 168]. For instance, many people decide to disclose their identifiable personal information, commonly considered as sensitive (e.g. full name, home address, phone number), in exchange to shopping loyalty cards. They consider that sharing this personal information is an acceptable trade-off between their privacy and benefits, e.g. receiving discounts on items bought outweigh potential privacy risks [2]. Furthermore, users might choose to share their current GPS positions in order to better coordinate arrivals to an event or to intentionally reveal presence to co-workers or friends [98]. Finally, other examples regard USN, where users decide to disclose their personal information, such as favourite music and movies or career abilities and expectations, to others for receiving new potential professional and personal networking benefits in exchange to their data disclosure [168].

The illustrated cases introduce the second aspect of privacy in ubicomp, as reviewed by Iachello and Hong [100], which emphasizes that the main challenge for privacy management systems of ubicomp is shifting from hiding personal data to ensuring successful management of disclosure of users' personal information. This leads to Westin's interpretation of privacy: "*The right to select what personal*

information about me is known to what people” [196]. Hong [98] supported Westin’s interpretation of privacy by introducing the concept of *personal privacy* that focuses on empowering the users to share the right information, with the right people under the right circumstances. Specifically, the author defined personal privacy as follows: *”The processes by which people selectively share personal information, such as email address, career skills and abilities, to organizations and to other people”*.

In contrast to data protection, personal privacy cannot be seen as a static notion where users set rules and enforce them, but rather it should be considered as a dynamic process representing continuous negotiation and management of the boundaries that shape personal data disclosure [4, 5, 100, 149]. Westin [196] discussed how people continuously encounter internal conflicts in finding optimal trade-offs between the desire for privacy and the one for disclosure and consequent communication. Moreover, Darrah et al [55] noticed that people tend to utilize different strategies for providing as little personal information to others as possible, while simultaneously seeking to maximize their communication opportunities. Goffman [79, 80] also provided insights into the management of disclosure of personal information by observing that people present different personalities (or personas) under different circumstances. For example, an individual might project a professional personality to colleagues at work, while being more informal with others in social environments. These observations are also supported by several investigations on preferences of personal data disclosure under different circumstances in ubiquitous social computing environments [45, 57, 108, 123, 148, 170, 199]. The results of these investigations proved that people’s data disclosure to others is highly situational, as individuals preferred sharing different sets of personal information under different

1. INTRODUCTION

circumstances.

In conclusion, it must be noted that when reviewing the two introduced aspects of privacy in the field of ubicomp, Iachello and Hong [100] highlighted that data protection and personal privacy present relevant differences, which would directly affect the design of ubicomp applications. The authors discussed that managing users' data privacy according to data protection guidelines would lead to setting rigid rules and policies to prevent potential malicious attacks. For applications that focus on leveraging interpersonal affinities in physical environments, rules that regulate users' data disclosure under different circumstances are not straightforward and they cannot be algorithmically modeled using rigid privacy policies. Instead, the authors suggested that the design of these environments should enable a constant and arbitrary selection of disclosed data for allowing users to easily present a desired image of themselves under different situations.

This dissertation agrees with the position of Iachello and Hong [100], who do not consider data protection and personal privacy as adversary concepts. In the case of USN technology, both concepts are crucial for ensuring its long-term success. For example, a USN application should include data protection solutions for securing safe collection and dissemination of users' personal information. As well, it needs solutions regarding personal privacy for helping users to accurately select the right information to be disclosed to the right person under the right circumstances.

Even if data protection is a complex problem worth investigating for the USN technology, it will not be further discussed in this dissertation. This research assumes a non-malicious infrastructure and targets at preventing accidental data disclosure, where personal information is unintentionally revealed, with or without

previous inquiry. Thus, the focus has been given on the management of personal privacy, which requires constant and arbitrary negotiation about data sharing preferences in USN environments.

1.2 Problem formulation

The key problem, addressed in this dissertation, is the difficulty of promoting privacy-aware social networking in ubicomp environments that focus on maximizing potential networking benefits while minimizing users' privacy concerns. To contribute to solving this problem, this research focuses on design, architecture, implementation and evaluation of USN concept and related services aiming at effortless disclosure of relevant, but not sensitive, users' personal information according to the different circumstances. Therefore, the main research question, posed to this dissertation, is the following:

Q: How can personal privacy be managed in ubiquitous social networking environments?

The management of personal privacy in USN should be researched in two different directions, as the selection of users' personal information to be disclosed in such environments occurs in two different situations. The first situation is related to the selection of the overall users' personal information to be available for the USN services. For example, when people sign up to any online social networks sites, they have to fulfil a user profile. Such kind of data disclosure decisions are principally made before actual use of the systems and would be highly influenced by designing

1. INTRODUCTION

services that would increase users' comfort with sharing of their personal information [12, 100, 120, 122]. Therefore, the first direction related to the management of personal privacy in USN environments leads to the following subquestion:

Q1: How can ubiquitous social networking be designed to increase users' comfort with sharing of their personal information for effortlessly maximizing potential networking benefits while preserving users' personal privacy?

The second situation regards the selection of personal data that should be disclosed to others under different circumstances. For example, after registering to an online social networks site, users have usually the possibility to customise their data disclosure decisions, based on different categories of inquirers, e.g. co-workers, friends and strangers. However, such kind of data disclosure decisions in USN are not influenced by only the identity of the inquirer, but they are also impacted by the current circumstances of the users' encounters, especially when taking into consideration the specific focus on supporting social interactions between strangers [45, 170]. In fact, meeting a stranger at work would probably lead to different data disclosure decisions than encountering the same person at friday night in a bar in the center of the city [64]. Thus, the complexity of evaluating which kind of users' personal data that should be disclosed to others is significantly increased. In these situations, personal privacy cannot be managed with predefined preferences often applied in online social networks, such as *private*, *work* or *public*, but it requires a more fine grained selection of personal data. This dissertation addresses the necessity to gain an extensive comprehension of the variation of human data sensitivity that affects information disclosure under different circumstances in USN

environments [170]. It suggests managing personal privacy as individuals do in ordinary human interactions, where they intuitively evaluate various determinants and unconsciously choose what personal information to disclose during face-to-face interactions [26, 122]. Therefore, the second direction related to the management of personal privacy in USN environments leads to the following subquestion:

Q2: Which are the influential factors that would affect the variation of human data sensitivity and consequent data disclosure decisions in ubiquitous social networking during users' encounters?

The following sections present the research background, related to the introduced research questions.

1.2.1 Designing for personal privacy in ubiquitous computing

As earlier introduced, users are nowadays more willing to accept potential privacy threats if the advantages for sharing their personal information are perceived to be higher than potential privacy risks [119, 160]. However, one of the major concerns highlighted by Bellotti and Sellen [12] in regard to technologically sophisticated environments, such as ubicomp, is that these technologies do not provide adequate support for avoiding violation of privacy and considerably disempower users' control over the management of their personal information. Thus, the authors suggested that *privacy should be a central design issue in its own right*.

Bellotti and Sellen [12] emphasized that design of ubicomp should address disembodiment and dissociation privacy threats. The former regards the danger that users would not be able to present themselves to others as they do in face-to-face

1. INTRODUCTION

interactions. Dissociation refers to the threat that the results of actions are possible to be seen, while the actions themselves are invisible. In order to avoid such privacy threats, the authors proposed that the design of ubicomp should comprise control and feedback principles. The control principle should allow users to decide what to disclose and whom to disclose, while ensuring subsequent feedback about data disclosure decisions should provide opportunities for users to be aware of when and what data is being shared and who can access it.

Palen and Dourish [149] provided more theoretical insights into the disembodiment privacy risk. The authors were inspired by the work of Altman [4, 5], who describes privacy as a dynamic process, representing continuous negotiation and management of the boundaries that shape data disclosure. Palen and Dourish identified three dynamic boundaries for negotiation of users' personal data disclosure. Firstly, the privacy and publicity boundary separates personal information into the disclosed and retained data sets. Then, the identity boundary defines the role, represented by the user based on the time, place and situation contexts. Finally, the temporal boundary regards the past, present and expected future of the users. The authors concluded that data disclosure decisions are taken by continuously negotiating the internal conflicts between the elements of the three identified boundaries.

Jiang et al [104] attempted to refine the work of Bellotti and Sellen in relation to the dissociation privacy risk. The authors suggested to encourage minimum information asymmetry between the parties (i.e. data owner, data collector and data user), by either decreasing the flow of information from data owners to data collectors and users or otherwise increasing the flow of information back to the data owner. As well, Langheinrich [120] suggested that, among other principles, notice

and explicit consent should be included into the design of ubicomp environments. The former refers to the right for end users to be notified when other entities collect or disseminate personal information, while the latter empowers people to provide explicit consent on a case-by-case basis. These principles constitute recommendations that privacy models would demand the attention of the individuals under circumstances that would threaten their privacy.

In addition to general principles, Lederer et al [122] suggested specific guidelines for ubiquitous social computing environments. The authors argued that users must be empowered to take informed data disclosure decisions, by proposing to address the socio-technical gap, introduced by Ackerman [1]. The socio-technical gap refers to the division between “*what we know we must support socially and what we can support technically*”. Lederer et al discussed that in case an intermediary point of the socio-technical gap is not found, the user would be either overwhelmed or disempowered, which would both result in uninformed and impulsive data disclosure choices. To find this balance, the authors attempted to reconcile Palen and Dourish’s theoretical insights [149] with Bellotti and Sellen’s [12] technical solutions. As well, Lederer et al took into account the Fair Information Practices, outlined by Langheinrich [120] and attempted to encourage minimal information asymmetry between the parties, as suggested by Jiang et al [104]. These privacy guidelines firstly focus on enabling users to understand the actual and potential impact of their data disclosure, by informing who is the recipient and what information is disclosed, as well as about the privacy implications of their data disclosure, e.g. how the information is shared, the presence of third party observers, etc. Moreover, these guidelines also aimed at allowing users to perform natural social actions by (i)

1. INTRODUCTION

managing users' personal privacy as a natural consequence of their normal engagement with the environments, (ii) providing a binary choice for halting and resuming data disclosure and (iii) giving the opportunity for users to transfer established social practice, such as ambiguous information and plausible deniability, to ubicomp. The disclosure of ambiguous information would empower users to provide imprecise data about themselves, while the plausible deniability refers to the possibility for users to deny data disclosure, without revealing whether it was intentional or not.

These design guidelines, proposed by Lederer et al [122], indisputably comprise aspects that are crucial for ubicomp environments that support social interactions between their inhabitants. However, other important design factors can be identified when taking into account the particular focus on promoting privacy-aware social networking among inhabitants of ubicomp. For instance, as suggested by Lederer et al, this dissertation acknowledges the importance of informing users about future potential implications of their data disclosure, e.g. informing them whether their data is as well shared with third parties. On the other hand, it also seeks the necessity to prevent such implications, because a set of data that is not considered sensitive today might create major user's privacy worries in the future [6, 84, 91, 130, 168]. Further, this dissertation recognizes the importance of taking into account that many users might be displeased when too much users' intervention and attention is required. For example, as suggested by Langheinrich [120], rigid rules for ensuring choice and notice might not be possible for efficient and reliable implementations of ubicomp environments. They would probably result in too many interactions between the users and systems for either approving users' data disclosure decisions or informing when the collection and dissemination

of their personal information occur. While these principles are crucial for ensuring better management of users' personal privacy, they must as well be designed to embrace Weiser's vision of calm technology [191, 193]. Lastly, the reviewed guidelines do not balance privacy concerns with potential networking benefits, as the human data sensitivity variation depending on different circumstances of users' encounters is not taken into consideration for data disclosure decisions. For instance, if the users decided to share detailed profiles for maximizing potential networking benefits, it could result in disclosure of personal information, which is sensitive under certain circumstances. Contrarily, if the users provided only a limited profile in order to avoid including personal information, which is too sensitive to be disclosed under some circumstances, it could result in loss of potential networking benefits in other situations where such personal information is not considered to be sensitive anymore.

For the reasons presented above, this dissertation addresses the necessity to enhance the current design guidelines, by additionally suggesting others to be taken into consideration when designing for personal privacy in USN environments. These guidelines should aim at addressing potential users' concerns arising as a result of their participation in USN environments by creating more functional and privacy-oriented services, where users would feel more comfortable with sharing their personal information for gaining networking benefits in exchange.

1.2.2 The variation of human data sensitivity

A clear understanding of the degree of sensitivity of users' personal information would provide a significant input for designing systems, which manage personal

1. INTRODUCTION

privacy in ubicomp environments focusing on promoting privacy-aware social networking. However, attitudes towards data disclosure decisions vary across different people and situations with diverse levels of trust and needs [90, 100]. Some types of information are perceived to be more sensitive than others, e.g. personal interests data is generally less sensitive than contact information [2]. Furthermore, different information within the same data type was also found to have varying sensitivity levels, e.g. phone number was discovered to be more sensitive than personal email address [49]. Finally, even the sensitivity of the same personal information was as well observed to vary under different circumstances [45, 123, 169, 170]. These studies highlight that disclosure of personal information in ubicomp environments cannot be seen as a static notion. Instead, different situations have different privacy implications and the consequent data disclosure decisions rely on the individuals' acceptance of actual risks, presented at the moment of the disclosure [170].

Inspired by the work of Goffman [79, 80], who highlighted the needs for users to project different personas under different circumstances, Lederer et al [122] attempted to incorporate such research into practical solutions. The authors designed and evaluated privacy management systems that rely on either *predefined sharing preferences* or *ad hoc privacy control*. The predefined sharing preferences approach attempts to predict all the potential circumstances and associated data sharing decisions a priori the actual data disclosure. On the contrary, the ad hoc privacy control solutions support data disclosure decisions *in situ*, i.e. at the moment of actual disclosure.

Firstly, Lederer et al proposed a predefined sharing preference model, called Faces [124]. Similarly to other solutions [42, 103, 111, 144], Faces users are asked to

indicate "who" can access "what" and "when", a priori any data disclosure. Thus, users set their privacy rules through a desktop application. These rules support four predefined levels of privacy protection, ranging from "undisclosed" that defines absolute confidentiality, to "precise" that allows openness of entire user's personal information. When the users meet one of the predefined circumstances, the Faces repository, where all the predefined preferences are stored, is queried and it discloses the corresponding users' predefined sharing preferences. Additionally, in case of unknown inquirers or situations, a default data disclosure decision is provided. After testing the Faces privacy model, the authors concluded that privacy models should avoid prediction of all the potential situations and associated data sharing decisions a priori the actual data disclosure, as they observed that users encountered situations where data disclosure decisions were not accurately predictable. In such cases, the authors emphasized that rules indicated by predefined data disclosure decisions would probably lead to invasion of privacy, because they would not meet the actual users' sharing preferences.

Lederer et al [122] subsequently argued that ubicomp privacy management systems should be designed to facilitate the dynamic and intuitive aspects of privacy and thus allow users to adjust their decisions while meeting the actual circumstances. The authors upgraded the Faces model into the ad hoc privacy control, called Precision Dial. In comparison to Faces, Precision Dial removed the preconfigured privacy preferences and added a quick manual selection of one of the four privacy protection levels, introduced in the Faces privacy model. In Precision Dial, while encountering different circumstances, the user has the opportunity to manually adjust his privacy settings when needed, similarly to the practice of adjusting

1. INTRODUCTION

ring volume of mobile devices [122].

Bünning and Cap [27] agreed with the vision of Lederer et al [122] and as well suggested that privacy management systems should be designed to closely reflect on people's natural privacy handling by enabling ad hoc data disclosure decisions. However, the authors highlighted that, if not well designed, ad hoc privacy control models might present crucial disadvantages, as they might require too much users' attention and intervention. For example, despite achieving the goal of allowing users to take ad hoc data disclosure decisions, the Precision Dial privacy model demands a considerable amount of users' attention and intervention, because users are continuously required to adjust their data disclosure settings. Moreover, if the user forgot to update the current privacy protection level when encountering different circumstances, it might result in unintentional data disclosure decisions and consequent invasion of personal privacy or loss of potential networking possibilities. The former would occur in case users had previously selected a more open privacy protection level than the desired one, while the latter would happen, if they had chosen a more confidential privacy protection level than the preferred one.

To address these limitations, Bünning [26] suggested a privacy model, called Disclosure Decision Model (DDM) that focuses on relieving the users from frequent data disclosure decisions. Specifically, DDM can be considered as an agent that manages information disclosure on behalf of the user by relying on previous data disclosure decisions. In DDM, users' attention and intervention is only expected in case of user's disagreement with the automated data disclosure decisions. The disagreement between the user and the DDM model about data disclosure decisions can be considered as the main challenge of such privacy control mechanisms. This

dissertation seeks the necessity to investigate solutions for reducing such interaction at minimum. As earlier introduced, it suggests analysing the variation of human data sensitivity under different circumstances, with the target of identifying and analysing the influential factors that would impact data disclosure decisions, during users' encounters in USN environments.

In previous studies, Lederer et al [125] identified three different levels of abstractions for determining data disclosure decisions in ubicomp: inquirer, situation and accuracy preferences. The inquirer is considered to be the individual that the user is interacting with and the situation is defined according to the circumstances at that time. The authors claimed that the accuracy of information disclosed is influenced by the identity of the inquirer and the situation at the time of the inquiry. In [123], Lederer et al provided more insights into these factors and determined the identity of the inquirer to be the most important factor, influencing the users' data disclosure decisions, followed by the situation as parameter of secondary significance. However, despite the current trend to indicate the identity of the inquirer as the most crucial determinant for data disclosure in ubicomp, e.g. [57, 108, 148, 199], Consolvo et al [45] also discovered other factors, such as purpose of disclosure, current activity and mood, which impact disclosure of personal information in ubiquitous social computing environments. Following these results, as well as due to focus on USN and consequent disclosure of personal information to strangers, this dissertation advances the attention to the current circumstances at the moment of data disclosure (e.g. current activity and location) as well as other information, which the user has in common with the inquirer, such as similar movie and sport preferences or mutual friends.

1.3 Research design and approach

This dissertation researches a complex situation in a real world setting, i.e. promoting privacy-aware social networking in ubicomp environments. Thus, it applied a methodology that aimed at improving management of users' data disclosure to maximize potential networking benefits, while minimizing users' privacy concerns. This methodology comprised iterative analysis, design, development and implementation, leading to contextually-sensitive results [190].

As the intrinsic nature of the research problem is embedded in the real world, the research design took a pragmatic stance focusing on developing theories and designs in a parallel manner through the research process, in order to best link theory to practice [22, 60, 190]. USN can be considered as a new emerging technology - at the beginning of my PhD studies, only little research had been carried out and to the best of my knowledge, only one application was commercialized, i.e. Aka-Aki¹. These considerations were confirmed by the results of one of my investigations [166], carried out in 2009 with students of the Aalborg University, who were experienced with mobile services and with at least one of the available online social networks sites. At that time, 90% of the students did not know about the existence of similar services and 75% of them found USN to be innovative idea and were looking forward for these services to become widespread on mobile devices.

For phenomenological study characterized by lack of previous research, such as USN, Creswell [50] provided methodological recommendations emphasizing the need to identify the essence of human experiences about the phenomenon observed

¹<http://www.aka-aki.com>

1.3 Research design and approach

during empirical research. Riemen's work [161] agreed with Creswell's suggestions and as well proposed to investigate the research problem with studies that closely simulated the experience of the users, rather than on hypothetical basis. This dissertation applied the methodological recommendations for phenomenological research of Creswell and Riemen and thus targeted at studying this phenomenon based on perceptions of participants observed during empirical investigations. This approach naturally comprised complexities, dynamics and limitations of authentic practices, including many expected and unexpected variables.

To cope with such research problem, this research followed an integrative design as a wide range of techniques were adopted to probe personal privacy dynamics [100]. Mixed methods comprising both quantitative and qualitative forms were utilized, depending on the needs of the research, in order to gain a broad understanding of the research problem and ensure greater overall objectivity, validity and applicability of results [50, 51]. For answering the research question Q1, as suggested in [50, 142], qualitative research was used to better understand the USN phenomenon and identify the variables for designing privacy-aware USN that would allow users to feel more comfortable with sharing of their personal data.

In relation to the research question Q2, this study acknowledged that previous works already provided relevant insights into the factors for the variation of human data sensitivity, related to users' participation in ubiquitous social computing environments. However, this dissertation identified the need for further investigation of these factors with particular focus on USN services. In this case, quantitative and qualitative investigations were required to both increase generalizability of the findings as well as develop a detailed view of meaning of a phenomenon for individuals

1. INTRODUCTION

[50].

The studies, included in this dissertation, were cross-sectional and thus aimed at investigating users' perceptions and behaviors in regard to data sharing preferences in USN at one specific point in time. Notably, this dissertation acknowledges that a prospective longitudinal approach might as well provide valuable insights into the researched phenomenon, as privacy expectations and perceptions in relation of personal data disclosure decisions might change over time, due to increased familiarity with a particular technology. However, this research did not focus on the temporal tendencies of users' perceptions and behaviors, but instead it targeted at accurate assessment of current privacy perceptions for the acceptance of the emerging USN technology [54].

The research design applied in this dissertation was not only integrative, because it comprised different study methods, but it was also flexible and iterative. In fact, it provided great flexibility for progressively refining theories and practices throughout iterative revisions of analysis, design, evaluation and redesign [44, 190]. At beginning of this research, the initial research plan was not detailed to account for managing personal privacy in USN environments and when necessary changes were introduced and implemented in the upcoming investigations of this study. All the empirical studies were firstly designed according to the analysis of the available literature, however outcomes from previously conducted investigations led to new expectations that become the main target during the next cycle of analysis.

The obtained outcomes of the different empirical studies were specifically related to the research settings where the investigations were carried out. However, this research not only provided findings for contributing to the problem of enabling

privacy-aware USN, but it also aimed at describing in details the problem setting for (i) helping to better interpret the findings as well as (ii) guiding further research for both evolving relevant theory and generating new findings [190]. Finally, when possible, results of this dissertation were validated in more than one context for enabling increased generalizability [60].

The following sections provide more insights into the selection of the participants, research process as well as quantitative and qualitative methodologies, applied in this dissertation.

1.3.1 The selection of participants

The method for selecting the participants of the empirical investigations followed the suggestions of Von Hippel [97], in relation to evaluation of new emerging technologies, e.g. USN. The author recommended to contact and interview the most advanced users in the field of interest. The reason for recruiting this type of users, called lead users, is motivated due to their interests and predisposition to innovative product ideas, as they are capable of facing needs long time before others encounter them, and have already found solutions to address potential concerns. In [115], Kujala and Kauppinen supported the findings of Von Hippel by suggesting the recruitment of lead users specifically intended for field studies, as they discovered that one lead user provided as much information and ideas as five ordinary users did.

In order to identify lead users, previous studies recommended to take into consideration participants who (i) perceive certain demands earlier than others and (ii) expecting high benefits from a new emerging technology to their needs [76, 96, 132, 187]. For these reasons, the selection of participants for the empirical

1. INTRODUCTION

investigations of this research was limited to Facebook² users. It was determined this category to be lead users of the emerging USN technology, because they were expected to perceive potential networking benefits in exchange for their information disclosure better than ordinary users, even if the perception towards the networking services might vary between virtual and physical worlds. Further, Facebook users have already encountered relevant privacy concerns in relation to their data disclosure in such online services and found strategies to balance their privacy concerns with potential networking benefits.

The participants were asked to provide information about their demographic characteristics and to indicate their privacy preferences on visibility of their own personal data (e.g. user profile, pictures, posts) in the Facebook online social networks site. Based on these answers, it was possible to observe patterns among data disclosure attitudes and divide the participants into three privacy clusters, following the Westin/Harris privacy segmentation model [198]: fundamentalists, pragmatists and unconcerned. The fundamentalists were considered to be extremely concerned about sharing of their personal information with any other user. Pragmatists were also carrying about loss of privacy due to disclosure of their personal information, however they often had specific concerns and particular strategies for addressing them. Finally, unconcerned participants were very open to share their personal data, as they believe that their privacy was not jeopardized.

During the last quantitative investigation of this research, it was identified the 17.8% of Facebook users to be fundamentalists, the 64.4% of them to be pragmatists and 17.8% privacy unconcerned. Afterwards, during the recruitment of the partic-

²www.facebook.com

ipants of the qualitative investigations, this work aimed at achieving stratification between participants' privacy clusters to ensure that specific characteristics of individuals are represented in the sample in accordance to the proportion in the entire population [75]. The target was to obtain similar proportions of participants' privacy clusters in reference to the last quantitative investigation of this study, where a random sample was selected. Finally, in all the qualitative studies of this research, the selection of participants was specifically restricted to the ones, who presented high interest on this new emerging technology, e.g. they claimed to be potential users of USN services.

The number of participants for the qualitative investigations was difficult to be defined, because the larger sample is always considered to be the better, but many research projects, as the one documented in this dissertation, have a restricted amount of available resources [115]. This research followed the suggestions of previous studies that recommended a number between 6 and 20 participants in order to collect significant useful information about the product development of new emerging technologies [15, 85, 116]. Finally, due to anonymity of the responses and different timeframes between the quantitative investigations and the other studies of this research, it cannot be ensured that the participants of the qualitative investigations were as a subset of the quantitative analysis. However, a slightly overlap might be expected.

1.3.2 Research process

As shown in Figure 1.3, the research design was composed of three different phases. The first phase comprised a literature review that provided the research background

1. INTRODUCTION

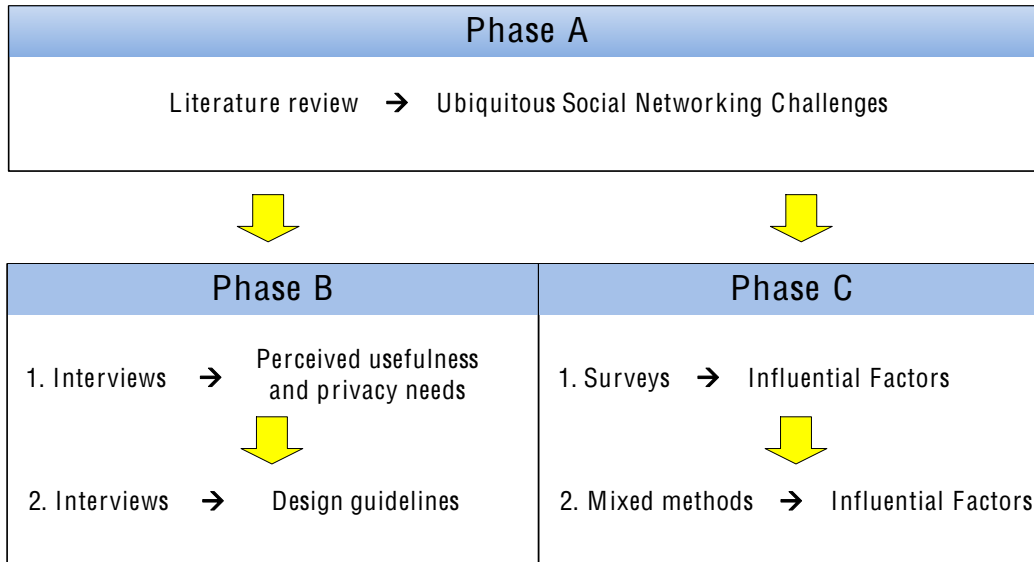


Figure 1.3: Research process

as well as the inputs to the other two phases, which were carried out at the same time. Phase 2 and Phase 3 were composed of empirical studies to investigate the research questions, introduced in Section 1.2. Details about the three different phases related to this research project are described in the following.

Phase 1 Initially, the available literature, related to the actual research problem, was identified and reviewed in order to gain insights into the status quo of USN. The literature review identified three different challenges for the further development of ubiquitous computing environments that target at promoting social networking: context acquisition, enhancing social networking and privacy. The analysis of the available literature led to the identification of a relevant gap, which if addressed, would significantly contribute to the development of USN. The identified gap refers

to the management of users' personal privacy in ubicomp environments with the goal of supporting social interactions between the users. In order to contribute to the identified gap of research, this phase further focused on the review of literature in the relevant field by surveying previous works on topics, related to management of users' personal privacy. Furthermore, it recommended upcoming research to study both technical and human aspects, related to the disclosure of users' personal information in USN environments in exchange to potential networking benefits.

Phase 2 The second phase was designed based on the outcomes of the literature review and it is related to the technical aspects of data disclosure in USN environments. During the first phase of research, relevant privacy guidelines for ubiquitous social computing environments were reviewed and suggested to be taken into consideration when designing USN environments. These guidelines target at empowering users to take informed data disclosure decisions. However, in Section 1.2, this dissertation highlighted the necessity to enhance the reviewed design guidelines, by additionally suggesting others that would increase the users' comfort with the selection of their personal information before the actual usage of the system.

As shown in Figure 1.3, the first study of Phase 2 analyzed the users' perceptions towards privacy-aware social networking and their related potential privacy risks. This investigation followed the work of Hayes and Abowd [94], who indicated that users would be more inclined to accept potential privacy risks, if they perceived the technology to be useful by enhancing their networking performance [56]. Further, it focused on discovering the users' privacy needs for ensuring acceptance of crucial requirements for the establishment of USN. These requirements refer to the exchange

1. INTRODUCTION

of users' personal information with strangers, announcement of users' presence as well as potential initiation of face-to-face interactions. In total 16 participants were recruited and introduced to the USN concept and various prototypes, with particular focus on potential networking benefits and privacy threats. Afterwards, the participants were interviewed to analyze (i) the perceived usefulness and (ii) users' privacy needs for the acceptance of the USN technology.

The result of this qualitative investigation provided inputs to the second study of Phase 2. The second investigation researched solutions for the design of privacy-aware USN services with the target at maximizing potential networking benefits while preserving users' privacy. When surveying existing USN prototypes, many crucial usability and privacy limitations, related to the discovered privacy needs, were identified and not to be taken into account in the existing privacy guidelines of ubicomp. Consequently, the second study of Phase 2 focused on updating the existing privacy design guidelines of ubicomp environments by additionally proposing others that are suggested to be taken into consideration when designing for personal privacy in USN. Further, it evaluated whether the users' perceptions of personal privacy were positively affected when the design of USN followed the proposed guidelines by conducting a qualitative investigation with 15 new participants. The selected sample of participants did not overlapped the one recruited for the first study of Phase 2. The decision to recruit new participants was taken in order to ensure the validity of answers, because the qualitative investigation, carried out during the second study of Phase 2, was mostly based on the identified participants' privacy needs, discovered during the analysis of the first study of Phase 2.

Phase 3 The third phase of research focused on the human aspects of data disclosure in USN environments. It aimed at investigating the relevant factors for the variation of human data sensitivity under different circumstances, as suggested during the literature review of Phase 1.

The first study of Phase 3 consisted of identification of the determinants that might shape human data sensitivity in USN. The factors that were taken into consideration were already found to be relevant in analysis of users' data disclosure decisions occurring in ubiquitous social computing environments. These determinants were subsequently investigated through a quantitative study with the focus of promoting networking in ubicomp environments. Two online surveys were carried out to probe preferences of data disclosure under different circumstances. The questionnaires were distributed to 500 potential respondents. In total 121 complete answers were received for the first survey and 101 answers for the second survey. Due to anonymity of the responses and different timeframes of the surveys, it cannot be ensured that the participants of both surveys completely match, however a significant overlap is expected. The findings of the two online surveys provided statistically significant results to contribute to the design of privacy management systems in USN environments. However, this dissertation acknowledges the necessity to additionally investigate the identified influential factors with users' sharing preferences, made at the moment of actual disclosure. This was necessary in order to increase realism and validity of the analysis about participants' data sharing preferences under different circumstances, because predefined data disclosure decisions might present only participants' attitudes, which could differ from users' actual behaviour [13, 100].

1. INTRODUCTION

The second study of Phase 3 was designed based on the results of the first study and it aimed at investigating users' ad hoc data disclosure decisions. In total 13 participants were recruited. The selected sample was a subset of the one recruited for the qualitative investigation, carried out during the first study of Phase 2. Unfortunately, it was not possible to recruit the complete sample of participants, because three of them could not take part in this investigation. The decision to recruit the same participants from the first qualitative investigation, carried out during Phase 2, was influenced by the following reasons. First, the participants have already gained experienced about USN and they were identified to be lead users of the USN technology. Second, both investigations adopted the same mobile prototype for data collection of users' sharing preferences, thus it was possible to analyse the collected data with different purposes, related to the different goals of the two studies. Finally, the knowledge acquired by the participants during the qualitative investigation of the first study of Phase 2 were not considered to bias the results of the analysis of Phase 3, as the two studies presented completely different targets.

The 13 participants engaged in a sequential mixed methods study, including quantitative and qualitative investigations. Firstly, a quantitative research investigated the relationship between the identified influential factors and participants' sharing preferences, based on a collection of a large amount of ad hoc data disclosure decisions. Information, acquired during the first phase of the mixed methods study, was explored further in the second phase, where qualitative interviews were conducted to better understand the impact of the influential factors on participants' personal data disclosure decisions, as well as to research on subjective motivations

causing the quantitative results.

1.3.3 Investigation methodology

In the following, the details about the methodological choices regarding the quantitative and qualitative investigations, carried out during this research, are provided. However, more information and motivations behind the methodological quantitative and qualitative choices are presented in the individual papers, included in this dissertation.

Qualitative investigations

As shown in Figure 1.3, qualitative studies were carried out in Phase 2 and in the second study of Phase 3. At beginning of each of the qualitative investigations, participants were helped to get more familiar with the USN concept. Each of them was introduced with the existing USN prototype Spiderweb [166] as well as its services (presented also in this video³) and other USN applications, already available in the market, i.e. Sonar⁴ and Aka-Aki⁵. Participants had also the opportunity to learn how potential networking benefits can be gained through USN, as shown in this video about Aka-Aki⁶. Further, they were presented to different scenarios from everyday lives, where these services might be applied, such as professional areas, dating and big events, as described in [65, 168]. Finally, it was discussed with the participants the potential networking benefits in the identified application areas as

³<http://www.youtube.com/watch?v=DgeVNv10CIM>

⁴<http://www.sonar.me>

⁵<http://www.aka-aki.com/>

⁶<http://www.youtube.com/watch?v=mvRgtT4LawU>

1. INTRODUCTION

well as possible privacy threats that might arise as a result of the information disclosure in USN. In the following, the techniques utilized for collection of qualitative data are described, followed by the strategy for data analysis.

Data collection Qualitative interviews were preferred alternatively to other investigation methods, such as handing out questionnaires or establishing a focus group interview. This method was chosen because of the following two reasons: (i) lack of participants' extensive experience in utilizing USN services and (ii) potential misinterpretation of the research questions due to their complexity and ambiguity, which might be caused when evaluating new emerging technologies. Moreover, it was decided to run semi-structured interviews to better understand the motivation behind the participants responses and ensure that general areas of information are collected from each participant, however still allowing adaptability of the interview process [50, 118, 136]. The interviews were audio taped for a duration of 30-60 minutes. Questions were related to the different targets of the investigations that are briefly recalled in the following:

Phase 2: Questions were firstly related to the perceived usefulness of USN services in order to investigate the degree to which participants believe that using a particular technology would enhance their networking performance with strangers [56]. Further, other questions were concerning the acceptance of crucial prerequisites for the establishment of USN services, i.e. announcement of user's presence, disclosure of personal information and potential initiation of face-to-face interactions.

In the second study, questions were related to the identified design guidelines,

suggested to be taken into consideration when designing for personal privacy in USN. Per each of the suggested privacy guidelines, participants were presented two different scenarios of USN services: the first scenario was based on a design of USN services that do not follow the proposed guideline, while the second scenario was based on a design of USN, which respects the proposed guideline. Thus, it was discussed with the participants whether they would feel more comfortable with sharing of their personal information in USN when the proposed guidelines were followed for better protecting their personal privacy.

Phase 3: Questions were related to the selected influential factors and respective statistical results, obtained during the quantitative analysis, carried out during the second study of the Phase 3. Per each of the influential factors, it was asked to the participants to reflect on how important each of the factors was for their personal data disclosure decisions and to elaborate on the reasons. Moreover, after showing the statistical results of the quantitative investigation to the participants, it was inquired whether they could confirm these results and comment on any surprising outcomes, obtained during the analysis of the quantitative investigation.

Data analysis The strategy utilized for analysis of the information, collected during the qualitative interviews, follows a hierarchical approach, illustrated in Figure 1.4. At beginning the qualitative interviews were transcribed (step 1) and reviewed in order to gain a generic understanding of the participants' attitudes towards data disclosure decisions in USN (step 2). In step 3, the transcribed answers were organized in different segments based on the parameters of the research phase (e.g.

1. INTRODUCTION

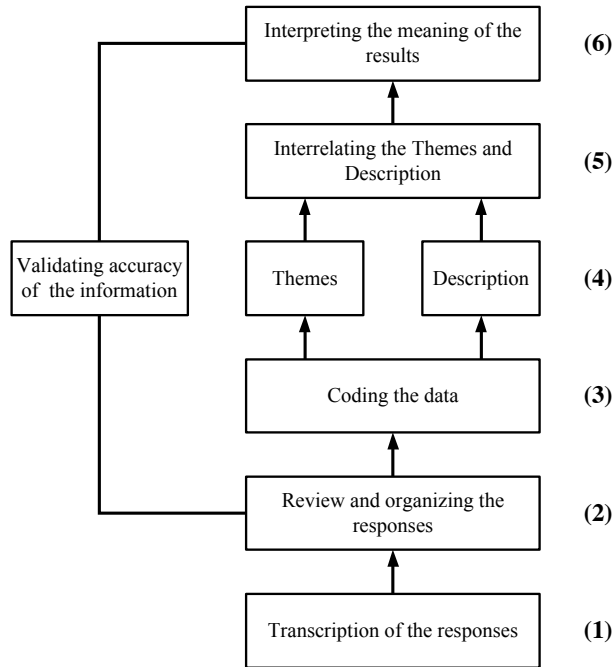


Figure 1.4: Qualitative data analysis

influential factors for Phase 3) and, in the next step, the descriptions of participants and themes were generated. The former regards information about the participants (e.g. gender, privacy clusters), while the latter refers to the categories of the major research findings. In step 5, the themes and descriptions data categories were interrelated and, in the last step, the qualitative data were interpreted while taking into account, when relevant, the interconnection between themes and descriptions.

In order to ensure accuracy of the findings, two different techniques were applied during the analysis of the qualitative data, as shown in Figure 1.4. The first one was the triangulation of different data sources, only carried out in Phase 3. This technique was utilized during the second step of the data analysis, as illustrated

1.3 Research design and approach

in Figure 1.4. Particularly, it was provided to participants a questionnaire for classifying the influential factors according to the impact that they had on their personal data disclosure decisions. Consequently, it was possible to understand whether the first review of the transcripts resulted in correct assumptions.

The second technique, i.e. member checking, was carried out during the last step of the data analysis in both phases of this research. This approach was useful for determining the interpretations' accuracy of the collected qualitative responses. Specifically, the final findings of the qualitative investigations were sent out back to the participants in order to get feedback on the accuracy of interpretation. When needed, follow-up interviews with the participants were conducted to give them the opportunity to additionally comment on the findings.

Quantitative investigations

Quantitative investigations were carried out only in Phase 3, which was related to the study of influential factors for analyzing the variation of human data sensitivity under different circumstances. In order to gain insight into human data sensitivity, the participants were asked to indicate personal information that they would like to share under different circumstances of their lives. They were informed that sharing of personal data is motivated by potential networking benefits, provided in return to disclosed information. Naturally, the benefits would be directly proportional to the amount of shared information, thus participants were asked to compromise between privacy risks and potential benefits. In the following, the techniques utilized for collection of quantitative data are described, followed by the strategy for data analysis.

1. INTRODUCTION

Data collection The first study of Phase 3 comprised two surveys that aimed at collecting a large number of answers regarding data disclosure decisions in USN with the goals of identifying the relevant influential factors that would impact users' sharing preferences in such environments. This investigation relied on the predefined sharing preferences approach, which attempts to associate data disclosure decisions to selected potential circumstances. The circumstances taken into consideration at this stage remained at a general level and they were grouped into the most common life situations, such as work and social environments.

The second study of Phase 3 targeted at complementing the outcomes of the previous quantitative investigation and it focused at collecting users' sharing preferences, made at the moment of actual disclosure. Participants were asked to utilize a mobile application that simulates the USN behaviour. It was preferred to provide a new mobile application, rather than utilizing the Spiderweb mobile social networks [166] or other existing USN applications, due to reasons that are explained in the following. First, Spiderweb and the other applications are not widely spread yet and participants would probably encounter difficulties in finding opportunities to disclose their personal information to other *real* users. Second, some components and services related to the automated creation of users' sub profiles were not implemented in the Spiderweb prototype and thus participants could not have the opportunity to customise their data disclosure decisions according to different circumstances, but they could only select two different level of privacy, i.e. public (visible to everyone) and private (visible to only friends). On the contrary, the provided USN prototype was explicitly designed to collect data about participants' information disclosure decisions to strangers and their related current circumstances

(e.g. location, activity, mood, number of mutual friends) for further statistical analysis.

Three screenshots of the USN prototype are shown in Figure 1.5. Several times a day, the USN prototype was randomly asking participants to specify their current circumstances and their related ad hoc data disclosure decisions. Differently from the data collection adopted in the two online surveys, the quantitative investigation, carried out during the second study of Phase 3, comprised a broader selection of the current circumstances. For example, Figure 1.5-A describes a manual input insertion of participants' current activity. The participants could either select one of the available options or manually include an unrestricted description to indicate their current activity. After the participants provided information about the current circumstances (e.g. their current location, mood, activity) and, in some cases, being aware about information that they have in common with the hypothetical

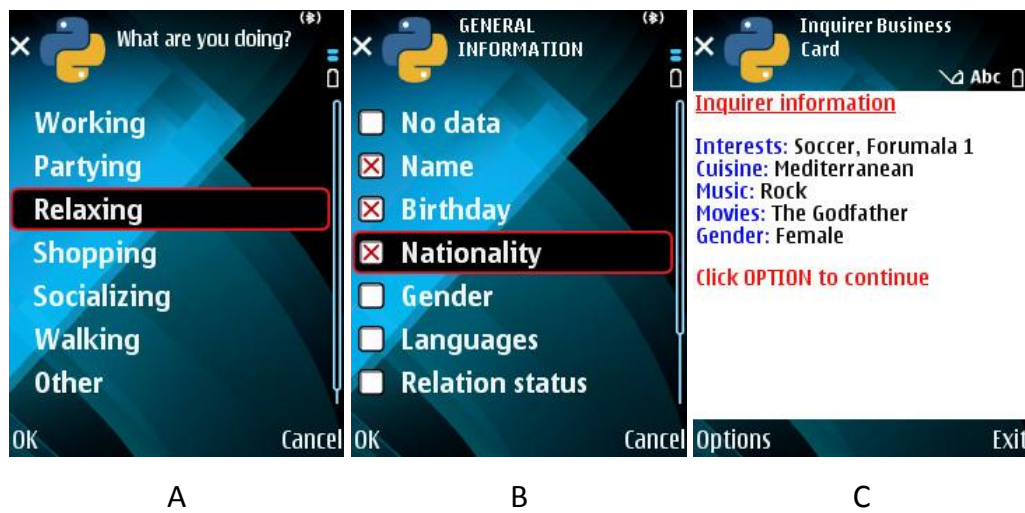


Figure 1.5: Three screenshots of the provided mobile prototype for collecting ad hoc data disclosure decisions

1. INTRODUCTION

inquirer (e.g. personal preferences or number of common friends, as illustrated by an example illustrated in Figure 1.5-C) the USN prototype guided the participants to express their ad hoc data disclosure preferences, as shown in Figure 1.5-B. The selection of data types to be disclosed was provided in accordance to data categorization in popular online social networks sites (e.g. gender, age and favorite music). The detailed description of the provided data types can be found in [170] and it was used in all the empirical investigations, carried out during this project. The selected ad hoc data disclosure decisions were stored in the local memory of the provided mobile phone together with the respective circumstances in order to be applied for further statistical analysis.

Data analysis In the first study of Phase 3, the collected predefined data disclosure decisions were analyzed by applying the Wilcoxon Signed Rank test. It was preferred to utilize a non-parametric test, because the responses of both surveys, grouped by different clusters (e.g. gender, age, occupation, privacy), were not normally distributed. In such case, a non-parametric statistical test was suggested to be selected, due to expected higher precision of the results in comparison to the parametric tests [141]. Further, the Wilcoxon Signed Rank test was specifically chosen, because the investigation aimed at analyzing two datasets of users' sharing preferences under different circumstances and evaluate whether their population means differ [201]. When analyzing more than 2 datasets, the Bonferroni correction was used to evaluate the results in order to avoid potential type I errors [195].

In the second study of Phase 3, the collected ad hoc data disclosure decisions were analyzed by applying the logistic regression method. This approach was se-

lected, because it does not require strict assumptions as other statistical methods like ordinary least squares regression or linear discriminant function analysis [151]. In contrast to the other two mentioned methods, the logistic regression does not assume linearity between independent and dependent variables nor normality or equal variance within each group of the independent variables [28, 71, 158, 182]. Moreover, it was decided to run a Binary Logistic Regression, instead of other kinds of logistic regression methods, such as multiple or ordinary, because the dependent variable was dichotomous, i.e. either disclose the information or not, and the categorical typology of the independent variables, e.g. environment, activity, mood. The research hypothesis posed to the data was that the likelihood of a USN user to disclose specific personal information is dependent on the investigated influential factors.

1.4 Research contributions

The technological development of ubicomp technologies inspired researchers to look for new opportunities for enhancing human communication in the physical world. This led to numerous research challenges, which range from incorporating users' personal information into ubicomp environments to ensuring accurate selection of users' personal data to be disclosed under different circumstances. When addressing these challenges, a contribution to the long-term success of USN can be achieved by promoting privacy-aware social networking among inhabitants of ubicomp environments. The major contributions of this work are summarized in relation to the five papers, included in this dissertation, as follows:

1. INTRODUCTION

1. The results of the article, presented in Appendix A, identify the major challenges for ensuring privacy in USN environments and review already implemented solutions. The state of the art, presented in this paper, is useful for informing about an overall perspective of the challenges regarding privacy-aware USN;
2. The results of the article, presented in Appendix B, describe in details the concept of USN through an example scenario, USN application areas as well as present the design and services of the Spiderweb prototype for the establishment of social networking in ubicomp environments. Afterwards, it comprises results of a qualitative investigation that aimed at evaluating participants' perceived usefulness of USN and identifying users' privacy needs for accepting the crucial requirements for the establishment of USN. The analysis of qualitative data, described in this paper, informs about the participants' perception of USN and circumstances where they find such services to be useful. Further, it provides an understanding of privacy needs that users have in respect to the USN technology;
3. The results of the article, presented in Appendix C, describe a set of drawbacks that are recommended to be avoided when designing for privacy in USN. These privacy design guidelines attempt to overcome crucial usability and privacy limitations, identified during the analysis of already existing USN prototypes. Moreover, the paper presents results of a qualitative analysis for investigating users' perceptions towards protection of their personal privacy in relation to the identified drawbacks. The findings, described in this paper, are useful

for informing about common USN design mistakes as well as design solutions that could be applied for avoiding them.

4. The results of the article, presented in Appendix D and Appendix E, outline and investigate relevant influential factors, which were considered to be relevant for data disclosure in USN. The analysis of quantitative and qualitative data, described in these two articles, provides significant input for the design and development of privacy management systems for USN environments.

1.5 Dissertation outline

This dissertation is further composed of the following chapters:

Chapter 2: This chapter provides an overview of the motivations, methodologies and main results of the five journal articles that are included in this dissertation.

The five articles are listed below:

1. Privacy and technology challenges for ubiquitous social networking. The full text of the article is presented in Appendix A;
2. Ubiquitous social networking: concept and evaluation. The full text of the article is presented in Appendix B;
3. Designing for privacy in ubiquitous social networking. The full text of the article is presented in Appendix C;
4. Privacy analysis in mobile social networks: the influential factors for disclosure of personal data. The full text of the article is presented in Appendix D;

1. INTRODUCTION

5. The influential factors for variation of human data sensitivity in ubiquitous social networking. The full text of the article is presented in Appendix E.

Chapter 3: This chapter reviews and connects the findings of the five journal papers in order to answer the research questions, introduced in Section 1.2. It introduces and discusses the major challenges to be addressed for the further development of privacy-aware USN. Further, this chapter presents results of a qualitative analysis that aimed at investigating the privacy needs for ensuring the acceptance of USN services. It reviews relevant guidelines for empowering users to take informed data disclosure decisions in ubicomp environments and introduces new guidelines as well as their related design solutions. Moreover, it describes a privacy model, called Diverged Personalities, which targets at enabling users to disclose different profiles under different circumstances. Finally, this chapter presents and discusses results of empirical investigations, which focus on researching whether USN environments following the proposed design solutions and the identified determinants for variation of human data sensitivity would impact users' data disclosure decisions in USN.

Chapter 4: This chapter presents final conclusions in relation to the promotion of privacy-aware social networking in ubicomp environments that target at maximizing potential networking benefits, while minimizing users' privacy concerns. Particularly, it summarizes the answers in relation to the research questions, posed to this dissertation, and recommends further research areas within the USN topic.

2

Summary of the papers

In this chapter, an overview of the five journal papers that form the foundation of this dissertation is provided. These articles target at addressing the research questions, posed to this dissertation in Section 1.2, related to support of privacy-aware social networking in ubicomp environments that focus on maximizing potential networking benefits while preserving users' privacy.

For each of the papers, this chapter discusses the brief motivations that led focus on the specific research areas and problems that the papers address. Afterwards, it presents methods and principles applied in order to gain relevant insights for contributing to the selected research areas by addressing the identified research problems. Finally, a summary of the major findings and contributions is presented and discussed.

More detailed motivations, methodology and results are provided in the individual papers presented in the appendix of this dissertation.

2. SUMMARY OF THE PAPERS

2.1 Paper A

Privacy and technology challenges for ubiquitous social networking

Antonio Sapuppo and Boon-Chong Seet

Motivations

Ubiquitous social networking can be seen as an evolution of ubiquitous computing supporting the social well-being of people in their everyday lives. The vision of ubiquitous social networking focuses on enhancing social interactions among its participants during users' physical meetings. This target is leading towards important challenges such as social sensing, enabling social networking and privacy protection. Consequently, it is necessary to survey previous literature regarding these three challenges and to gain insights into existing potential solutions for further contributing to the long-term success of ubiquitous social networking environments.

Methodology

The paper reviews previous studies about the three identified challenges for enabling ubiquitous computing environments with emphasis on networking. It firstly focused on sensing of the relevant context for promotion of sociability among users. It presents test results of existing implemented solutions, based on wearable sensors, such as mobile phones. Afterwards, it reviews different design architectures that aim at supporting social networking between people in the physical proximity. Finally, the article reviews design guidelines for ensuring protection of users' privacy and presents existing privacy models that focus on managing information disclosure in

such environments, by empowering users to disclose different profiles under different circumstances.

Results

This paper draws attention to privacy as the main challenge for the development of ubiquitous social networking environments. Particularly, users' privacy must be ensured during acquisition, management and disclosure of users' personal data. It describes existing privacy protection laws and reviews previous works that adapted these laws to privacy system design principles. Further, the article presents existing privacy design guidelines and privacy models for ensuring both understanding of privacy implications of participation in such environments as well as possibilities to conduct socially meaningful actions. However, privacy was not considered as a standalone challenge, because it was found to be directly dependent on other two challenges, i.e. acquisition of the relevant context and design of software architectures for enhancing social networking. Thus, the paper also reviews methods and technologies for acquisition of users' identities and relationships, activities as well as data from online social networks. Afterwards, it describes diverse designs for promoting social networking, based on centralized, decentralized and hybrid architectures. Finally, the article suggests further research to investigate solutions for ensuring better management of users' personal privacy with recommendations to study both technical and human aspects, related the sharing of personal data in USN environments.

2.2 Paper B

Ubiquitous social networking: concept and evaluation

Antonio Sapuppo

Motivations

Despite the great success of online social networks, there is still no automated way to facilitate communication between people in the physical environments. Thanks to their wireless technologies, smartphones are capable of enabling opportunistic networks, where nodes are wirelessly connected and have the possibility to identify each other as well as exchange contents in a short communication range [95, 102, 153]. When users' personal data is incorporated into opportunistic networks, they can be perceived as an important tool for addressing sociability issues in the physical world, as they enable the establishment of USN services [10, 59, 65, 153, 183]. In order to contribute to the further development of USN, it is important to investigate whether users would perceive the usefulness of USN and understand their privacy needs for accepting the necessary requirements for the establishment of these services.

Methodology

The paper firstly describes a solution for enabling USN services and identifies three crucial requirements for their establishment: announcement of users' presence, disclosure of personal data and potential initiation of face-to-face interactions. Afterwards, a qualitative investigation with 16 participants was carried out. At beginning of the study, the participants were helped to get more familiar with the USN concept

during an introductory meeting, where several prototypes of USN services were introduced. Furthermore, different everyday life USN scenarios were described to the participants as well as they were introduced to potential networking benefits and privacy threats that this technology might raise. Finally, they also had the opportunity to exploit a USN prototype. Subsequently, they were interviewed with focus on perceived usefulness of USN services and acceptance of the three requirements.

Results

All the participants appreciated the possibility to be connected with other people nearby and especially with those who share distinctive interests and goals. They indicated professional purposes as the most relevant potential application areas for USN services and discussed that they would probably need time to get used to these services before utilizing them also for facilitating social interactions. Among the 16 participants, only two of them claimed that they would not be potential users of USN, if they had to accept the third requirement, which might lead to undesired face-to-face interactions. However, all the others accepted the possibility to initiate a face-to-face interaction with other users as long as they had a coarse-grained control over the USN services. Finally, participants did not present any crucial concerns about announcement of users' presence and data disclosure to other end users. As well, they commonly appreciated the possibility to utilize ad hoc privacy control for sharing different profiles under different circumstances. However, participants emphasized the needs to limit the autonomy of such privacy management systems in case of inquiry for highly sensitive data as well as to modify their personal data, even after actual disclosure.

2. SUMMARY OF THE PAPERS

2.3 Paper C

Designing for privacy in ubiquitous social networking

Antonio Sapuppo and João Figueiras

Motivations

As disclosing personal information is an intrinsic part of USN, these services are subject to crucial privacy threats [171]. Despite ongoing legal [62, 82] and academic [12, 104, 120, 122, 149] discussions about disclosure of personal information, the current designs of USN environments do not provide adequate personal privacy management for their inhabitants. When analyzing the design of existing USN prototypes, many privacy and usability limitations can be identified. If improperly addressed, these usability and privacy concerns could discourage users from disclosing their personal information and consequently threaten the further development of USN applications.

Methodology

The paper reviews existing ubicomp design guidelines that aim at allowing users to take informed data disclosure decisions. Afterwards, it presents current USN design solutions and identifies additional crucial usability and privacy limitations, which might discourage users from disclosing personal data in such environments. Based on these findings, it depicts four drawbacks that should be taken into consideration when designing for privacy in USN and suggests design solutions for avoiding the identified drawbacks. Afterwards, a qualitative analysis was carried out in order to

evaluate users' perceptions towards the protection of personal privacy, in relation to the proposed privacy guidelines. The analysis focused on investigating whether the proposed design solutions, heeding the drawbacks, are (i) *must have*, i.e. participants would not disclose their data if their design solutions did not avoid the drawbacks, (ii) *nice to have*, i.e. participants prefer services that avoid the drawbacks, however they would still share some of their information, even if the drawbacks were not heeded, or (iii) *indifferent*, i.e. participants do not consider that avoiding the drawbacks would provide any advantages, related to their personal privacy.

Results

The paper acknowledges the need for USN users to be empowered to take informed data disclosure decisions. However, when reviewing prototypes for exploiting USN services that provide means of informed data disclosure decisions, additional crucial usability and privacy limitations were identified. Thus, the paper depicts four drawbacks, suggested to be avoided when designing for privacy in USN. The drawbacks are: (1) ignoring the variation of human data sensitivity, (2) embracing disclosure to third parties, (3) requiring too much user intervention and (4) lacking user's personal data control. Subsequently, the article proposes the design of a privacy-aware USN platform, which is engineered both to comply with the existing relevant privacy guidelines and to avoid the four drawbacks. Finally, the design solutions of the proposed USN platform were evaluated during a qualitative investigation. The majority of participants indicated the proposed design solutions as *must have* and claimed that they would not feel comfortable with sharing their personal data, if the drawbacks were not avoided in the design of USN.

2.4 Paper D

Privacy analysis in mobile social networks: the influential factors for disclosure of personal data

Antonio Sapuppo

Motivations

When users disclose their personal data to others in USN, the shared information is tied to a physical person and immediately available for the recipient [171]. Thus, the data disclosure can be directly translated into physical contact and potentially undesired or unpleasant face-to-face interactions [168]. To address these privacy concerns, privacy management systems should protect users' personal data privacy as individuals do in ordinary human interactions [25, 26, 99]. In fact, during face-to-face communication, people intuitively evaluate various determinants and unconsciously choose what personal information to share. Thus, the factors that might influence users' data disclosure decisions must be depicted and evaluated for enabling privacy management systems to take automated data disclosure decisions.

Methodology

The paper reviews previous qualitative and quantitative investigations in ubiquitous social computing environments for identifying relevant influential factors that might impact users' personal data disclosure decisions in USN. Afterwards, an empirical investigation, comprising two online surveys, was carried out in order to evaluate whether the identified influential factors can be considered relevant for data dis-

closure in USN. Participants of the two surveys were asked to indicate personal information that they would like to share under different circumstances of their lives by compromising between privacy risks and potential benefits. More than 100 responses were collected in each of the two surveys and the Wilcoxon Signed Rank statistical test was applied to examine whether the identified influential factors impact users' personal data disclosure decisions.

Results

The following influential factors were identified: location familiarity, current activity, mutual friends, familiar strangers, purpose of disclosure and access & control. According to the results of this analysis, the purpose of data disclosure was found to be the most important determinant for selecting privacy preferences among the ones tested. As well, having the opportunity to modify personal data, even after actual disclosure, was proven to help users to feel more secure to share their personal information to others. Further, the paper also suggests designers of privacy management systems to consider the other identified influential factors, however as indexes of secondary importance. Particularly, the familiarity with the current location was commonly approved by all the respondents who indicated the tendency to be more open to share their personal data in more familiar locations. This analysis also proved that knowing beforehand information about the inquirer, such as number of mutual friends or previous encounters, relevantly impacted participants' data disclosure decisions. Finally, the activity factor was observed to be significantly influential only on disclosure of data related to work activities, if compared to sharing of data related to social activities.

2.5 Paper E

The influential factors for the variation of data sensitivity in ubiquitous social networking

Antonio Sapuppo

Motivations

This study complements the findings of the previous paper, introduced in Section 2.4, which identified influential factors for data disclosure in USN, based on predefined privacy preferences. In the previous investigation, participants were asked to predict their sharing preferences a priori the actual data disclosure in relation to proposed user scenarios. The results of the investigation, presented in Section 2.4 presented statistically significant results, obtained from a large number of participants. However, this paper acknowledged that there might be a difference between what people say they want to share and what they actually do share in practice when encountering different circumstances [13, 100]. Thus, it was important to analyze if the previously identified factors also impact users' data disclosure decisions made at the moment of actual disclosure as well as to gain an extensive understanding of people' attitudes and motivations that govern such sharing preferences.

Methodology

The paper classified previously identified factors into three different groups, i.e. contextual data, interrelated attributes and design properties. Afterwards, it focused on the first two groups, due to the need for specific analysis of the variation of hu-

man data sensitivity under different circumstances. A sequential two-phase mixed methods study was carried out to explore ad hoc data disclosure preferences in USN. In the first phase, a quantitative research investigated the relationship between the identified influential factors and ad hoc data disclosure decisions. The participants' ad hoc data disclosure decisions were collected by exploiting a USN prototype. The collected data was analyzed by applying the Binary Logistic Regression statistical model for examining whether the selected influential factors could be considered as predictors for data disclosure decisions in USN. Information, acquired during the first phase of the study, was further explored in the second phase, where qualitative interviews were used to gain in-depth understanding of different aspects and motivations of users' data disclosure in USN.

Results

The findings of this study show that the sensitivity of participants' personal data decreases as the relevance of data disclosure for initiation of networking increases. Among the influential factors, the current environment contextual data and purpose of disclosure interrelated attribute primarily guided the participants in evaluation of their data sensitivity and relevance for exploiting USN services. Further, the other contextual data (i.e. current activities, mood and location familiarity) and interrelated attributes (i.e. mutual friends and familiar strangers) influential factors are suggested to be considered as indexes of secondary importance. In fact, these predictors were generally discovered to either refine grained selection of disclosed personal data or provide a feeling of increased comfort as well as motivate curiosity to start an interaction with other users.

2. SUMMARY OF THE PAPERS

3

Results

The literature review, carried out during the first phase of this project, identified three crucial challenges for the development of USN environments, which are following presented:

1. Context acquisition: USN environments must be capable of acquiring the relevant context in order to promote sociability among its participants. Further, an evaluation of the obtained context must be carried out to elaborate its significance and relevancy, which conduces to learning users' behavioral and social patterns for personalizing social networking services;
2. Social networking: USN environments must be capable of enabling social interactions between its participants. Moreover, those services have to be applied not only among acquaintances but also between strangers with interpersonal affinities. Hence it would lead to highlighting relevant social paths between users in the physical world, that would remain hidden otherwise;

3. RESULTS

3. Privacy: USN must be capable of providing a secure and safe collection and dissemination of participants' personal information as well as ensuring an accurate selection of users' personal information to be disclosed to others. This challenge arises due to the fact that the foundation of USN is based on sharing of participants' personal information, which could provoke potential privacy threats.

Among these challenges, the results presented in Appendix A called attention to ensuring users' privacy as the main obstacle for the long-term success of USN. Further, it specifically focused on the management of users' personal privacy, because it assumed a non-malicious infrastructure, targeting at preventing accidental data disclosure, where personal information is unintentionally revealed with or without previous inquiry. Thus, it concentrated on ensuring an accurate selection of users' personal information to be disclosed to others.

In order to manage personal privacy in USN environments, a privacy model called Diverged Personalities (DiP) was designed. In the DiP privacy model, the user's profile is diverged into different user's personalities to be presented under different circumstances. The most suitable personality for each circumstance is generated by the process shown in Figure 3.1. The central component of the DiP is the Personality Logic, which receives as input the unified user profile (UUP) that is composed of a collection of various available user's profiles. Moreover, the Personality Logic also processes the inquirer's social information (e.g. stranger, co-worker, number of mutual friends) and context information (e.g. current activity, location, time). Based on these inputs, the Personality Logic should automatically provide the most suitable personality to be shared with other users. The disclosed

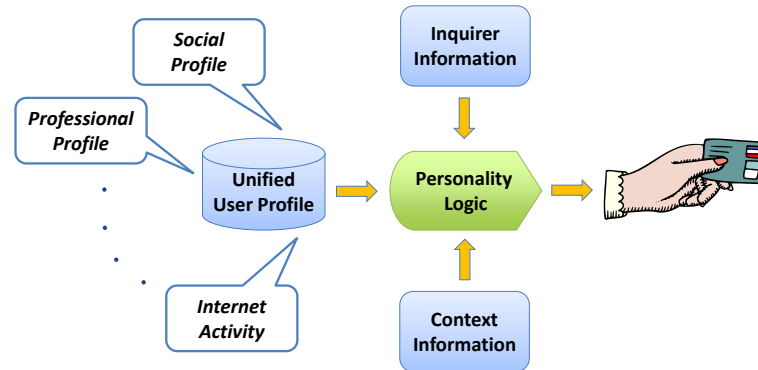


Figure 3.1: Diverged personalities model process

personality can be compared to a business card (BC) that people often exchange for interpersonal benefits and it must be composed of relevant, but not sensitive, personal information for the encountered circumstances.

When taking into consideration the DiP privacy model, it is possible to identify that the management of personal privacy occurs in two different moments of the users' engagement with USN environments. Firstly, it regards whether users' personal data should or should not be included in the UUP. Secondly, the management of personal privacy is also related to the selection of information that should be included into users' sub-profiles (or personalities) to be disclosed to other end users under the different circumstances of their encounters. In the following sections, this dissertation presents the results obtained during the empirical investigations of this work, which are related to the two directions for managing users' personal privacy in USN, followed by a discussion about the investigation limitations for better interpreting the outcomes of this dissertation.

3. RESULTS

3.1 Privacy design guidelines

As introduced in Section 1.2, the first direction to manage personal privacy in USN environments is related to data disclosure decisions that are principally made before the actual use of the system. Such data disclosure decisions would be highly influenced by designing USN services that would increase users' comfort with sharing of their personal data in such environments.

During Phase 2 of this project, it was firstly analyzed whether users would accept potential privacy risks and thus be inclined to share their personal information in USN environments. Results of this qualitative investigation, presented in Appendix B, proved that users were inclined to accept potential privacy risks, because participants perceived USN to be useful for providing valuable networking benefits in exchange to their data disclosure. All the participants of this investigation acknowledged the potential of USN services for improving their everyday communication, as they could foresee that these services would help them to connect with others nearby, who have matching personal and professional interests and goals. Participants claimed that they would be potential users of USN services, as long as such services would respect end users privacy and minimize unintentional data disclosure, which might lead to possible unpleasant face-to-face interactions.

The qualitative investigation regarding the perceived usefulness of USN environments highlighted the need for the development of privacy-aware USN environments. This dissertation agrees with the position of Bellotti and Sellen [12], who suggested that privacy should be a central design issue in its own right for such kind of invasive technologies. Thus, it was crucial to discover, understand and satisfy the

users' privacy needs related to the acceptance of the USN technology for minimizing potential invasions of users' personal privacy.

Privacy needs for ensuring acceptance of ubiquitous social networking

The qualitative investigation, presented in Appendix B, further focused on discovering the privacy needs of the participants for ensuring acceptance of the prerequisites for the establishment of USN services. Three prerequisite were identified as follows:

1. Announcement of users' presence: users must accept to inform others nearby about their whereabouts;
2. Disclosure of personal data: users must accept to share their personal information with others;
3. Potential initiation of face-to-face interactions: users must accept possible immediate face-to-face interactions with other users, when notified about potential profile similarities.

Table 3.1 illustrates the identified users' privacy needs for ensuring the acceptance of the three prerequisites for the establishment of USN services. The first requirement, related to the announcement of users' presence, presented participants'

Table 3.1: Participants' privacy needs for the establishment of ubiquitous social networking services

| Requirement | Privacy needs |
|---|--|
| Announcement of users' presence | Only to end users |
| Disclosure of personal data | Individual participation; Sharing different profiles; Ad hoc privacy control |
| Potential initiation of face-to-face interactions | Coarse-grained control |

3. RESULTS

concerns, provoked by unnecessary disclosure of their current positions. They discussed location based services as a negative example - when exploiting these services, users must also share their current position with third party entities, which then will notify them when their friends are nearby [206]. Instead, participants commonly preferred to disclose this information only to other end users, who are located in their close proximity.

Further, in relation to the second requirement, i.e. disclosure of personal data, participants accepted to disclose their personal information in order to gain potential networking benefits in exchange. However, they expressed relevant concerns about loss of permanent control over their data after actual disclosure. They claimed to be more comfortable with sharing their personal data, if these services would respect an important privacy principle, noted in the Fair Information Practices, called individual participation. This privacy protection principle is essential for personal data disclosure and it is already incorporated into all major privacy laws worldwide, such as [62, 82]. The individual participation regards the right of the user to always be able to see and correct any data disclosure decisions. Further, participants also appreciated the possibility to customize their sharing preferences according to different circumstances, and thus rejected the option of disclosing a static user profile in all the situations, often applied in the majority of current USN applications and prototypes, such as [10, 65, 153, 171, 183]. In order to share different profiles under different circumstances, they commonly favored ad hoc privacy control over the predefined privacy preferences. This choice was motivated by reasons, similar to the ones already introduced in Section 1.2.2, which regarded the outcomes of a qualitative investigation, carried out by Lederer et al [122] and reviewed in Appendix A. In

fact, after exploiting the provided mobile prototype simulating the USN behavior, participants discussed that it would be very difficult for them to define in advance what to disclose per each circumstance, as they expected to encounter situations where data disclosure decisions would not be accurately predictable in advance.

Lastly, analysis of qualitative data, related to the third requirement about potential initiation of face-to-face interactions with others, highlighted the need for users to have a coarse-grained control over the USN services. This principle was already introduced in the privacy guidelines of ubicomp, reviewed in Appendix A, which suggested that users should always have a binary choice for halting and resuming participation in such environments. A few participants, however, refused their potential participation in USN environments, due to serious concerns arising in regard to this requirement. In fact, they were worried that someone would unnecessarily disturb them, just because of the information that they had shared. They suggested reconsidering their potential participation in USN, in case these services would enable an invisible mode option and disclose their information only after user's approval, which implied manual evaluations of the trade-offs between potential networking benefits and privacy risks.

The invisible mode option is a relevant design characteristic, already applied in some services, such as online instant messaging, which are typically exploited within circles of acquaintances. For example, users of Microsoft Windows Live Messenger can enter an invisible mode that allows them to discover other online users, without revealing their actual presence. Afterwards, they must change their current status from invisible to online for initiating an interaction with any other online user. The change of visibility status occurs after an evaluation of trade-offs between desired

3. RESULTS

communication and announcement of their availability to all the users. However, such evaluation of trade-offs in USN would present increased complexity and require too much user's attention and intervention, due to focus on initiation of relationship between strangers. Thus, application of this option would not lead to a calm USN technology, where users could effortlessly exploit these services [193].

Privacy guidelines for ubiquitous social networking The identified privacy needs, illustrated in Table 3.1, provided the input for designing more functional and privacy-oriented USN environments where users can effortlessly exploit USN services and feel more comfortable with sharing of their personal information. In Appendix C, it was analyzed whether already existing USN applications, such as [10, 65, 153, 171, 183], would satisfy the discovered privacy needs as well as respect existing privacy guidelines for managing personal privacy in ubicomp environments. These guidelines, proposed by Lederer et al [122], were already introduced in Section 1.2.1 and they are summarized in Table 3.2. They describe a set of pitfalls that should be avoided for empowering users of ubicomp to take informed data disclosure decisions. In order to achieve this target, these guidelines focus on helping people to gain relevant understanding and support actions that are needed in order to manage personal privacy in such environments. In fact, by heeding the first two pitfalls, these environments would enable users to understand privacy implications of their data disclosure. Furthermore, users are allowed to conduct socially meaningful actions through the system, if the last three pitfalls are avoided [122].

As a result, even when the reviewed mobile prototypes, such as [65], provide means for taking informed data disclosure decisions by respecting the privacy guide-

3.1 Privacy design guidelines

Table 3.2: Five pitfalls to be avoided in the design of ubiquitous computing [122]

| Pitfall | | Description |
|---------|---------------------------------------|---|
| 1 | Obscuring potential information flow | Ubicomp should not obscure the nature and extent of data disclosure. Users should easily comprehend, for example, what kind of information is disclosed and to whom, how the information is shared, the presence of third-party observers and the potential for unintentional disclosure. |
| 2 | Obscuring actual information flow | Ubicomp should not obscure the actual disclosure of information. The disclosure should be obvious to the user as it occurs, however without overwhelming his attention. When immediate notice is not feasible, then it must be ensured with a reasonable delay. |
| 3 | Emphasizing configuration over action | Ubicomp should not require exaggerated manual configuration to manage personal privacy. Instead, users' privacy should be managed as a natural consequence of their normal engagement with the environments. |
| 4 | Lacking coarse-grained control | Ubicomp should not forgo a binary choice for halting and resuming data disclosure. |
| 5 | Inhibiting established practice | Ubicomp should not inhibit users from transferring established social practice to emerging technologies. For example, ubicomp should enable disclosure of ambiguous information as well as ensure plausible deniability. |

lines illustrated in Table 3.2, many other crucial usability and privacy limitations related to the discovered users' privacy needs, shown in Table 3.1, were identified to lead to potential invasions of users' personal privacy. In fact, the existing guidelines, proposed by Lederer et al [122], do not consider that users were found to be averse to disclose their personal information to third parties [6, 84, 91, 130, 168], as they were concerned about potential future implications.

Secondly, the disclosure of different profiles under diverse circumstances is not considered in these guidelines, despite the fact that the human data sensitivity was found to vary upon different situations [45, 123, 167, 169]. Thirdly, the aforementioned guidelines also lack attention to the individual participation privacy principle.

3. RESULTS

Having the possibility to keep control over their personal data, even after actual disclosure, was discovered to influence users' sharing preferences [167, 168].

Lastly, with respect to usability, users might be displeased when too much user intervention is required for the accomplishment of USN. As earlier discussed, design features for minimizing potential unpleasant face-to-face interactions like invisible mode option should not be adopted in USN environments. Instead, design solutions for addressing these users' concerns must be identified by targeting at the development of a calm technology [168]. Furthermore, in Appendix C, other design characteristics that would require a considerable amount of users' intervention, were discovered in the designs of Nokia Sensor [153] and Bluedating [10], and consequently advised to be avoided. First, they refer to requiring the users to manually invoke a Bluetooth discovery for finding others nearby as well as to carry out manual profile comparisons for evaluating whether the encountered users present relevant profile similarities. Second, such design characteristics refer to only informing users about potential profile similarities with others without initiating any connections between the users, which would additionally require considerable user's intervention especially in crowded places.

For the reasons presented above, this dissertation seeks the necessity to update the current privacy guidelines by introducing four additional drawbacks that should be avoided when designing for privacy in USN environments. The proposed guidelines, shown in Table 3.3, aim at designing functional USN services that respect the privacy of end users and help them to feel more comfortable with data disclosure in such environments targeting at maximizing potential networking benefits while minimizing users' privacy concerns. In fact, when avoiding Drawback 2 and Draw-

3.1 Privacy design guidelines

Table 3.3: Four drawbacks to be avoided when designing for privacy in ubiquitous social networking

| Drawback | | Description |
|----------|--|---|
| 1 | Ignoring the variation of human data sensitivity | USN should not disclose personal information without taking into consideration the human data sensitivity of the current circumstances. Instead, different sets of personal information should be disclosed upon different circumstances. |
| 2 | Embracing disclosure to third parties | USN should avoid disclosure of users' personal information to third-party entities. Contrarily, disclosure should occur towards other users, whose profiles might lead to potential mutual interests and networking benefits. |
| 3 | Requiring too much user intervention | USN should not require too much user intervention. Connections between users must be created with minimal efforts of end users, thus allowing technologies to operate seamlessly in the background. |
| 4 | Lacking user's personal data control | USN should not lead to loss of permanent control over personal data. On the contrary, users must always have opportunities to modify any piece of information even after actual data disclosure. If desired, the updated data should be effortlessly synchronised to all relevant peers, who have permission to access to it. |

back 4 users' privacy concerns are reduced, because their personal data becomes available only for other end users, who present relevant profile similarities, and with the possibility to be modified when desired. The latter might also increase potential networking benefits to even a greater extent by allowing users' information to be continuously updated. Moreover, heeding Drawback 3 would embrace Weiser's vision of a calm technology and thus allow users to effortlessly exploit USN services [191, 193]. When avoiding Drawback 1, the design of USN would allow to prevent potential invasions of users' privacy as well as to motivate users to disclose their personal information considered as too sensitive to be shared in some

3. RESULTS

Table 3.4: Interdependencies between the identified drawbacks

| Interdependencies | | Description |
|-------------------|--|---|
| 1 | Ignoring the variation of human data sensitivity Requiring too much user intervention | In order to evaluate which data should be shared under the current circumstances, USN users might be requested to take data disclosure decisions at the moment of actual disclosure. However, such kind of approach would indisputably need a considerable amount of users' attention and intervention. |
| 2 | Embracing disclosure to third parties Lacking user's personal data control | A decentralized approach might disclose users' personal data only to other end users, but it would lack control of users' personal data after actual disclosure. Contrarily, a centralized architecture might be ideal for enabling users to modify at any time their personal data, but it would certainly require disclosure to third-party components. |

circumstances⁷.

Finally, privacy designers of USN are as well advised to carefully evaluate how to avoid these drawbacks, because successfully heeding one drawback might result in the risk of falling into another. They should find solutions for avoiding the proposed drawbacks by also taking into consideration the interdependencies between them illustrated in Table 3.4. When that is not possible, designers are challenged to find solutions that would represent an optimal trade-off between the risks that the drawbacks might impose.

In order to investigate whether the perception of users' personal privacy is enhanced when the design of USN avoids the four identified drawbacks, this dissertation proposes the design of a privacy-aware USN platform. Afterwards, it presents

⁷The reader should note the difference between Drawback 1 and ambiguous data disclosure, included in Pitfall 5 illustrated in Table 3.2. For example, a design solution that discovers users with similar interests and discloses to them only their related affinities, provides means of ambiguous data disclosure, because it enables dynamic sharing of profiles that varies depending on the different similarities between the encountering users. However, such design falls into Drawback 1, as it does not take into consideration the variation of human data sensitivity, according to the different circumstances. More details are provided in the individual paper in Appendix C.

results of a qualitative investigation that analyzed whether participants feel more comfortable with sharing of their personal information when the four identified drawbacks were avoided in the design of USN services.

The design of a privacy-aware social networking platform The proposed privacy-aware social networking (PAUSN) is designed to overcome both privacy pitfalls and drawbacks. In order to comply with the existing and proposed privacy guidelines, PAUSN utilizes a third-party entity that receives encrypted (thus incomprehensible) profiles from the encountering users in order to calculate the profile similarities. The third-party is capable of comparing the encrypted profiles and computing similarities between the two users, as profiles are ciphered with the same security key⁸. When the similarity scores, defined by the users, exceed their corresponding threshold values, users are notified and their personal information is disclosed to each other, however it still remains incomprehensible for the third-party. At any time, PAUSN users are empowered to modify their data disclosure decisions, thanks to its centralized architecture.

⁸The PAUSN design adopts an asymmetric cryptography method, which utilizes two different keys, referred as public and private keys. The public key is used for encrypting data and it can be sent to anyone. Contrarily, the private key is used for decrypting the data and it is never revealed to another party. However, this is not the only cryptography method that can be adopted in PAUSN. Probably, there are more efficient and secure cryptography methods that can be taken into consideration in the design of PAUSN, e.g. Shared Secret Key and Private Set Intersection. For example, a Shared Secret Key can be established by adopting a key agreement protocol, such as a variant of Elliptic Curve Diffie Hellman (ECDH) [88]. The shared key can be subsequently applied to encrypt the communication messages. These messages can be then decrypted by the other parties, utilizing their private key. The Private Set Intersection, instead, would allow the two users to identify their profile similarities, based on their inputs that they exchange in a peer-to-peer way [58]. Consequently, this cryptography method would present the advantage to at least reduce the interaction between the users and the third party component for the identification of users' similarities. Advantages and disadvantages of these cryptography methods are suggested to be investigated when analyzing PAUSN with respect to security.

3. RESULTS

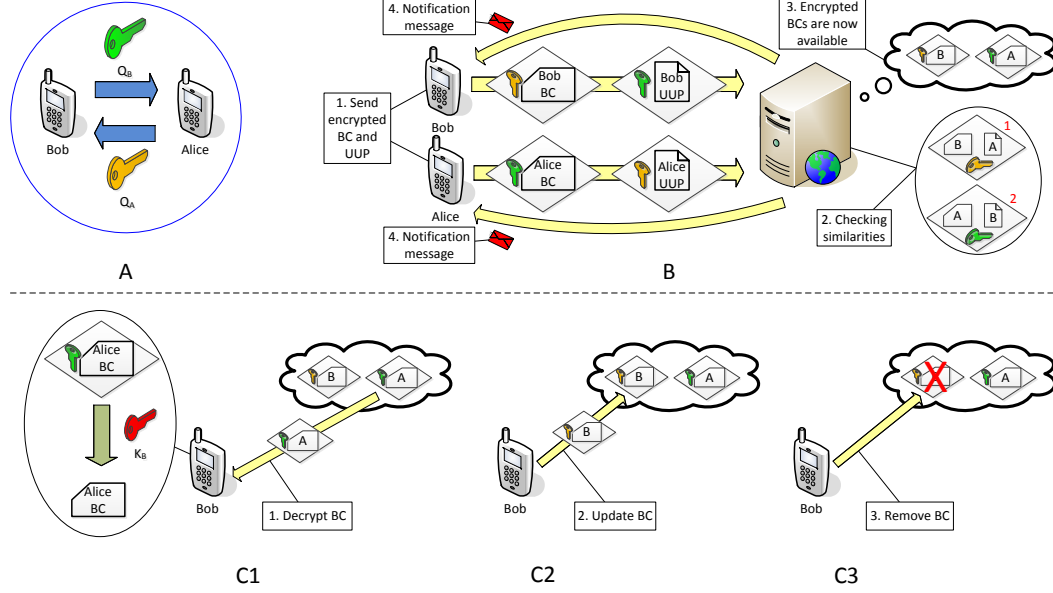


Figure 3.2: PAUSN design

In Figure 3.2, an example scenario of the establishment of USN services between two users, i.e. Bob and Alice, is illustrated to better describe the design characteristics of PAUSN. As shown in Figure 3.2-A, when Bob and Alice enter each other's wireless range, they exchange their public keys through direct ad hoc links, i.e. Bob sends to Alice his public key (Q_B) and Alice sends to Bob her public key (Q_A). Afterwards, in Figure 3.2-B, Bob encrypts his UUP using his Q_B and his BC using Alice's public key, i.e. Q_A . Accordingly, Alice encrypts her UUP using her public key, i.e. Q_A and her BC using Bob's public key, i.e. Q_B ⁹. Finally, the encrypted profiles are submitted to a third-party (step 1) using their own broadband infras-

⁹The UUP of Bob and Alice is composed of a collection of various available users' profiles (e.g. social profile, work profile), while their BC is a subset of the UUP, which is created by the DiP privacy model, shown in Figure 3.1. As earlier introduced, the users' BC is composed of personal data, considered to be relevant, but not sensitive, for the current circumstances of their encounter.

structure link, i.e. Internet connection. Even if the third-party is not able to access the contents of the encrypted profiles, it checks for similarities by comparing Bob's UUP with Alice's BC, as both profiles are encrypted with the same Bob's public key, i.e. Q_B (see Section C.5.3 for details). The third-party also compares Bob's BC with Alice's UUP, both encrypted with the Alice's public key, i.e. Q_A (step 2). If no match is found, the third-party deletes the received encrypted profiles. Otherwise, if similarities are discovered, the third-party deletes the users' UUPs and stores only the encrypted BCs of Bob and Alice (step 3)¹⁰. Afterwards, the third-party sends a notification to both users (step 4).

As illustrated in Figure 3.2-C1, Bob is now able to access Alice's BC from the third-party entity and decrypt it by using his private key, i.e. K_B , as Alice's BC was previously encrypted using Bob's public key, i.e. Q_B . Even if not shown, Alice is also able to access and decrypt Bob's BC using her private key. Moreover, the users are always able to modify or revoke their data disclosure decisions. For example, if Bob is interested to networking with Alice, he can modify his BC by encrypting his new BC with Alice's public key (i.e. Q_A) and replacing it on the server, as shown in Figure 3.2-C2. The contents of the new Bob's BC would still be inaccessible for the third-party and available only for Alice. Finally, as shown in Figure 3.2-C3, Bob has the opportunity to revoke his BC. In this case, any information about Bob will not be available anymore, Alice would be optionally notified and the link between the two users will be deleted. When desired, any piece of information that

¹⁰The strategy to disclose users' BCs, after identifying relevant users' similarities between the UUP of a user with the BC of the other (and vice versa) was adopted in PAUSN, because it was found to increase potential networking benefits, while decreasing accidental invasions of privacy. More details about the advantages and disadvantages of the selected strategy for comparing users' profiles in PAUSN are provided in Section C.5.1.

3. RESULTS

is updated or removed by Bob can be effortlessly synchronized not only with Alice, but also with all the other users, who have access to that specific information.

Analysis and evaluation of PAUSN design solution Table 3.5 presents an analysis of the PAUSN design solutions for avoiding the five pitfalls, while Table 3.6 introduces the PAUSN design characteristics for overcoming the four identified drawbacks.

The PAUSN design also takes into account the identified interdependencies between the drawbacks. Firstly, applications of ad hoc privacy control systems should provide an acceptable balance in relation to the interdependency between Drawback 1 and Drawback 3. In fact, PAUSN takes into account the variation of human

Table 3.5: PAUSN design solutions for avoiding the five pitfalls

| Pitfall | Design solution |
|---------|---|
| 1 | PAUSN deliberately constrains the potential information flow to intentional disclosure of personal data between users with interpersonal affinities. |
| 2 | PAUSN makes the actual information flow evident through the notifications from the third-party entity. Moreover, users are also aware about who can access their profiles, because they are stored by the third party component, after discovery of relevant profile similarities. |
| 3 | PAUSN requires minimum configuration to manage users' personal privacy, due to possible application of the Diverged Personalities privacy model that discloses personal information on users' behalf. |
| 4 | PAUSN provides a coarse-grained control for halting and resuming information flow, thanks to application of mobile devices' exit and power buttons. |
| 5 | PAUSN supports ambiguous data disclosure and plausible deniability. The former is ensured, because PAUSN does not disclose static detailed users' profiles, as application of ad hoc privacy control allows arbitrary customization of disclosed users' personal data. The latter is also supported, because a user never knows the true reasons why another user discloses a specific subset of his UUP in detriment of other pieces of information. Reasons might consist of different subjective levels of the users' privacy perceptions on different environments, time circumstances or simply the desire to be left alone. |

3.1 Privacy design guidelines

Table 3.6: PAUSN design solutions for avoiding the four drawbacks

| Drawback | Design solution |
|----------|--|
| 1 | PAUSN does not ignore variation of human data sensitivity, because it provides opportunity to customize users' profiles according to different circumstances, due to adoption of the Diverged Personalities privacy model. |
| 2 | PAUSN does not disclose any users' personal information to the third-party entities. Instead, users send encrypted, thus incomprehensible, profiles to third parties. |
| 3 | PAUSN does not require too much users' intervention, as it automates the process of finding users' profile similarities. When a match between users is found, PAUSN automatically notifies users and establishes connections between them, by storing users' profiles in third party entities. |
| 4 | PAUSN does not lack user's personal data control, as it follows the individual participation principle by allowing users to see, modify and, if desired synchronise, the contents of their disclosed profile, even after actual disclosure. |

data sensitivity (avoiding Drawback 1), while relieving the users from frequent data disclosure decisions (avoiding Drawback 3). This is achieved by applying the Diverged Personalities privacy model that manages personal data privacy on the users' behalf. However, some users' intervention might be expected in some situations. For example, users should always be empowered to adjust their data disclosure, in case of disagreement with the automated sharing decisions. Furthermore, if highly sensitive data is inquired to be shared, the ad hoc privacy control should limit its autonomy and require users' approval before any actual disclosure.

Secondly, PAUSN relies on a centralized architecture that stores only encrypted profiles, which should also provide an acceptable balance in relation to the interdependency between Drawback 2 and Drawback 4. In fact, the proposed platform only shares personal information with end users (avoiding Drawback 2) and allows personal data control, even after actual disclosure (avoiding Drawback 4). However,

3. RESULTS

this design solution might present challenges for discovering users' with similar interests and goals. PAUSN might limit discovery of profile similarities, because it might be challenging to compare complex preferences between the encountering users, due to profile comparison on the encrypted domain. As an example, it might be difficult to apply sophisticated profile matching methods, usually utilized in online dating sites.

The above design solutions for avoiding the identified drawbacks and addressing their interdependencies were taken into account during the qualitative investigation, presented in Appendix C. This study targeted at evaluating users' perceptions towards management of personal privacy in relation to the proposed privacy guidelines. Participants indicated that they would tend not to disclose their personal information, if the identified drawbacks were not avoided in the design of USN. The most consistent results were observed in regard to Drawback 3, as participants acknowledged that it is crucially important to be able to effortlessly exploit USN services. In fact, none of the participants would disclose their personal data in USN, in case these services would require a considerable amount of user intervention, such as manual profile comparison.

Contradictory outcomes were found in regard to embracing disclosure to third parties (Drawback 2). The majority of participants appreciated the possibility not to share their personal data with third parties. For example, some of the participants considered their political views as not sensitive to be disclosed to other end users, who share the same political preferences. However, they claimed to reconsider such decision in case it was necessary to share their political views also with a third party, due to concerns about potential negative future implications.

On the other hand, it is also important to note that a few of the participants did not perceive any advantages of not disclosing their personal information to third parties, as they did not expect any possible negative future privacy implications.

Further, participants acknowledged the relevance of the centralized USN architecture, which would enable them to keep control over their data even after disclosure (Drawback 4). These results were also supported by findings of the quantitative investigation, included in Appendix D. When it was emphasized to the participants that the design of USN services would provide control over their data, even after disclosure, it was statistically proven that participants shared a larger amount of their personal data, in comparison to situations where such control was not provided.

Finally, participants also appreciated that USN would support application of ad hoc privacy control for customizing sharing preferences according to the current circumstances of the users' encounters (Drawback 1). In fact, without such option, participants argued that they would not include in their unified user profile some personal data (e.g. sexual orientation), which is considered to be sensitive in some circumstances (e.g. work environments), even if it is preferred to be shared in other situations (e.g. social environments).

In relation to the first interdependency between Drawback 1 and Drawback 3, participants were introduced to the possibility for users to be required to intervene in case of (i) disagreement between the users' actual preferences and automated data disclosure decisions and (ii) needed approval for disclosure of highly sensitive data. Participants agreed with the previous results of the qualitative investigation, included in Appendix B, and thus highlighted serious concerns, if they were

3. RESULTS

not allowed to interfere with automated data disclosure decisions, taken by the Diverged Personalities privacy model. Further, in relation to the second interdependency between Drawback 2 and Drawback 4, participants were introduced to the disadvantages of discovering less profile similarities with other users, due to profile comparison on the encrypted domain. This option would allow them not to share their personal information with third parties while keeping control over their data, even after disclosure. The participants, who preferred not to disclose personal information to third parties, indicated that they would accept to compromise the number of discovered users with relevant similarities, if the opposite would mean jeopardizing their personal privacy.

3.2 The influential factors for data disclosure

As introduced in Section 1.2, the second direction to manage personal privacy in USN environments is related to data disclosure decisions that occur during users' ad hoc meetings. In this case, people's data disclosure attitudes are highly situational and present continuous negotiation of the privacy boundaries at the moment of the actual disclosure [4, 5, 100, 149]. Thus, the third phase of this project focused on identifying and investigating the determinants for the selection of personal data to be shared with other end users.

Accurate selection of the most suitable personality to be disclosed for the current circumstances of the users' encounters is significantly important for maximizing potential networking benefits while minimizing unintended data disclosure decisions that might provoke serious users' privacy concerns. In order to achieve this goal,

3.2 The influential factors for data disclosure

the Personality Logic of the DiP privacy model (Figure 3.1) takes into consideration relevant influential factors, found to impact users' data disclosure decisions in USN. These influential factors were identified during an empirical investigation, presented in Appendix D. Afterwards, they were subsequently categorized and analyzed through a mixed methods study, included in Appendix E. As shown in Table 3.7, the identified influential factors determining the variation of human data sensitivity, were divided into two main categories¹¹: contextual information and interrelated attributes. The first group relates to influential factors regarding the current contextual circumstances of the users' encounters in USN environments, e.g. where is the user, what is he doing, etc. The second group of influential factors consists of information regarding what the user has in common with the inquirer, e.g. similar music preferences or number of mutual friends.

When analyzing the results of the investigations, included in Appendix D and in Appendix E, it was discovered that the disclosed personal information was selected by compromising between perceptions of data sensitivity for the current circumstances and evaluations of data relevance for gaining potential networking benefits. Participants' data sensitivity was found to decrease as the relevance of information disclosure for initiation of networking increases. Furthermore, the identified influential factors were discovered to have different influence on users' personal data

¹¹Appendix E as well identifies a third category of influential factors, i.e. design properties. This group of determinants corresponds to design solutions that should be taken into consideration when implementing USN services, e.g. not disclosing users' personal data to third parties. The data analysis, presented in Appendix E, does not include such determinants and focuses on contextual data and interrelated attributes influential factors, because it targets at in-depth analysis of the variation of human data sensitivity under different circumstances. In fact, as earlier mentioned in this dissertation, the design properties were found to influence overall data disclosure, rather than shaping data sensitivity under different circumstances [167, 168].

3. RESULTS

Table 3.7: The influential factors for data disclosure in ubiquitous social networking

| Category | Factor | Description |
|-------------------------|-----------------------|--|
| Contextual data | Environment | It is considered to be the current location of the users, grouped according to their ordinary activities in that location, e.g. work environments, social environments. |
| | Location familiarity | It is considered to be the users' familiarity with their current location, evaluated according to the amount of time that users usually spend in a specific location, e.g. daily, monthly, first time in this location. |
| | Activity | It is considered to be the current action of the user, e.g. working, relaxing. |
| | Mood | It is considered to be the users' current status of emotion, e.g. depressed, happy, sad, angry. |
| Interrelated attributes | Familiar strangers | It is considered to be the number of times that the users have already encountered the inquirer, e.g. 120 times in the last 3 weeks, etc. Notably, encountering does not necessarily imply interaction - they may have just passed by each other without noticing. |
| | Mutual friends | It is considered to be the number of mutual friends that the users have with the inquirer, e.g. 6 common friends. |
| | Purpose of disclosure | It is considered to be the reason why specific personal information is disclosed, e.g. potential networking benefits are foreseen, because users have matching interests or career abilities and expectations. |

disclosure decisions.

The current environment contextual data was found to be a crucial determinant for data disclosure, because it primarily guided the users in the evaluation of their data sensitivity and relevance for exploiting USN services. Similarly to the current environment, the purpose of disclosure interrelated attribute as well significantly guided the participants in taking their data disclosure decisions. Moreover, the purpose of disclosure was discovered to alter participants' data disclosure decisions, previously based on the contextual data influential factors (e.g. current environment, activity), when significant potential networking benefits could be clearly

3.2 The influential factors for data disclosure

foreseen.

Following the results of these investigations, it is also suggested to take into consideration the other two contextual data influential factors, i.e. current activities and location familiarity, however as indexes of secondary importance, if compared to the current environment. In fact, these two influential factors motivated participants to refine grained selection of disclosed personal information rather than being primary predictors for personal data disclosure. The current activity refined data disclosure decisions of the majority of the participants, while the location familiarity presented contradicting results where only a few of the participants were influenced. The last contextual data influential factor, i.e. mood, was found to impact users' participation in USN environments, rather than shaping the actual data sensitivity. Thus, it was suggested to utilize information about users' current mood as a trigger to interrupt their participation in USN environments.

Among the interrelated attributes, the numbers of mutual friends and previous encounters (i.e. familiar strangers) can be considered as relevant predictors for data disclosure in USN, because they were found to be statistically significant during the quantitative investigations. However, these determinants were proposed to be considered as indexes of secondary importance when compared to the purpose of disclosure. In fact, they provided a feeling of increased comfort with data disclosure as well as motivated curiosity to start an interaction with other users, rather than guiding the participants in the evaluation of data sensitivity and relevance for gaining networking benefits. Notably, the impact of familiar strangers influential factor was found to increase directly proportionally to the rising number of previous meetings, while such inclination was not observed in regard to the number of

3. RESULTS

mutual friends, i.e. only a slight difference in data disclosure was presented between varying numbers of mutual friends.

When taking into consideration the influential factors, listed in Table 3.7, significant prediction results were obtained by applying the logistic binary regression statistical model on participants' ad hoc data disclosure decisions. The overall prediction results presented an approximate accuracy of 90% and peaks of 93%, with potential for further increasing performance. Despite these good results, some of the respondents expressed their desire to limit the autonomy of the Diverged Personalities in some situations. Participants claimed that they would be uncomfortable with allowing the Personality Logic to take decisions on their behalf in case of either inquiry for highly sensitive personal data (e.g. political views) or some specific circumstances that are very important (e.g. attending a job interview). These concerns confirmed the findings of previous studies [24, 25, 27], where the authors advised to provide only suggested data disclosure choices while waiting for user's approval before any actual disclosure, in case of inquiry for highly sensitive data. However, even if such option can be considered very useful in relation to ad hoc privacy control, it is still recommended to be kept at minimum in the design of USN, in order not to require too much users' attention and intervention.

When comparing the results presented in this section to previous analysis of data disclosure predictors in other research settings, it can be noted that some of the influential factors, introduced in Table 3.7, were as well found to impact users' data disclosure decisions in ubiquitous social computing and online social networks. For example, in [45] Consolvo et al investigated the willingness to disclose users' current position, based on different granularity of users' location (e.g. specific, vague)

3.2 The influential factors for data disclosure

to other known people (e.g. friends, co-workers). Even if defining the identity of the inquirer as a crucial parameter for their information disclosure, their results confirmed that knowing the particular reason for data disclosure also significantly motivated users to share their current position to others. Similarly, the authors also discussed that users differentiated their data disclosure decisions upon different activities and mood. Participants were more inclined to share their personal information in some activities (e.g. exercising) rather than others, e.g. studying, as well as they were most willing to disclose their personal data when depressed, in contrast to being angry. Finally, the results related to the mutual friends influential factor were also confirmed in analysis of data disclosure decisions in online social networks, where users were proven to be much more likely to disclose their personal data to strangers, if they had at least one friend in common [145].

In relation to the different privacy clusters (i.e. fundamentalists, pragmatists and unconcerned), the results of the empirical investigations [167, 170], carried out during this project, show a relation between users' personal privacy preferences in online social networks and in USN. Fundamentalists users were found to share less personal information than pragmatists and unconcerned users. Further, unconcerned users commonly disclosed an higher amount of personal information when compared to pragmatists. These results contrast with the findings of Consolvo et al [45], where the participants' privacy classification was not found to be as a relevant factor of users' data disclosure rate in ubiquitous social computing environments. This can be explained for the following reason: the authors measured attitudes about disclosure of their current location to acquaintances, which completely differs from sharing personal data, such as home address and personal phone number,

3. RESULTS

to strangers in USN environments. This identified relationship between users' data disclosure decisions in online social networks and USN might also provide relevant input for the design of privacy management systems of USN environments. For example, it was discovered that the pragmatists users were overall the most affected by the influential factors, in comparison to the other two privacy clusters. Furthermore, the fundamentalist privacy cluster was the most influenced by the purpose of disclosure determinant, even if they were only slightly impacted by the other factors. The unconcerned privacy users, instead, were less impacted on differentiating their sharing preferences according to different circumstances, because of their inclination to share large sets of personal data in all the circumstances. Unfortunately, it was not possible to further investigate such relation between users' data disclosure decisions in online social networks and USN in the mixed methods analysis, included in Appendix E, due to the restricted amount of recruited users, i.e. 13.

3.3 Investigation limitations

The results presented in this dissertation pose the question regarding the validity and reliability of the analysis. This is related to different limitations of the empirical investigations, carried out during this project. In the following, these limitations are listed in order to help the reader for better interpreting the results, presented in this study.

In relation to the qualitative investigations, it was acknowledged that a larger number of participants might lead to a more reliable analysis. Unfortunately, this

3.3 Investigation limitations

was not possible due to limited resources. As introduced in Section 1.3.1, this dissertation attempted to address this issue, by following the suggestions of previous studies that advised to recruit a number between 6 and 20 of lead users for the evaluation of new emerging technologies, as this type of users were discovered to provide as much information and ideas as five *ordinary* users did [115].

Furthermore, the outcomes obtained through analysis of face-to-face interviews might have been biased by the the presence of the interviewer. This is a well known problem, already discussed in previous literature, and identified to be difficult to be avoided for such kind of investigations [163].

The results presented in Appendix C, which regard selection of users' personal data that should or should not be included in the unified user profile are based on predefined sharing preferences: participants were asked to predict their sharing preference in relation to the proposed user scenarios. This dissertation acknowledged that there might be a difference between what people say they want to share and what they actually do share in practice [100]. Similar limitations can also be identified in the results, presented in Appendix D, where users selected different subsets of their unified user profile to be disclosed to other end users according to different circumstances, presented as user scenarios. In this case, the results of this quantitative investigation were complemented with an analysis of users' ad hoc data disclosure decisions, i.e. selection of sharing preferences made at the moment of actual disclosure, presented in Appendix E. However, it must be noted that the collection of this data was not taking into consideration subsequent potential face-to-face interactions, arising from the sharing of users' personal data. It might be possible that after adopting the USN services, users would have decided to differen-

3. RESULTS

tiate their data disclosure decisions, selected at the moment of acceptance of USN, after experience the first face-to-face interactions with other users.

In both cases presented above, these studies would have benefited from an implementation of the PAUSN application in order to investigate users' actual behaviours in USN environments. Unfortunately, due to the limited resources and number of participants, this approach was not feasible at the time of this project, because it would have required a fast adoption of the PAUSN in order to allow the participants with the possibility to disclose their personal information to other PAUSN users in their everyday lives.

Finally, the restricted number of participants in the mixed methods, presented in Appendix E limited further analysis on the relation between users' privacy attitudes about data disclosure in online social networks and USN. Although the collected ad hoc data disclosure decisions were quantitatively analyzed based on the number of data disclosure decisions per each of the respondents, this dissertation acknowledged that a larger number of participants would have also increased generalizability of the findings.

4

Conclusions

The key problem that this dissertation addresses is the difficulty of promoting privacy-aware social networking in ubiquitous computing environments that focus on maximizing potential networking benefits while minimizing users' privacy concerns. In order to contribute to the further development of these environments, this dissertation focused on the management of personal privacy - the processes by which people selectively share personal information, e.g. interests, phone number, career skills and expectations, to organizations and to other people [98]. This focus was motivated by the fact that data disclosure in ubiquitous social networking cannot be considered as a straightforward concept and consequently modeled using rigid privacy policies. Instead, the intrinsic negotiation about data disclosure must be constant and arbitrary in order to allow users to share personal data, customized according to the current circumstances. This dissertation contributes to this problem, by taking into consideration two different directions for ensuring the management of users' personal privacy. They refer to the design of privacy-aware ubiquitous

4. CONCLUSIONS

social networking for enabling more functional and privacy-oriented services and further identification of the determinants for evaluating the human data sensitivity, according to the different circumstances, in order to select user' personal data to be disclosed in such environments.

The first direction to manage users' personal information in ubiquitous social networking is related to the challenge of designing ubiquitous computing environments that would increase users' comfort with sharing of their personal information for effortlessly maximizing potential networking benefits while preserving users' personal privacy. To address this challenge, this dissertation acknowledged the importance for users to take informed data disclosure decisions and thus suggested privacy designers to follow the guidelines, described in Table 3.2.

Further, inspired by results of a qualitative investigation of users' privacy needs for accepting the requirements for the establishment of such services, this dissertation discovered additional usability and privacy limitations of existing ubiquitous social networking solutions, which were not taken into consideration in the reviewed privacy guidelines. Consequently, it proposed to update the current design guidelines for ubiquitous computing environments, by additionally suggesting to avoid four drawbacks when designing for privacy in ubiquitous social networking environments. The four drawbacks are: (i) ignoring variation of human data sensitivity, (ii) embracing disclosure to third parties, (iii) requiring too much user intervention and (iv) lacking user's personal data control.

In order to evaluate users' perceptions towards the management of personal privacy in relation to systems that follow the identified privacy guidelines, this dissertation suggested the design of a privacy-aware ubiquitous social networking

platform that overcomes the four drawbacks and allows users to take informed data disclosure decisions. The platform relies on a centralized architecture that enables disclosure of personal user information upon comparison of encrypted profiles. As a result, dynamic user sub-profiles, created by taking into account the human data sensitivity of the current circumstances, are effortlessly disclosed only to other end users, who hold a profile that might lead to potential networking benefits. Furthermore, thanks to its centralized architecture, the platform empowers users to modify the contents of their disclosed profiles at any time.

During qualitative interviews focusing on the proposed design solutions, the participants perceived such services to be more functional and more privacy-oriented, which consequently led them to feel more conformable with the disclosure of personal information in ubiquitous social networking environments. These findings strongly encourage privacy designers of ubiquitous computing environments, which target at supporting social networking between their inhabitants, to take into account the four drawbacks, illustrated in Table 3.3, additionally to the guidelines for enabling users to make informed data disclosure decisions, illustrated in Table 3.2.

The second direction to manage users' personal privacy in ubiquitous social networking is related to the challenge of identifying the influential factors that would affect the variation of human data sensitivity and consequent data disclosure decisions in such environments during users' encounters. To address this challenge, this dissertation identified 7 predictors for data disclosure, i.e. environment, location familiarity, activity, mood, familiar strangers, mutual friends and purpose of disclosure. These determinants were grouped into two different categories, i.e. contextual data and interrelated attributes. The former regards influential factors

4. CONCLUSIONS

related to the current contextual circumstances of the users' encounters in USN environments, while the latter consists of information that the user has in common with the inquirer.

When analyzing the identified influential factors through quantitative and qualitative investigations, it was discovered that participants selected the disclosed personal information by compromising between perceptions of data sensitivity for the current circumstances and evaluations of data relevance for gaining potential networking benefits. The users' data sensitivity was discovered to decrease as the relevance of information disclosure for initiation of networking increases. Among the identified influential factors, the purpose of disclosure and current environment primarily guided the participants in evaluation of their data sensitivity and relevance for exploiting ubiquitous social networking services. Furthermore, the other investigated influential factors were found to be predictors of secondary importance. In fact, the mood influential factor was impacting the potential participation in such environments, while all the other contextual data and interrelated attributes were discovered to only refine selection of disclosed personal information, rather than guiding the users in evaluating data relevance for better exploiting these services. When taking into account these influential factors, significant data disclosure prediction results were obtained with an approximate accuracy of 90% and potential for further increasing performance. These findings are useful for providing significant input for the design of ad hoc privacy management systems of ubiquitous social networking environments.

As future work, additional analysis is encouraged in relation to the Diverged Personalities privacy model. Large scale investigations are required to confirm the

results regarding the variation of human data sensitivity, presented in Appendix E. Such analysis should be carried out to provide further insight into the relationship between users' privacy attitudes about data disclosure in online social networks and ubiquitous social networking as well as many aspects of the investigated influential factors that still need further attention. Firstly, participants of qualitative interviews emphasized that the duration of the current activity might differently influence their data disclosure decisions. Especially in case of activities with very long duration, it is suggested to analyze whether the current activity might impact the evaluation of data disclosure relevance more than the current environment influential factor. Secondly, it is also important to statistically investigate whether knowing the identity of the mutual friends would influence users' sharing preferences. Two relevant features are suggested to be taken into consideration: self-reported closeness and clustering of users' friends into manageable categories (e.g. co-workers). Lastly, further insight is needed into the contradictory results about the location familiarity influential factor. Moreover, participants of qualitative investigations strongly appreciated options to provide their approval for disclosure of highly sensitive personal information. However, privacy designers of ubiquitous social networking are suggested to identify and investigate potential design solutions that would enable such feature, without requiring too much users' attention and intervention.

Finally, the findings reviewed in this dissertation should be supported with results of longitudinal investigations about users' data disclosure. Thus, it is suggested an implementation of the proposed privacy-aware ubiquitous social networking platform that should be adopted by a large number of users in order to collect users's

4. CONCLUSIONS

ad hoc data disclosure decisions for further statistical analysis. This study would be useful for researching the evolution of privacy attitudes and behaviors over time in ubiquitous social networking environments. Acceptance of potentially intrusive technologies can be observed to follow two phases - pessimistic, where the technology receives alarmist reactions from the potential users and optimistic, where users become more comfortable with the new technology [100].

As ubiquitous social networking is still in the early phase of its life cycle, it is crucial to identify and further investigate factors that would contribute to address users' concerns in order to facilitate the transition from the pessimistic phase to the optimistic one. This dissertation already provides many valuable insights into the users' privacy and usability concerns, however it can be expected that these topics are not exhaustive and thus additional research areas need to be identified and analyzed. Among others, it is strongly suggested to investigate the affect of social influence on perceptions of privacy in ubiquitous social networking [189]. Moreover, other factors might regard the topic of trust, as it was discovered to be an important aspect in users' acceptance of systems [150]. A special focus is suggested on users' personal experience in ubiquitous social networking, because the increasing familiarity with such services might have a significant affect on the acceptance and further adoption of this technology. For example, users might be at first very conservative about data disclosure in ubiquitous social networking, but might change this attitude and become more open about it over time [157].

References

- [1] M. S. Ackerman. The intellectual challenge of cscw: The gap between social requirements and technical feasibility. *Human-Computer Interaction*, 15(2):179–203, 2000. 17, 150, 211
- [2] A. Adams. Users’ perceptions of privacy in multimedia communications, 2001. 7, 10, 20
- [3] R. Agrawal, J. Kiernan, R. Srikant, and Y. Xu. Hippocratic databases. In *Proceedings of the 28th international conference on Very Large Data Bases*, pages 143–154. VLDB Endowment, 2002. 265
- [4] I. Altman. The environment and social behavior: Privacy, personal space, territory, and crowding. *CA: Brooks/Cole Publishing*, 1975. 11, 16, 78, 144, 150, 211, 264
- [5] I. Altman. Privacy regulation: culturally universal or culturally specific? *Journal of Social Issues*, 33(3):66–84, 1977. 11, 16, 78, 144, 150, 211, 264
- [6] M. Aspan. How sticky is membership on facebook? just try breaking free. *The New York Times*, 2008. 18, 67, 208, 221, 224
- [7] J. Barbosa, R. Hahn, D. N. F. Barbosa, and C. F. R. Geyer. Mobile and ubiquitous computing in an innovative undergraduate course. In *ACM SIGCSE Bulletin*, volume 39, pages 379–383. ACM, 2007. 2

REFERENCES

- [8] L. Barkhuus, B. Brown, M. Bell, S. Sherwood, M. Hall, and M. Chalmers. From awareness to repartee: sharing location within social groups. In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 497–506. ACM, 2008. 267, 302
- [9] L. Barkhuus and A. Dey. Location-based services for mobile telephony: a study of users privacy concerns. In *Proc. Interact*, volume 2003, pages 709–712. Citeseer, 2003. 7
- [10] R. Beale. Supporting social interaction with smart phones. *Pervasive Computing, IEEE*, 4(2):35–41, 2005. 50, 64, 66, 68, 170, 181, 183, 194, 207, 215
- [11] M. Beetz, B. Kirchlechner, and M. Lames. Computerized real-time analysis of football games. *Pervasive Computing, IEEE*, 4(3):33–39, 2005. 2
- [12] V. Bellotti and A. Sellen. Design for privacy in ubiquitous computing environments. In *Proceedings of the third conference on European Conference on Computer-Supported Cooperative Work*, pages 77–92. Kluwer Academic Publishers, 1993. 14, 15, 17, 52, 62, 149, 150, 208, 210, 211
- [13] B. Berendt, O. Günther, and S. Spiekermann. Privacy in e-commerce: stated preferences vs. actual behavior. *Communications of the ACM*, 48(4):101–106, 2005. 33, 56, 297
- [14] A. R. Beresford and F. Stajano. Location privacy in pervasive computing. *Pervasive Computing, IEEE*, 2(1):46–55, 2003. 8, 300
- [15] Hugh Beyer and Karen Holtzblatt. *Contextual design: defining customer-centered systems*. Morgan Kaufmann, 1997. 29
- [16] E. O. Blaß and M. Zitterbart. Towards acceptable public-key encryption in sensor

REFERENCES

- networks. In *ACM 2nd International Workshop on Ubiquitous Computing*, pages 88–93, 2005. 8
- [17] N. Bouillot, J. R. Cooperstock, E. Gressier-Soudan, R. Pellerin, T. Pietkiewicz, Z. Settel, and M. Wozniowski. Soundpark: Towards highly collaborative game support in a ubiquitous computing architecture. *IFIP Lecture Notes in Computer Science (LNCS)*, 5523(5523):157–170, 2011. 2
- [18] D. M. Boyd and N. B. Ellison. Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1):210–230, 2008. 3, 4, 121, 207
- [19] L. D. Brandeis and S. D. Warren. The right to privacy. *Harv.L.Rev.*, 4:193, 1890. 8, 143
- [20] P. Bresciani, A. Perini, P. Giorgini, F. Giunchiglia, and J. Mylopoulos. Tropos: An agent-oriented software development methodology. *Autonomous Agents and Multi-Agent Systems*, 8(3):203–236, 2004. 265
- [21] J. G. Breslin, S. Decker, M. Hauswirth, G. Hynes, D. Le Phuoc, A. Passant, A. Polleres, C. Rabsch, and V. Reynolds. Integrating social networks and sensor networks. In *W3C Workshop on the Future of Social Networking, Barcelona, January*, pages 15–16, 2009. 6, 124, 133
- [22] A. L. Brown and J. C. Campione. *Psychological theory and the design of innovative learning environments: On procedures, principles, and systems*. Lawrence Erlbaum Associates, Inc, 1996. 24
- [23] E. Bruns, B. Brombach, T. Zeidler, and O. Bimber. Enabling mobile phones to support large-scale museum guidance. *Multimedia, IEEE*, 14(2):16–25, 2007. 2

REFERENCES

- [24] C. Bünnig. Learning context based disclosure of private information. In *The Internet of Things & Services - 1st Intl.Research Workshop*, 2008. 82, 157, 196
- [25] C. Bünnig. Simulation and analysis of ad hoc privacy control in smart environments. *Intelligent Interactive Assistance and Mobile Multimedia Computing*, 53:307–318, 2009. 54, 82, 156, 158, 160, 196, 197, 242, 262, 264, 296, 297, 301
- [26] C. Bünnig. Smart privacy management in ubiquitous computing environments. *Human Interface and the Management of Information.Information and Interaction*, 5618:131–139, 2009. 15, 22, 54, 156, 157, 194, 235, 262, 264, 296, 297, 301
- [27] C. Bünnig and C. H. Cap. Ad hoc privacy management in ubiquitous computing environments. In *2009 Second International Conference on Advances in Human-Oriented and Personalized Mechanisms, Technologies, and Services*, pages 85–90. IEEE, 2009. 22, 82, 156, 158, 194, 196, 235, 264, 297, 301
- [28] R. B. Burns and R. A. Burns. *Business research methods and statistics using SPSS*. SAGE Publications Ltd, 2008. 43, 310, 311
- [29] J. W. Byun, E. Bertino, and N. Li. Purpose based access control of complex data for privacy protection. In *Proceedings of the tenth ACM symposium on Access control models and technologies*, pages 102–110. ACM, 2005. 265, 299
- [30] J. W. Byun and N. Li. Purpose based access control for privacy protection in relational database systems. *The VLDB Journal - The International Journal on Very Large Data Bases*, 17(4):603–619, 2008. 265, 299
- [31] I. Calemis, A. Kameas, C. Goumopoulos, and E. Berg. Astra: An awareness connectivity platform for designing pervasive awareness applications. *Innovations and Advances in Computer Sciences and Engineering*, pages 185–190, 2010. 134

REFERENCES

- [32] A. T. Campbell, S. B. Eisenman, N. D. Lane, E. Miluzzo, and R. A. Peterson. People-centric urban sensing. In *Proceedings of the 2nd annual international workshop on Wireless internet*, page 18. ACM, 2006. 7, 124
- [33] A. T. Campbell, S. B. Eisenman, N. D. Lane, E. Miluzzo, R. A. Peterson, H. Lu, X. Zheng, and M. Musolesi. The rise of people-centric sensing. *IEEE Internet Computing*, pages 12–21, 2008. 7, 124, 136
- [34] A. Candiotti and N. Clarke. The future of ubiquitous computing on campus. *Communications of the ACM*, 41(1):41, 1998. 2
- [35] J. C. Cano, P. Manzoni, and C. K. Toh. Ubiquimuseum: A bluetooth and java based context-aware system for ubiquitous computing. *Wireless Personal Communications*, 38(2):187–202, 2006. 2
- [36] M. Chalmers, I. MacColl, and M. Bell. Seamful design: Showing the seams in wearable computing. In *Euroearable, 2003. IEE*, pages 11–16. IET, 2003. 2
- [37] G. Chen and F. Rahman. Analyzing privacy designs of mobile social networking applications. In *IEEE/IFIP International Conference on Embedded and Ubiquitous Computing, 2008. EUC'08*, volume 2, 2008. 189
- [38] A. D. Cheok, K. H. Goh, W. Liu, F. Farbiz, S. W. Fong, S. L. Teo, Y. Li, and X. Yang. Human pacman: a mobile, wide-area entertainment system based on physical, social, and ubiquitous computing. *Personal and Ubiquitous Computing*, 8(2):71–81, 2004. 2
- [39] E. H. Chi. Introducing wearable force sensors in martial arts. *Pervasive Computing, IEEE*, 4(3):47–53, 2005. 2
- [40] D. K. W. Chiu, Y. T. F. Yueh, H. Leung, and P. C. K. Hung. Towards ubiquitous tourist service coordination and process integration: A collaborative travel

REFERENCES

- agent system architecture with semantic web services. *Information Systems Frontiers*, 11(3):241–256, 2009. 2
- [41] W. Chou. Elliptic curve cryptography and its applications to mobile devices. *University of Maryland, College Park*, 2003. 237
- [42] S. Clauß, A. Pfitzmann, M. Hansen, and E. Van Herreweghen. Privacy-enhancing identity management. *The IPTS Report*, 67:816, 2002. 20, 154
- [43] S. Cohen. Social relationships and health. *American Psychologist*, 59:676–684, 2004. 7, 124
- [44] The Design-Based Research Collective. Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, pages 5–8, 2003. 26
- [45] 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*, pages 81–90. ACM, 2005. 11, 14, 20, 23, 67, 82, 83, 160, 194, 236, 263, 265, 296, 299
- [46] D. J. Cook, A. Crandall, G. Singla, and B. Thomas. Detection of social interaction in smart spaces. *Cybernetics and Systems*, 41(2):90–104, 2010. 131
- [47] R. Cornejo. Integrating older adults into social networking sites through ambient intelligence. In *Proceedings of the 16th ACM international conference on Supporting group work*, pages 341–342. ACM, 2010. 2
- [48] S. Counts and K. E. Fisher. Mobile social networking: An information grounds perspective. In *Hawaii International Conference on System Sciences, Proceedings of the 41st Annual*, page 153. IEEE, 2008. 3, 169, 261

REFERENCES

- [49] L. Faith Cranor, J. Reagle, and M. S. Ackerman. Beyond concern: Understanding net users' attitudes about online privacy. In *In Proceedings of the Telecommunications Policy Research Conference*, 1999. 20
- [50] J. W. Creswell. *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage Publications, Inc, 2009. 24, 25, 26, 36, 182, 245, 313
- [51] J. W. Creswell and V. L. Plano Clark. *Designing and conducting mixed methods research*. Thousand Oaks (California): Sage Publications, 2007. 25, 305
- [52] C. Crook and D. Barrowcliff. Ubiquitous computing on campus: Patterns of engagement by university students. *International Journal of Human-Computer Interaction*, 13(2):245–256, 2001. 2
- [53] D. Dagger, A. O'Connor, S. Lawless, E. Walsh, and V. P. Wade. Service-oriented e-learning platforms: From monolithic systems to flexible services. *Internet Computing, IEEE*, 11(3):28–35, 2007. 2
- [54] A. Dale and R. B. Davies. *Analyzing social and political change: a casebook of methods*. Sage Publications, 1994. 26
- [55] CN Darrah, J. English-Lueck, and J. Freeman. Families at work: An ethnography of dual career families. *Report for the Sloane Foundation*, pages 98–96, 2001. 11
- [56] F. D. Davis. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 13(3):319–340, 1989. 31, 36, 142, 181, 184
- [57] S. Davis and C. Gutwin. Using relationship to control disclosure in awareness servers. In *Proceedings of Graphics Interface 2005*, pages 145–152. Canadian Human-Computer Communications Society, 2005. 11, 23, 160, 236, 263, 264, 296, 299

REFERENCES

- [58] Emiliano De Cristofaro, Jihye Kim, and Gene Tsudik. Linear-complexity private set intersection protocols secure in malicious model. In *Advances in Cryptology-ASIACRYPT 2010*, pages 213–231. Springer, 2010. 71
- [59] F. Delmastro, M. Conti, and A. Passarella. Social-aware content sharing in opportunistic networks. In *Sensor, Mesh and Ad Hoc Communications and Networks Workshops, 2009. SECON Workshops '09. 6th Annual IEEE Communications Society Conference on*, pages 1–3, 2009. 50, 170
- [60] J. Van den Akker. Principles and methods of development research. *Design methodology and developmental research in education and training*, pages 1–14, 1999. 24, 27
- [61] A. K. Dey and G. D. Abowd. Towards a better understanding of context and context-awareness. In *CHI 2000 workshop on the what, who, where, when, and how of context-awareness*, volume 4, pages 1–6. Citeseer, 2000. 3, 121, 295
- [62] EU Directive. 95/46/ec of the european parliament and of the council of 24 october 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. *Official Journal of the EC*, 23, 1995. 52, 64, 145, 146, 148, 151, 208, 209, 212, 222, 224, 266
- [63] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern classification*, volume 2. Citeseer, 2001. 131
- [64] N. Eagle. Mobile phones as social sensors. *The Handbook of Emergent Technologies in Social Research*, 2010. 6, 14, 124
- [65] N. Eagle and A. Pentland. Social serendipity: Mobilizing social software. *IEEE Pervasive Computing*, 4(2):28–34, 2005. 3, 4, 35, 50, 64, 66, 121, 122, 136, 139, 169, 170, 180, 181, 183, 194, 207, 208, 219, 244, 262, 296, 302, 305

REFERENCES

- [66] N. Eagle and A. Pentland. Reality mining: sensing complex social systems. *Personal and Ubiquitous Computing*, 10(4):255–268, 2006. 130
- [67] D. Easterly and A. Blachnitzky. Designing and evaluating ubiball: a ubiquitous computing game for children. *International Journal of Arts and Technology*, 4(3):276–293, 2011. 2
- [68] S. R. Eddy. Hidden markov models. *Current opinion in structural biology*, 6(3):361–365, 1996. 130, 131
- [69] D. Edwards, Food Ontario. Ministry of Agriculture, and Rural Affairs. *Personal Information Protection and Electronic Documents Act*. Ontario, Ministry of Agriculture, Food and Rural Affairs, 2005. 145, 146, 148
- [70] W. Edwards and R. Grinter. At home with ubiquitous computing: Seven challenges. In *Ubicomp 2001: Ubiquitous Computing*, pages 256–272. Springer, 2001. 2
- [71] B. Efron. The efficiency of logistic regression compared to normal discriminant analysis. *Journal of the American Statistical Association*, pages 892–898, 1975. 43, 310
- [72] K. Einarsdottir and Y. Li. Scatterfriend: A mobile social network using the beddernet p2p middleware, 2010. 139, 140, 142
- [73] S. B. Eisenman, N. D. Lane, E. Miluzzo, R. A. Peterson, G. S. Ahn, and A. T. Campbell. Metrosense project: People-centric sensing at scale. In *First Workshop on World-Sensor-Web (WSW2006)*. Citeseer, 2006. 7, 124
- [74] T. Erl. *Service-oriented architecture: concepts, technology, and design*. Prentice Hall PTR Upper Saddle River, NJ, USA, 2005. 133, 134
- [75] J. Floyd and JR Fowler. *Survey research methods*. Thousands Oaks, Sage, 2002. 29, 248, 317

REFERENCES

- [76] Nikolaus Franke, Eric Von Hippel, and Martin Schreier. Finding commercially attractive user innovations: A test of leaduser theory*. *Journal of Product Innovation Management*, 23(4):301–315, 2006. 27
- [77] M. Fukase, R. Akaoka, L. Lei, C. T. Shu, and T. Sato. Hardware cryptography for ubiquitous computing. In *Communications and Information Technology, 2005. ISGIT 2005. IEEE International Symposium on*, volume 1, pages 478–481. Ieee, 2005. 8
- [78] A. Galati and C. Greenhalgh. Exploring shopping mall environment for ubiquitous computing. *UbiComp at a Crossroad*, 2009. 2
- [79] E. Goffman. *Behavior in public places: Notes on the social organization of gatherings*. Simon and Schuster, 1966. 11, 20
- [80] E. Goffman. *The presentation of self in everyday life*. Harmondsworth, 1978. 11, 20, 155
- [81] R. S. Gohs and S. R. Gunnarsson. Beddernet-bluetooth scatternet framework for mobile devices, 2010. 139
- [82] R. E. Gregg. The privacy act of 1974. *Army Law.*, page 25, 1975. 52, 64, 145, 146, 148, 151, 208, 209, 212, 222, 224, 266
- [83] A. Gupta, A. Kalra, D. Boston, and C. Borcea. Mobisoc: a middleware for mobile social computing applications. *Mobile Networks and Applications*, 14(1):35–52, 2009. 136, 169, 170, 296
- [84] S. Gürses, R. Rizk, and O. Günther. Privacy design in online social networks: Learning from privacy breaches and community feedback. In *Twenty Ninth International Conference on Information Systems*, 2008. 18, 67, 208, 221, 224
- [85] JoAnn T. Hackos and Janice Redish. User and task analysis for interface design. 1998. 29

-
- [86] T. Hall and L. Bannon. Designing ubiquitous computing to enhance children’s learning in museums. *Journal of Computer Assisted Learning*, 22(4):231–243, 2006. 2
- [87] A. Hammershøj, A. Sapuppo, and R. Tadayoni. Challenges for mobile application development. In *Intelligence in Next Generation Networks (ICIN), 2010 14th International Conference on*, page 1, oct. 2010. 142, 178
- [88] D. R. Hankerson, S. A. Vanstone, and A. J. Menezes. *Guide to elliptic curve cryptography*. Springer-Verlag New York Inc, 2004. 71, 237
- [89] R. H. R. Harper. Why people do and don’t wear active badges: a case study. *Computer Supported Cooperative Work (CSCW)*, 4(4):297–318, 1995. 7
- [90] R. H. R. Harper, M. G. Lamming, and W. M. Newman. Locating systems at work: Implications for the development of active badge applications. *Interacting with Computers*, 4(3):343–363, 1992. 20
- [91] J. Hart, C. Ridley, F. Taher, C. Sas, and A. Dix. Exploring the facebook experience: a new approach to usability. In *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges*, pages 471–474. ACM, 2008. 18, 67, 208, 221, 224
- [92] T. Hashem and L. Kulik. Safeguarding location privacy in wireless ad-hoc networks. *Ubicomp 2007: Ubiquitous Computing*, pages 372–390, 2007. 8
- [93] A. Hasswa and H. Hassanein. Using heterogeneous and social contexts to create a smart space architecture. In *Computers and Communications (ISCC), 2010 IEEE Symposium on*, pages 1138–1142. IEEE, 2010. 6, 124, 133, 135
- [94] G. R. Hayes and G. D. Abowd. Tensions in designing capture technologies for an evidence-based care community. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 937–946. ACM, 2006. 31

REFERENCES

- [95] A. Heinemann. *Collaboration in opportunistic networks*. Universitäts-und Landesbibliothek Darmstadt, 2007. 50, 136, 170, 175
- [96] Eric Von Hippel. Lead users: a source of novel product concepts. *Management science*, 32(7):791–805, 1986. 27
- [97] Eric Von Hippel. New product ideas from lead users. *Research Technology Management*, 32(3):24–27, 1989. 27, 246
- [98] J. I. Hong. An architecture for privacy-sensitive ubiquitous computing, 2004. 7, 10, 11, 87, 144
- [99] J. I. Hong, J. D. Ng, S. Lederer, and J. A. Landay. Privacy risk models for designing privacy-sensitive ubiquitous computing systems. In *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques*, pages 91–100. ACM, 2004. 54, 143, 262, 296
- [100] G. Iachello and J. Hong. End-user privacy in human-computer interaction. *Foundations and Trends in Human-Computer Interaction*, 1(1):1–137, 2007. 9, 10, 11, 12, 14, 20, 25, 33, 56, 78, 85, 92, 143, 149, 162, 255, 297
- [101] G. Iachello, I. Smith, S. Consolvo, M. Chen, and G. D. Abowd. Developing privacy guidelines for social location disclosure applications and services. In *Proceedings of the 2005 symposium on Usable privacy and security*, pages 65–76. ACM, 2005. 267, 302
- [102] S. Ioannidis and A. Chaintreau. On the strength of weak ties in mobile social networks. In *Proceedings of the Second ACM EuroSys Workshop on Social Network Systems*, pages 19–25. ACM, 2009. 50, 170
- [103] U. Jendricke, M. Kreutzer, and A. Zugenmaier. Pervasive privacy with identity man-

-
- agement. In *Proceedings of the Workshop on Security in Ubiquitous Computing, Ubi-comp*. ACM Press, 2002. 20, 153, 154, 156, 194, 235, 264, 297, 301
- [104] X. Jiang, J. Hong, and J. Landay. Approximate information flows: Socially-based modeling of privacy in ubiquitous computing. *UbiComp 2002: Ubiquitous Computing*, pages 176–193, 2002. 16, 17, 52, 149, 151, 208, 210, 212
- [105] M. De Jode, J. Allin, D. Holland, A. Newman, C. Turfus, I. Litovski, R. Hayun, G. Sewell, S. Lewis, and M. Aubert. *Programming Java 2 Micro Edition on Symbian OS*. John Wiley, 2004. 138
- [106] P. Johnson, A. Kapadia, D. Kotz, N. Triandopoulos, and NH Hanover. People-centric urban sensing: Security challenges for the new paradigm. Technical report, Technical Report TR2007-586, Dartmouth College, Computer Science, Hanover, NH, 2007. 8, 143
- [107] A. Joly. Leveraging semantic technologies towards social ambient intelligence. *Ubiquitous and Pervasive Computing: Concepts, Methodologies, Tools, and Applications*, pages 1643–1668, 2010. 133
- [108] Q. Jones, S. A. Grandhi, S. Whittaker, K. Chivakula, and L. Terveen. Putting systems into place: a qualitative study of design requirements for location-aware community systems. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, pages 202–211. ACM, 2004. 11, 23, 160, 236, 263, 265, 296, 299
- [109] E. Kaasinen. User needs for location-aware mobile services. *Personal and ubiquitous computing*, 7(1):70–79, 2003. 7
- [110] K. Kalaiselvi and GV Uma. Integrated knowledge management approach for academic improvement in ubiquitous computing. *International Journal of Computer Applications IJCA*, 11(3):24–28, 2010. 2

REFERENCES

- [111] A. Kapadia, T. Henderson, J. Fielding, and D. Kotz. Virtual walls: Protecting digital privacy in pervasive environments. *Pervasive Computing*, pages 162–179, 2007. 20, 153, 154, 194, 235, 264
- [112] J. P. Kaps, G. Gaubatz, and B. Sunar. Cryptography on a speck of dust. *Computer*, 40(2):38–44, 2007. 8
- [113] S. Khajuria and H. Tange. Implementation of diffie-hellman key exchange on wireless sensor using elliptic curve cryptography. In *Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology, 2009. Wireless VITAE 2009. 1st International Conference on*, pages 772–776. IEEE, 2009. 237
- [114] V. Kostakos and E. O’Neill. Cityware: Urban computing to bridge online and real-world social networks. *Handbook of Research on Urban Informatics: The Practice and Promise of the Real-Time City*, pages 195–204, 2008. 137
- [115] Sari Kujala and Marjo Kauppinen. Identifying and selecting users for user-centered design. In *Proceedings of the third Nordic conference on Human-computer interaction*, pages 297–303. ACM, 2004. 27, 29, 85, 255
- [116] Sari Kujala and Martti Mntyl. *How effective are user studies?*, pages 61–71. People and Computers XIV Usability or Else! Springer, 2000. 29
- [117] N. Kushwaha, M. Kim, D. Y. Kim, and W. D. Cho. An intelligent agent for ubiquitous computing environments: smart home ut-agent. In *Software Technologies for Future Embedded and Ubiquitous Systems, 2004. Proceedings. Second IEEE Workshop on*, pages 157–159. IEEE, 2004. 2
- [118] S. Kvale. Interviews: An introduction to qualitative research interviewing. *Evaluation and program planning*, 20(3):287–288, 2004. 36, 182, 245, 313

REFERENCES

- [119] J. Ladd. Computers and moral responsibility: a framework for an ethical analysis. In *Computerization and controversy*, pages 664–675. Academic Press Professional, Inc., 1991. 15
- [120] M. Langheinrich. Privacy by design - principles of privacy-aware ubiquitous systems. In *UbiComp 2001: Ubiquitous Computing*, pages 273–291. Springer, 2001. 7, 9, 14, 16, 17, 18, 52, 145, 146, 147, 148, 149, 208, 209, 210, 212, 264, 265, 266, 300
- [121] J. Lawrence, T. R. Payne, and D. De Roure. Co-presence communities: Using pervasive computing to support weak social networks. In *Enabling Technologies: Infrastructure for Collaborative Enterprises, 2006. WETICE'06. 15th IEEE International Workshops on*, pages 149–156. IEEE, 2007. 129
- [122] 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*, 8(6):440–454, 2004. 14, 15, 17, 18, 20, 21, 22, 52, 64, 66, 67, 149, 150, 151, 156, 157, 163, 208, 210, 211, 212, 223, 227, 235, 264, 297, 311
- [123] 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*, pages 724–725. ACM, 2003. 11, 20, 23, 67, 154, 160, 194, 236, 263, 264, 296, 298
- [124] S. Lederer, J. Mankoff, A. K. Dey, and C. Beckmann. *Managing personal information disclosure in ubiquitous computing environments*. Citeseer, 2003. 20, 154, 156, 194, 235, 264
- [125] Scott Lederer, Anind K Dey, and Jennifer Mankoff. A conceptual model and a metaphor of everyday privacy in ubiquitous computing environments. 2002. 23

REFERENCES

- [126] K. M. Leung. Naive bayesian classifier. Technical report, Technical Report, Department of Computer Science / Finance and Risk Engineering, Polytechnic University, Brooklyn, New York, USA, 2007. 131, 158
- [127] D. Lewis. Naive (bayes) at forty: The independence assumption in information retrieval. *Machine Learning: ECML-98*, pages 4–15, 1998. 130
- [128] L. Lilien, Z. H. Kamal, V. Bhuse, and A. Gupta. Opportunistic networks: the concept and research challenges in privacy and security. *Proc.of the WSPWN*, pages 134–147, 2006. 261
- [129] Y. Liu and S. K. Das. Information-intensive wireless sensor networks: potential and challenges. *Communications Magazine, IEEE*, 44(11):142–147, 2006. 6, 124
- [130] D. Lyons. The high price of facebook. *Newsweek*, 2010. 18, 67, 208, 221, 224
- [131] D. J. Malan, M. Welsh, and M. D. Smith. A public-key infrastructure for key distribution in tinyos based on elliptic curve cryptography. In *Sensor and Ad Hoc Communications and Networks, 2004. IEEE SECON 2004. 2004 First Annual IEEE Communications Society Conference on*, pages 71–80. IEEE, 2004. 237
- [132] Edwin Mansfield. *Industrial research and technological innovation: An econometric analysis*. Norton New York, 1968. 27
- [133] S. Mascetti and C. Bettini. A comparison of spatial generalization algorithms for lbs privacy preservation. In *Mobile Data Management, 2007 International Conference on*, pages 258–262. IEEE, 2007. 8
- [134] F. Massacci, J. Mylopoulos, and N. Zannone. Hierarchical hippocratic databases with minimal disclosure for virtual organizations. *The VLDB journal*, 15(4):370–387, 2006. 265

REFERENCES

- [135] I. Mavrommati and I. Calemis. Astra awareness connectivity platform based on service oriented concepts. *Constructing Ambient Intelligence*, pages 70–74, 2009. 133, 134
- [136] C. McNamara. General guidelines for conducting interviews. *Retrieved December, 20:2003, 1999.* 36, 182, 245, 313
- [137] F. Michahelles and B. Schiele. Sensing and monitoring professional skiers. *Pervasive Computing, IEEE*, 4(3):40–45, 2005. 2
- [138] S. Milgram. The familiar stranger: An aspect of urban anonymity. *The individual in a social world*, pages 51–53, 1977. 129, 270
- [139] E. Miluzzo, N. D. Lane, K. Fodor, R. Peterson, H. Lu, M. Musolesi, S. B. Eisenman, X. Zheng, and A. T. Campbell. Sensing meets mobile social networks: the design, implementation and evaluation of the cenceme application. In *Proceedings of the 6th ACM conference on Embedded network sensor systems*, pages 337–350. ACM, 2008. 8, 131, 133, 134, 235
- [140] M. F. Mokbel, C. Y. Chow, and W. G. Aref. The new casper: query processing for location services without compromising privacy. In *Proceedings of the 32nd international conference on Very large data bases*, pages 763–774. VLDB Endowment, 2006. 8
- [141] D. S. Moore and G. P. McCabe. *Introduction to the Practice of Statistics Chapters 14-17*. WH Freeman & Co, 2005. 42, 275
- [142] J. M. Morse. Approaches to qualitative-quantitative methodological triangulation. *Nursing Research*, 40(1):120–123, 1991. 25, 306
- [143] M. Musolesi, E. Miluzzo, N. D. Lane, S. B. Eisenman, T. Choudhury, and A. T.

REFERENCES

- Campbell. The second life of a sensor: Integrating real-world experience in virtual worlds using mobile phones. *Technology*, 100:3, 2008. 6, 124
- [144] G. Myles, A. Friday, and N. Davies. Preserving privacy in environments with location-based applications. *IEEE Pervasive Computing*, pages 56–64, 2003. 20, 153, 154, 194, 235
- [145] F. Nagle and L. Singh. Can friends be trusted? exploring privacy in online social networks. In *Social Network Analysis and Mining, 2009. ASONAM'09. International Conference on Advances in*, pages 312–315. IEEE, 2009. 83, 301
- [146] DH Nguyen and ED Mynatt. Privacy mirrors: Making ubicomp visible. In *Human Factors in Computing Systems: CHI 2001 (Workshop on Building the User Experience in Ubiquitous Computing)*, 2001. 224
- [147] T. Nicolai, E. Yoneki, N. Behrens, and H. Kenn. Exploring social context with the wireless rope. In *On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops*, pages 874–883. Springer, 2006. 137
- [148] J. S. Olson, J. Grudin, and E. Horvitz. A study of preferences for sharing and privacy. In *CHI'05 extended abstracts on Human factors in computing systems*, pages 1985–1988. ACM, 2005. 11, 23, 263, 265, 296, 299
- [149] L. Palen and P. Dourish. Unpacking privacy for a networked world. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 129–136. ACM, 2003. 11, 16, 17, 52, 78, 144, 149, 150, 208, 210, 211, 264
- [150] A. Patrick, S. Marsh, and P. Briggs. Designing systems that people will trust. 2005. 92
- [151] C. Y. J. Peng, K. L. Lee, and G. M. Ingersoll. An introduction to logistic regression

REFERENCES

- analysis and reporting. *The Journal of Educational Research*, 96(1):3–14, 2002. 43, 310, 311
- [152] J. Perkio, V. Tuulos, M. Hermersdorf, H. Nyholm, J. Salminen, and H. Tirri. Utilizing rich bluetooth environments for identity prediction and exploring social networks as techniques for ubiquitous computing. In *Web Intelligence, 2006. WI 2006. IEEE/WIC/ACM International Conference on*, pages 137–144. IEEE, 2007. 129
- [153] P. Persson and Y. Jung. Nokia sensor: from research to product. In *Proceedings of the 2005 conference on Designing for User eXperience*, page 53. AIGA: American Institute of Graphic Arts, 2005. 3, 50, 64, 66, 68, 121, 128, 138, 170, 175, 183, 194, 207, 214
- [154] M. Petkovic, D. Prandi, and N. Zannone. Purpose control: did you process the data for the intended purpose? *Secure Data Management*, pages 145–168, 2011. 299
- [155] A. K. Pietiläinen, E. Oliver, J. LeBrun, G. Varghese, and C. Diot. Mobiclique: middleware for mobile social networking. In *Proceedings of the 2nd ACM workshop on Online social networks*, pages 49–54. ACM, 2009. 141, 142, 143, 169, 170, 181, 296
- [156] WR Pires, A. A. F. Loureiro, and R. A. R. Oliveira. Using web technologies in assessment of context-aware pervasive/ubiquitous systems: A tourist guide service. In *Wireless and Mobile Computing, Networking and Communications (WiMob), 2010 IEEE 6th International Conference on*, pages 691–698. IEEE, 2010. 2
- [157] M. Prabaker, J. Rao, I. Fette, P. Kelley, L. Cranor, J. Hong, and N. Sadeh. Understanding and capturing peoples privacy policies in a people finder application. In *Proc. Workshop Ubicomp Privacy*, 2007. 92
- [158] S. J. Press and S. Wilson. Choosing between logistic regression and discriminant

REFERENCES

- analysis. *Journal of the American Statistical Association*, pages 699–705, 1978. 43, 310
- [159] J. Rana, J. Kristiansson, J. Hallberg, and K. Synnes. An architecture for mobile social networking applications. In *Computational Intelligence, Communication Systems and Networks, 2009. CICSYN '09. First International Conference on*, pages 241–246, 2009. ID: 1. 169
- [160] E. Richards. Proposed fbi crime computer system raises questions on accuracy, privacy. In *Computerization and controversy*, pages 436–438. Academic Press Professional, Inc., 1991. 15
- [161] DJ Riemen. The essential structure of a caring interaction: Doing phenomenology. *Nursing research: A qualitative perspective*, pages 85–105, 1986. 25
- [162] G. Riva and F. Gramatica. From stethoscope to ambient intelligence: the evolution of healthcare. *International Journal of Healthcare Technology and Management*, 5(3):268–283, 2003. 2
- [163] Colin Robson. *Real world research: A resource for social scientists and practitioner-researchers*, volume 2. Blackwell Oxford, 2002. 85, 256
- [164] K. S. Rook, S. Mavandadi, D. H. Sorkin, and L. A. Zettel. Optimizing social relationships as a resource for health and well-being in later life. *Handbook of health psychology and aging*, pages 267–285, 2007. 7, 124
- [165] N. Sadeh, J. Hong, L. Cranor, I. Fette, P. Kelley, M. Prabaker, and J. Rao. Understanding and capturing peoples privacy policies in a mobile social networking application. *Personal and Ubiquitous Computing*, 13(6):401–412, 2009. 153
- [166] A. Sapuppo. Spiderweb: a social mobile network. In *Wireless Conference (EW), 2010 European*, pages 475–481, 2010. 3, 24, 35, 40, 121, 142, 178, 182, 183, 261, 305

-
- [167] A. Sapuppo. Privacy analysis in mobile social networks: the influential factors for disclosure of personal data. *International Journal of Wireless and Mobile Computing*, 5(4):315–326, 2012. 67, 68, 79, 83, 160, 194, 209, 224, 236, 248, 297, 300, 301, 304, 317
- [168] A. Sapuppo. Ubiquitous social networking: Concept and evaluation. *Sensor Letters*, 10(8):1632–1644, 2012. 4, 10, 18, 35, 54, 67, 68, 79, 122, 208, 209, 224, 242, 244, 296, 302, 304, 305
- [169] A. Sapuppo. The influential factors for the variation of data sensitivity in ubiquitous social networking. *International Journal of Wireless and Mobile Computing*, 2013. 20, 67, 160, 194, 208, 236
- [170] A. Sapuppo and B. C Seet. An empirical investigation of disclosure of personal information in ubiquitous social computing. *International Journal of Computer Theory and Engineering*, 4(3):373–378, 2012. 11, 14, 15, 20, 42, 83, 159, 160, 183, 208, 223, 296, 309
- [171] A. Sapuppo and L. T. Sørensen. Local social networks. In *International Proceedings of Computer Science and Information Technology - Computer Communication and Management*, volume 5, pages 15–22, 2011. 3, 4, 52, 54, 64, 66, 122, 141, 142, 158, 207, 208, 223, 235, 261, 262, 263, 264, 296
- [172] R. Schollmeier. A definition of peer-to-peer networking for the classification of peer-to-peer architectures and applications. In *Peer-to-Peer Computing, 2001. Proceedings. First International Conference on*, pages 101–102. IEEE, 2002. 136, 138
- [173] T. E. Seeman. Social ties and health: The benefits of social integration. *Annals of Epidemiology*, 6(5):442–451, 1996. 7, 124

REFERENCES

- [174] B. C Seet and J. R. Casar. Physical object search and reminding in ambient intelligent spaces. In *Proceedings of Conference of the Spanish Association for Artificial Intelligence*, 2007. 2
- [175] J. Shimoyama. Development and demonstration of a ubiquitous system for tourist services. *Joho Chishiki Gakkaishi*, 20(3):231–238, 2010. 2
- [176] A. Smailagic and D. Kogan. Location sensing and privacy in a context-aware computing environment. *Wireless Communications, IEEE*, 9(5):10–17, 2002. 194, 235
- [177] H. J. Smith. Privacy policies and practices: inside the organizational maze. *Communications of the ACM*, 36(12):104–122, 1993. 9
- [178] I. Smith, S. Consolvo, A. Lamarca, J. Hightower, J. Scott, T. Sohn, J. Hughes, G. Iachello, and G. D. Abowd. Social disclosure of place: From location technology to communication practices. *Pervasive Computing*, pages 134–151, 2005. 267, 302
- [179] F. Stajano. *Front Matter and Index*. Wiley Online Library, 2002. 9
- [180] J. Suh and C. Woo. Design and development of a social intelligence based context-aware middleware using blackboard. In *Tools with Artificial Intelligence (ICTAI), 2011 23rd IEEE International Conference on*, pages 908–910. IEEE, 2011. 3, 121, 295
- [181] Y. Sumi, T. Etani, S. Fels, N. Simonet, K. Kobayashi, and K. Mase. C-map: Building a context-aware mobile assistant for exhibition tours. *Community computing and support systems*, pages 137–154, 1998. 2
- [182] B. G. Tabachnick, L. S. Fidell, and S. J. Osterlind. *Using multivariate statistics*. Allyn and Bacon Boston, 2001. 43, 310

REFERENCES

- [183] P. Tamarit, C. T. Calafate, J. C. Cano, and P. Manzoni. Bluefriend: Using bluetooth technology for mobile social networking. In *Mobile and Ubiquitous Systems: Networking & Services, MobiQuitous, 2009. MobiQuitous' 09. 6th Annual International*, pages 1–2. IEEE, 2009. 3, 50, 64, 66, 121, 139, 170, 183, 194, 207, 216, 296, 302
- [184] T. Terano. *New frontiers in artificial intelligence: joint JSAI 2001 workshop post-proceedings*, volume 2253. Springer Verlag, 2001. 3, 121, 295
- [185] Y. Tian, B. Song, and E. N. Huh. A privacy-aware system using threat-based evaluation and feedback method in untrusted ubiquitous environments. *Security Technology*, pages 193–200, 2009. 265, 299
- [186] A. Toninelli, A. Pathak, A. Seyedi, S. Cardoso, and V. Issarny. Middleware support for mobile social ecosystems. In *Computer Software and Applications Conference Workshops (COMPSACW), 2010 IEEE 34th Annual*, pages 293–298. IEEE, 2010. 138
- [187] Glen L. Urban and Eric Von Hippel. Lead user analyses for the development of new industrial products. *Management science*, 34(5):569–582, 1988. 27
- [188] S. A. Vanstone. Next generation security for wireless: elliptic curve cryptography. *Computers & Security*, 22(5):412–415, 2003. 237
- [189] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis. User acceptance of information technology: Toward a unified view. *MIS quarterly*, pages 425–478, 2003. 92
- [190] F. Wang and M. J. Hannafin. Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development*, 53(4):5–23, 2005. 24, 26, 27

REFERENCES

- [191] M. Weiser. The computer for the 21st century. *Scientific American*, 272(3):78–89, 1995. 2, 8, 19, 69, 143, 146, 209, 224, 295
- [192] M. Weiser. Open house. *Rank Xerox PARC*, 1996. 2, 295
- [193] M. Weiser and J. S. Brown. Designing calm technology. In *PowerGrid Journal*. Citeseer, 1996. 2, 19, 66, 69, 146, 209, 224, 295
- [194] M. Weiser, R. Gold, and J. S. Brown. The origins of ubiquitous computing research at parc in the late 1980s. *IBM Systems Journal*, 38(4):693–696, 1999. 7
- [195] E. W. Weisstein. Bonferroni correction. *MathWorldA Wolfram Web Resource*, 2004. 42, 275
- [196] A. F. Westin. *Privacy and freedom*, volume 97. London, 1967. 9, 11, 143, 144
- [197] A. F. Westin. *Information technology in a democracy*. Harvard University Press, 1971. 9, 144
- [198] A. F. Westin. Harris-equifax consumer privacy survey 1991. *Atlanta, GA: Equifax Inc*, 1991. 28, 184, 247, 272, 316
- [199] J. Wiese, P. G. Kelley, L. F. Cranor, L. Dabbish, J. I. Hong, and J. Zimmerman. Are you close with me? are you nearby? investigating social groups, closeness, and willingness to share. In *Proceedings of the 13th international conference on Ubiquitous computing*, pages 197–206. ACM, 2011. 11, 23, 160, 236, 299
- [200] G. Wijnalda, S. Pauws, F. Vignoli, and H. Stuckenschmidt. A personalized music system for motivation in sport performance. *Pervasive Computing, IEEE*, 4(3):26–32, 2005. 2
- [201] F. Wilcoxon. Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6):80–83, 1945. 42, 275

REFERENCES

- [202] D. Wright, S. Gutwirth, M. Friedewald, P. De Hert, M. Langheinrich, and A. Moscibroda. Privacy, trust and policy-making: Challenges and responses. *Computer Law & Security Report*, 25(1):69–83, 2009. 145, 264
- [203] H. C. Yang and W. Y. Wang. Facilitating academic service-learning with android-based applications and ubiquitous computing environment. In *Fourth International Conference on Ubi-Media Computing*, pages 191–196. IEEE, 2011. 2
- [204] G. Yee. Using privacy policies to protect privacy in ubicomp. In *Proceedings of the IEEE 19th International Conference on Advanced Information Networking and Applications (AINA 2005 Volume II)*, 2005. 145, 146, 147, 153, 264
- [205] M. Youngblood, D. J. Cook, and L. B. Holder. Seamlessly engineering a smart environment. In *Systems, Man and Cybernetics, 2005 IEEE International Conference on*, volume 1, pages 548–553. IEEE, 2006. 121, 295
- [206] N. D. Ziv and B. Mulloth. An exploration on mobile social networking: Dodgeball as a case in point. In *Mobile Business, 2006. ICMB'06. International Conference on*, page 21. IEEE, 2007. 3, 64, 169, 189, 261

REFERENCES

Appendix A

Privacy and technology challenges for ubiquitous social networking

Sapuppo Antonio Center for Communication, Media and Information Technologies - Aalborg University, Sydhavnsgade 17, Copenhagen 2450, Denmark - Email: antonio@cmi.aau.dk

Boon-Chong Seet Department of Electrical and Electronic Engineering - Auckland University of Technology, Private Bag 92006, Auckland 1142, New Zealand - Email: bseet@aut.ac.nz

Accepted for publication in International Journal of Ad Hoc and Ubiquitous Computing.

Abstract

Ubiquitous social networking can be seen as an evolution of ubiquitous computing supporting the social well-being of people in their everyday lives. The vision of ubiquitous social networking focuses on enhancing social interactions among its participants during users' physical meetings. This target is leading towards important challenges such as social sensing, enabling social networking and privacy protection. In this paper we firstly investigate the methods and technologies for acquisition of the relevant context for promotion of sociability among inhabitants of ubiquitous social networking environments. Afterwards, we review architectures and techniques for enabling social interactions between participants. Finally, we identify privacy as the major challenge for networking in ubiquitous social networking environments. Consequently, we depict design guidelines and review privacy protection models for facilitating personal information disclosure.

Keywords: Privacy; Ubiquitous Computing; Information Disclosure; Context awareness; Social Networking; Design Guidelines.

A.1 Introduction

The great popularity of online social networks has inspired ubiquitous computing researchers and practitioners to investigate possibilities for improving human communication by enhancing social networking and transferring online social networks benefits to the physical world [65, 153, 183]. To achieve this goal, it is essential for ubiquitous computing to embody social intelligence in order to intelligently and naturally support human communication in the physical world. Social intelligence can be defined as the ability of the environment to acquire and apply users' social context in order to foster social interactions among its inhabitants [61, 166, 180, 184]. This can be considered as an evolution of ubiquitous computing, where a social dimension has been introduced to respond to the social nature of the users and increase awareness, knowledge and intelligence of these environments [205].

This extension of ubiquitous computing can be defined as ubiquitous social computing and the networking services established in such environments as Ubiquitous Social Networking (USN). The term *networking* is preferred to the term *network*, as in the case of online social network sites, because such online services principally focus on supporting communication within users' existing social networks. On the contrary, *networking* emphasizes relationship initiation, often between strangers. It is related to the development of possible advantageous relationships such as friendships, partnerships and business relations by uncovering hidden connections that people share with others nearby. [18]. Particularly, services that focus on networking, e.g. USN, facilitate initialization of face-to-face interactions between strangers with similar interests, i.e. *people who do not know each other, but probably should*.

A. PRIVACY AND TECHNOLOGY CHALLENGES FOR UBIQUITOUS SOCIAL NETWORKING

As a result, the value of social networking is significantly enhanced and benefits are available immediately upon demand [65, 171]. Potential application areas of USN are numerous and they range from professional, where these services might lead to new opportunities such as connecting employers with potential employees, to big events, such as conferences, company events and exhibitions that usually comprise large amounts of participants who potentially share similar professional or social interests [65, 168].

In order to better explain the concept of USN services, a scenario is presented in Figure A.1. A user named Bob, who is marked in blue, is located in a public place, such as a canteen of an ordinary work place. Bob is surrounded by people whom he knows, marked in green, and people who are strangers to him, marked in white. Even if Bob does not interact with all people in the canteen, his ubiquitous devices do that for him by exchanging personal information with other people in his

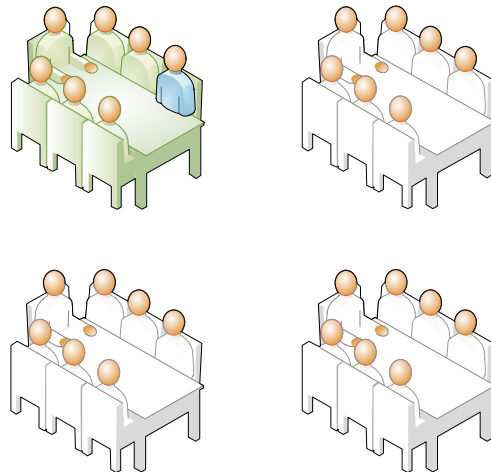


Figure A.1: Ubiquitous social networking example scenario

proximity, as shown in Figure A.2-A. Due to automatic exchange of users' personal information, this process does not interfere with the current users' activities, and it allows USN to develop an understanding about who the people nearby are, as well as their respective preferences.

As shown in Figure A.2-B, these services are capable of identifying users with similar interests, and thus highlighting relevant social paths between users that would remain hidden otherwise. When USN services find profile similarities between Bob and other users, who are highlighted in yellow in Figure A.2-B, users are notified about each others' presence and, therefore, have the opportunity to immediately initiate a face-to-face communication.

To translate the USN vision into reality, there is an indisputable need for social sensors, as sensing and inferring of the relevant environmental and social information of the users are crucial for reaching the primary goal to foster social interac-

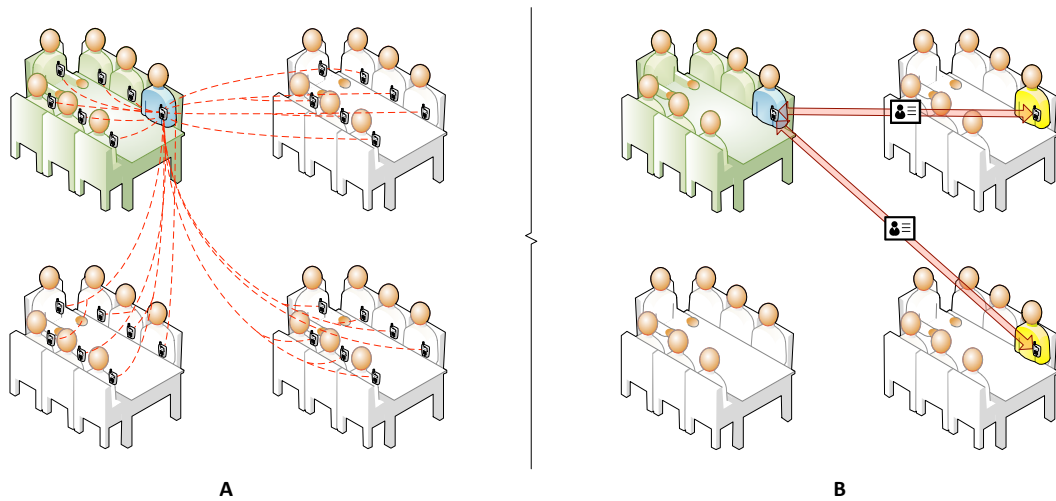


Figure A.2: Automated personal information exchange and user profile matching in ubiquitous social networking

A. PRIVACY AND TECHNOLOGY CHALLENGES FOR UBIQUITOUS SOCIAL NETWORKING

tions. Sensor networks have been deployed in ubicomp environments for obtaining context information. However, these sensors cannot be considered optimal due to their limited resources, such as restricted energy, networking and storage capabilities and they are also incapable of gathering users' social information, such as users' preferences [93, 129]. To overcome these limitations, latest research identifies smartphones as potential social sensors [64]. These devices are already equipped with GPS, accelerometer and other sensing components, enabling acquisition of environmental context information [21, 143]. They are also ideally suited to provide an insight into social behavior patterns, because they can be seen as wearable and inconspicuous sensors, constantly carrying users' personal information [21]. Finally, smartphones are also considered to be gateways to social networks, which are significant sources of users' personal information such as interests, preferences and list of friends [21, 64].

The vast amount of relevant context information, acquired by such social sensors, enables more personalized USN services, where inhabitants of USN environments are the primary targets of sensing. These services can be considered to be based on people-centric sensing approaches, because they focus mainly on social components, such as user preferences, identities and relationships between the users, rather than just machines and devices [32, 33, 73]. This collection and further dissemination of users' personal information provides a crucial foundation for USN services to better empower people in their social conduct and enhance their sense of social connectedness, which in turn will play an important role in their physical and cognitive well-being [43, 164, 173]. However, in order to enable those services, it is necessary to address the following mutually dependent technological and psychological

challenges:

1. Context acquisition: USN environments must be capable of acquiring the relevant context in order to promote sociability among its inhabitants. Further, an evaluation of the obtained context must be carried out to elaborate its significance and relevancy, which conduces to learning users' behavioral and social patterns for personalizing social networking services.
2. Social networking: USN environments must be capable of enabling social interactions between its participants. Moreover, those services have to be applied not only among acquaintances but also between strangers with interpersonal affinities. Hence it would lead to highlighting relevant social paths between users in the physical world, that would remain hidden otherwise.
3. Privacy: USN must be capable of providing a secure and safe collection and dissemination of participants' personal information as well as ensuring an accurate selection of users' personal information to be disclosed to others. This challenge arises due to the fact that the foundation of ubiquitous social networking is based on sharing of participants' personal information, which can provoke potential privacy threats.

In this paper, we review previous works, which focus on sensing the relevant context for promotion of sociability among USN users (Section A.2) and different software architectures for enabling social interactions (Section A.3). In Section A.4, we discuss design guidelines for protecting users' privacy and in Section A.5 we present existing privacy models that focus on managing information disclosure in

A. PRIVACY AND TECHNOLOGY CHALLENGES FOR UBIQUITOUS SOCIAL NETWORKING

ubiquitous social computing environments. Final conclusions and recommendations for future work are drawn in Section A.6.

A.2 Context acquisition

The nature of USN context is heterogeneous because it is composed of diverse environmental and social contexts of the users. It is considered to be the relevant context to be acquired by the environment for reaching the primary goal of fostering social interactions among its inhabitants. In this section we are going to review recent studies that provide significant contribution to the acquisition and inference of the USN context, shown in Figure A.3.

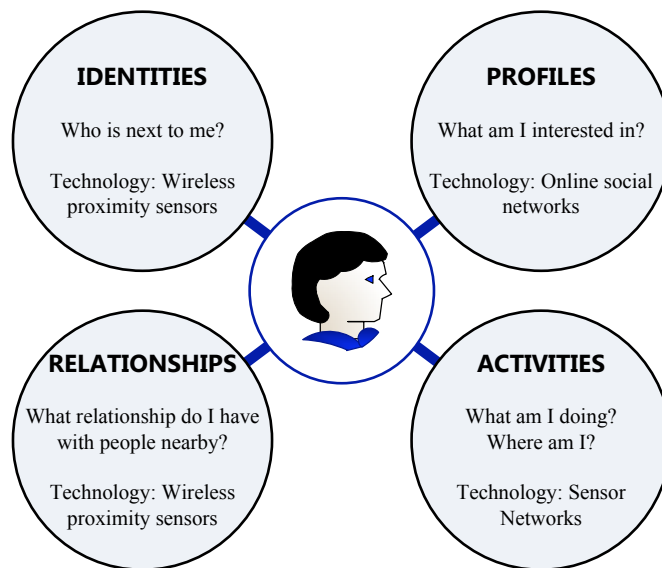


Figure A.3: Relevant context to be acquired by the USN environment to foster interactions among its inhabitants through ubiquitous social networking

Specifically, we will be focusing on works that:

- present test results of existing implemented solutions, which can be applied for improvement of users' social conduct through the USN environments;
- take into account only approaches that are based on wearable sensors with most focus on mobile phones.

Moreover, the selected papers were analyzed by categorizing them into three areas:

1. Users' identities and type of relationships: we investigated acquisition of users' identities and understanding types of relationships between the users, by reviewing only relevant methods that utilize the Bluetooth technology as wireless proximity sensors. Importantly, despite the Bluetooth limitations (e.g. power consumption), we focus on this solution rather than others (e.g. conversation network analysis for inference of relationship between users), because it is the most widely applied in recent studies, especially with test results acquired from large number of participants. We did not consider solutions that utilize Near Field Communication (NFC), although NFC has been recently utilized for exchanging social business cards in mobile social networks, such as Poken¹². Due to its limited range, NFC was not found to be suitable for ubiquitous computing environments that target at promoting social interactions between strangers with similar interests, i.e. USN. For example, users of Poken are capable of exchanging their digital business cards by only touching each others' mobile phones. This usually occurs after having an ordinary

¹²<http://www.poken.com>

A. PRIVACY AND TECHNOLOGY CHALLENGES FOR UBIQUITOUS SOCIAL NETWORKING

face-to-face interaction between the users, who decided to keep in contact afterwards. As the main goal of USN is to promote initiation of relationship, other solutions based on Bluetooth technology were found to be more relevant. The range of Bluetooth was already discussed to be short enough to ensure that users are in the proximity of each other. At the same time it was found to be long enough for users to exchange personalized contents without being adjacent to each other [153];

2. Users' activities: we present two different studies on inference of users' activities. We concentrate on these solutions because they enable anticipation of users' social activities in two different perspectives. The first approach infers prediction of generic information about the current users' activities focusing on individual activities of a single user. The second approach, instead, supplies anticipations of more detailed collaborative social activities between users in such environments;
3. Users' online profiles: we explore works that integrated online social networks in ubiquitous social computing environments to retrieve users' profiles. Relevantly, online social networks are a vast source of users' personal information (e.g. list of friends, preferences, etc), which can be successfully applied for promoting social interactions in USN environments.

A.2.1 Users' identities and type of relationships

In regard to acquisition of users' identities and type of relationships, mobile phones have been exploited as social proximity sensors because of their wireless capabilities.

Among the wireless technologies available, Bluetooth was usually preferred due to its common adoption in mobile phones and characteristics such as device and service discoveries. In this case, the acquisition range of USN context depends on the power classes of the Bluetooth radio. Obviously, a more powerful radio implies larger power consumption. Mobile phones are typically equipped with class 2 Bluetooth radios, which provide a range of up to 10 meters for use as social proximity sensors.

Bluetooth-enabled mobile phones were adopted as social proximity sensors by Lawrence et al in their research on the subject of familiar strangers [121]. Two people are classified as familiar strangers if they encounter regularly without interacting or forming an explicit relationship of a social nature [138]. The challenge of the authors was to identify the social groups among familiar strangers. They defined those groups as co-presence communities, i.e. groups of individuals who regularly share a particular location at the same time. Members of co-presence communities are not required to have any social interactions, but due to repeated collocations of individuals, they were expected to have common interests in the functioning of the local area (e.g. punctuality of a bus among the people who tend to take the same bus). Thus, the authors implemented an Ambient Information Dissemination Environment (AIDE), which is able to acquire and exploit the relevant context. AIDE detects the mobile devices, identifies users and the community to which they belong, and automatically disseminates content in the background, based on the preferences specified by the users.

Acquisition and applications of users' identities were further explored by Perkio et al in [152]. The main objective of their research was to predict users' locations by monitoring their Bluetooth neighborhoods in a work environment. The environ-

A. PRIVACY AND TECHNOLOGY CHALLENGES FOR UBIQUITOUS SOCIAL NETWORKING

ments were equipped with several passive Bluetooth beacons. The purpose of the beacons was to scan the surroundings and detect Bluetooth-enabled mobile phones, carried by users as wearable sensors. The collected data was stored in a back-end server, and subsequently used as input data for training Naive Bayes models [127] in order to predict inhabitants' location in the environments. Test results showed that they were able to reach an average prediction accuracy of 94.8% in their office environment, using a time-dependent dynamic model.

Similar methodology and objectives have been employed by Eagle and Pentland. However, in contrast to the previously discussed project, the authors focused on investigation of social relationships, in addition to predicting users' locations [66]. Specifically, three locations were considered: at work, home and elsewhere. Moreover, the Bluetooth neighborhood of the users was not only investigated by using passive Bluetooth beacons. Mobile phones were also scanning, detecting and identifying other users in opportunistic meetings, thus providing more complete information about the users' Bluetooth surroundings. Apart from data related to Bluetooth neighborhoods, the collected information further included call logs, cell tower IDs, application usage and phone status to be used in order to provide insight into both the individuals and their communities. Thus, the acquired data was applied not only to predict user locations, but as well to anticipate the probability of user meetings and define types of relationships between them. The authors employed a Hidden Markov Model [68], which was trained for 1 month and afterwards resulted in accuracy greater than 95% to anticipate users' locations. Bayes' rules were used to predict the probability of user encounters with an accuracy of 90%. And finally, the social proximity context information was used in order to investigate

types of relationships between users. Particularly, the nature of the relationships between the people was deduced by combining users' proximity information with temporal and spatial information. Specifically, it was expected that being near someone in a canteen at 3 pm implies different relationship than being detected in the proximity downtown on a late weekend night. Based on these assumptions, the authors trained a Gaussian mixture model [63], which achieved a prediction accuracy of over 90%, with regards to the identification of social relationships between users.

A.2.2 Users' activities

Identifying users, predicting their locations and types of participants' relations were not the only challenges addressed in ubiquitous computing environments. In [46], the authors further explored context acquisition by attempting to detect users' social activities. They collected context data of two inhabitants using wearable sensors, by monitoring interactions with objects in the environment, and stored it in a SQL database. In order to recognize activities that occurred in the environments, they applied Hidden Markov [68] and Naive Bayesian classifier [126] statistical models. The former model achieved an average activity recognition accuracy of 90%, while the latter one obtained accuracy level of 49%. Thus, the Hidden Markov model can be assumed to be a more effective approach for this type of classification problem. However, it must be noted that attempting to recognize social activities between more than just two users would result in a significant increase of complexity of the system.

Predicting users' activities was also the objective of the CenceMe project [139].

A. PRIVACY AND TECHNOLOGY CHALLENGES FOR UBIQUITOUS SOCIAL NETWORKING

However, the authors used a different methodology than Cook et al to acquire and apply users' context. In fact, they exclusively relied on sensing components of mobile phones to anticipate users' current activities. Particularly, CenceMe uses the following sensing components of smartphones:

- accelerometers to predict the current condition of the user (e.g. sitting, standing, walking, running, etc);
- microphone to recognize quiet or noise environments and conversation between people;
- Bluetooth as social proximity sensor to classify the neighborhood members, by identifying the identity of users nearby and crossmatching it with online social networks data (e.g. friends, strangers, etc);
- GPS to identify the current position of the user and mobility patterns, such as traveling in a vehicle or not, being stationary, walking, running, etc.

Therefore, CenceMe mobile application is capable of predicting the current users' activities by utilizing the combination of data, acquired from the mobile sensing components, listed above. For example, by combining the current location of the user, information from the proximity sensor (e.g. co-workers in the neighborhood) as well as information from the accelerometers (e.g. sitting) and microphone (e.g. talking), it can be predicted that the user is participating in a business meeting.

A.2.3 Users' online profiles

As previously discussed, sensor networks have the possibility to acquire context such as users' identities, relationships and current activities within the USN environment. However, they do not develop awareness of users' interests and preferences, which would present much added-value for such environments. Indisputably, online social networks is a vast source of users' personal information. Thus, by integrating online social networks into sensor networks, USN environments can become more intelligent and context-aware consequently more advanced solutions can be provided [21, 93].

The integration of online social networks and sensor networks technologies could be accomplished by modeling them through semantic web technologies. Subsequently, it would be possible to develop a unified layer of USN context on the top of existing applications. This approach was discussed in studies, such as [107] and [21].

Milluzzo et al [139] investigated the acquisition of the USN context through online social network sites, in addition to the prediction of users' activities, discussed in Section A.2.2. Particularly, CenceMe retrieves the user's list of friends from social network sites, such as Facebook and MySpace, to update them about his current activities. CenceMe presents new opportunities to keep friends close by providing constantly accessible information about users' locations and activities.

Another approach, which exploits data from online social network sites, is the Astra project [135]. Astra is based on Service Oriented Architecture Principles [74], which enables effective service delivery in dynamic environments. Particularly,

A. PRIVACY AND TECHNOLOGY CHALLENGES FOR UBIQUITOUS SOCIAL NETWORKING

Astra has the purpose to accommodate users' wishes through available network technologies [135]. A model to manage the users' information and preferences was implemented, based on the concepts of Nimbus and Focus [31]. While Nimbus represents the information that the user would like to share, Focus represents the set of information that the user is interested in receiving. Thus, a customized response to each user's preferences can be created by exploiting the USN context gathered from smart objects and sensors in the users' surroundings. According to users' preferences, automatic updates about the users' availability for interaction (e.g. going for a walk or a phone conversation) can be delivered to the community of the users (e.g. friends, family members, etc).

A.3 Software architectures

As noted in the previous section, some studies already acquired and applied USN context from online social networks in order to enhance social network services in the physical environment. Specifically, Astra [74] and CenceMe [139] focused on keeping friends and family members close by updating the user about their current activities, locations, availability, etc. However, these approaches cannot be identified as "networking" due to their principal focus on existing acquaintances. As earlier introduced, USN must be capable of promoting sociability not only among friends, but between strangers as well. Consequently, we are going to review different studies that aim at promoting USN between people in the physical proximity. Notably, the reviewed studies are presented according to the type of software architectures, i.e. centralized, decentralized and hybrid.

A.3.1 Centralized architecture

In this section we present solutions based on centralized architecture that promote social interactions among USN participants. The first approach that we describe is proposed by Hasswa and Hassanein [93] who promoted USN between users with similar interests. The authors utilized mobile devices both as wearable sensors and as gateways to social networks in order to obtain USN context. The proposed architecture comprises Internet access points in order to enable the USN environments, as it is shown in Figure A.4-A. When mobile phones connect to the access point, users give permissions to retrieve their data, stored in online social networks. Thus, users' preferences, interests, personal data, list of friends become available for that specific USN environment, which in turn is able to provide personalized services to its inhabitants. For example, the USN environment is capable of classifying the participants into different groups based on relationships and similar interests, and consequently promote networking among the identified groups (e.g. exchange of

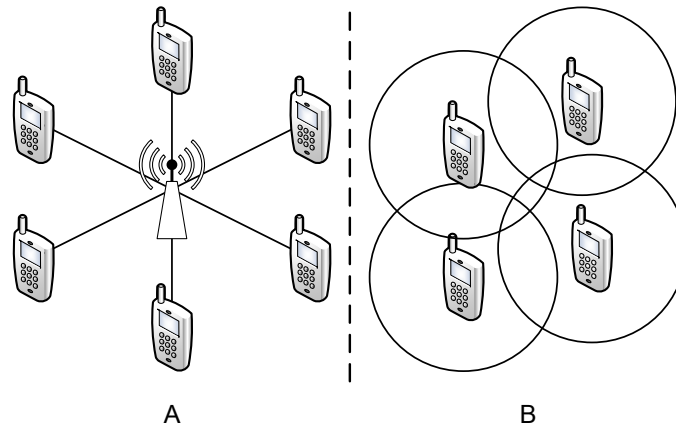


Figure A.4: Methods for establishing connections among inhabitants of USN environments

A. PRIVACY AND TECHNOLOGY CHALLENGES FOR UBIQUITOUS SOCIAL NETWORKING

public, private and group messages, pictures, etc).

As shown in Figure A.4-A, Hasswa and Hassanein's work is based on an architecture, which exclusively relies on a central unit. In this case, the USN environment is limited to a particular physical location and the offered services are restricted to a certain network [172]. However, new software architectures were inspired by the elevated mobility of mobile phone users that could potentially enable unrestrained USN environments [33]. We further investigated USN that is established during purely ad-hoc meetings, as shown in Figure A.4-B. This solution leads towards sociable opportunistic networks, where nodes are wirelessly interconnected and have the possibility to identify each other [95]. Sociable opportunistic networks contribute to addressing sociability issues by enabling dynamic, circumscribe and mobile USN environments, applied in everyday physical world.

MIT Serendipity project [65] was the pioneer of sociable opportunistic networks. The software architecture of Serendipity is shown in Figure A.5. When Serendipity users randomly meet, they exchange their Bluetooth identification (step 1). This information is sent to a central server (step 2), which contains all the Serendipity users' profiles along with the matchmaking preferences. The server evaluates similarities between encountering users (step 3), and if the similarity score identifies a mutual match, the server notifies both users about their presence and the related affinities (step 4).

Adopting a similar approach is the MobiSoc [83] mobile application, implemented by Gupta et al. When MobiSoc peers meet, the application captures and manages the social context of physical communities such as social profiles, people-to-people and people-to-place affinities. Further, MobiSoc exploits learning algorithms

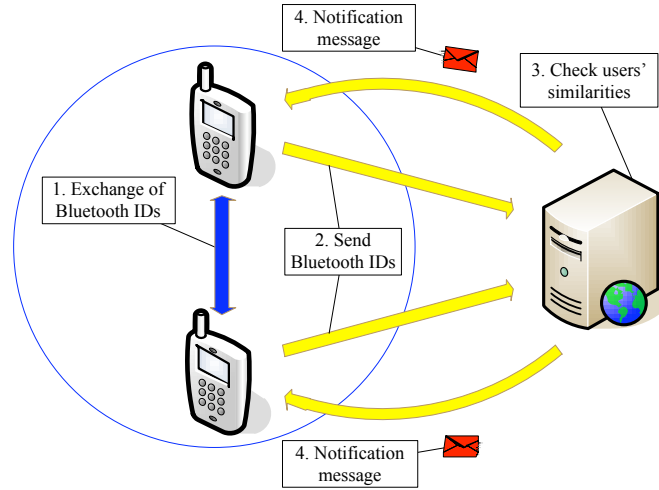


Figure A.5: Centralized solution for ad-hoc USN environments

to notify users about relevant matches according to their preferences.

Other works, which have also investigated sociable opportunistic networks that rely on a central unit are Cityaware [114] and the Wireless Rope [147]. The former aims at displaying users' encounters by storing proximity data on a central server and then visualizing them through a Facebook application, while the latter targets at analyzing users' personal social networks. Specifically the Wireless Rope collects data of other Bluetooth devices in the proximity and stores it in a database for further analysis. Consequently, Wireless Rope users are able to investigate their neighborhood and identify the relations with other users in their proximity.

A.3.2 Decentralized architecture

Even if relying on opportunistic meetings, the previously discussed solutions still comprise interactions between the nodes and a central server, as shown in Figure

A. PRIVACY AND TECHNOLOGY CHALLENGES FOR UBIQUITOUS SOCIAL NETWORKING

A.5. Recently, those centralized solutions were replaced with new decentralized approaches, which exploit dynamic mobile connectivity in terms of peer to peer communications. In fact, peer to peer solutions are considered to be more suitable for USN environments, due to the fact that inhabitants engage in opportunistic ad-hoc social interactions [186].

Figure A.6 shows an example of peer to peer communication [172] on mobile phones, using the Bluetooth technology. Each node of the peer to peer network can simultaneously play two different roles: server and client. The task of the server is to publish a service and accept concurrent connections, whereas the task of the client is to search and connect to services in order to exchange data [105].

One of the well-known implementations of decentralized architecture was Nokia Sensor [153], which relies exclusively on opportunistic connections between devices nearby. Using the Bluetooth technology, similarly to Figure A.6, Nokia Sensor users are able to discover each other within a short communication range and exchange users' contents, stored in the local memory of the mobile phones (e.g. profiles, pictures, etc).

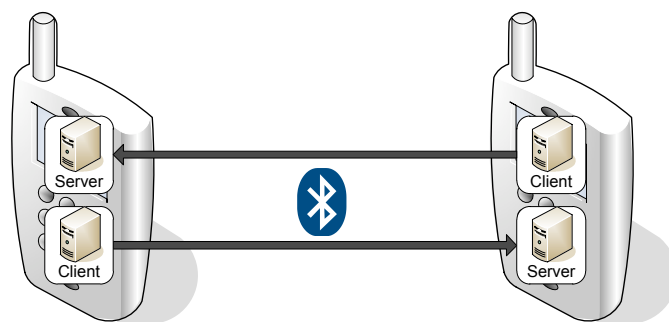


Figure A.6: Decentralized solution for ad-hoc USN environments

A similar approach was proposed in [183]. The authors presented an application for mobile devices called BlueFriend. Comparably to Serendipity [65], the main goal of BlueFriend is to match interests of users when they encounter each other, while relying on a decentralized architecture. To achieve this goal, the application splits user's personal information into two main categories: profile and personal data. While the former includes public data (e.g. users' preferences), the latter consists of data that users generally prefer to keep private (e.g. contact details, user's description and photo). Consequently, when BlueFriend users encounter each other, they firstly exchange their users' profiles in order to find similarities. If the matching index is found to be above a threshold value, defined by both users, then they also exchange their personal data.

While the main advantage of the previous described decentralized approach is the possibility to directly communicate without requiring a third party, the major limitation is the restriction of the communication range. In fact, these solutions utilize a Bluetooth communication between the users, which is limited to an approximate range of 10 meters. To overcome those limitations, in [72], the authors proposed a mobile application, called ScatterFriend, which focuses on extending the reach of Bluetooth. Einarsdottir and Li applied the Beddernet [81] peer to peer middleware in order to enable ad hoc networks consisting of two or more piconets, called scatternets. Figure A.7 shows examples of a piconet and scatternet.

A piconet is composed of two or more connected Bluetooth devices, as shown in Figure A.7-A. The device which starts the connection is called master while the other device is called slave. Master/Slave is a model for communication protocol where one entity (master) controls other entities (slaves). Importantly, only one

A. PRIVACY AND TECHNOLOGY CHALLENGES FOR UBIQUITOUS SOCIAL NETWORKING

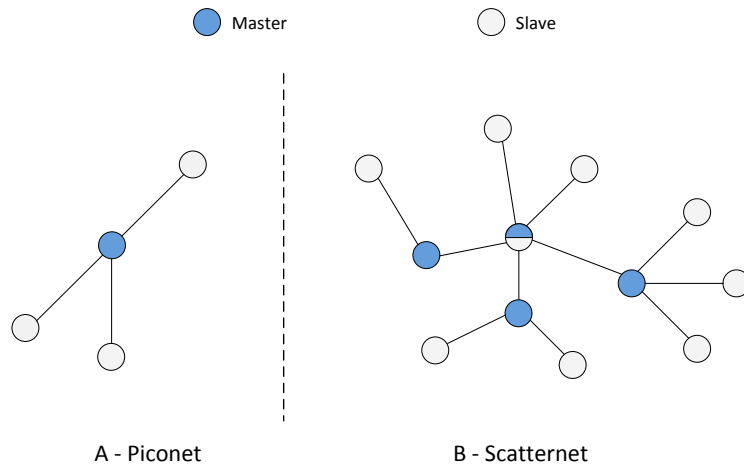


Figure A.7: Bluetooth piconet and scatternet

device can play the role of master and this device can be connected at maximum to 7 slaves. A scatternet, shown in Figure A.7-B is composed of interconnected piconets. Specifically, scatternets are formed when a device of one piconet (either the master or a slave) connects as a slave in a second adjacent piconet. This device is considered to be the linking point between the two piconets, as it can communicate with participants of both piconets. Due to application of scatternets, ScatterFriend users are able to exchange their personal contents, such as contact invitations, public and private messages, not only over direct links, but also over multihop links within the scatternet.

The ScatterFriend was implemented on the Android platform and tested by first users. Despite the perceived usefulness of ScatterFriend, the participants indicated the stability of the network connection and time consumption of the Bluetooth discovery as key weaknesses of the application [72].

A.3.3 Hybrid architecture

Among the sociable opportunistic networks, there are other solutions that rely on hybrid architecture, enabling data portability of social networks. Specifically, user profiles of these solutions are stored in online social networks and synchronized with the local memory of the mobile devices. However, in the offline mode, the application does not rely on any central unit, but it utilizes the decentralized architecture, shown in Figure A.6, in order to discover and identify users as well as exchange personal contents [155, 171].

As illustrate in Figure A.8, a hybrid architecture is based on an integration of online social networks and opportunistic networks. Importantly, such an integration enables mobile social network users to exploit social networking benefits in the physical world, rather than only in the virtual world. This integration has been previously introduced as local social networks [171], in which physically close nodes are linked to online social networking profiles and wirelessly interconnected

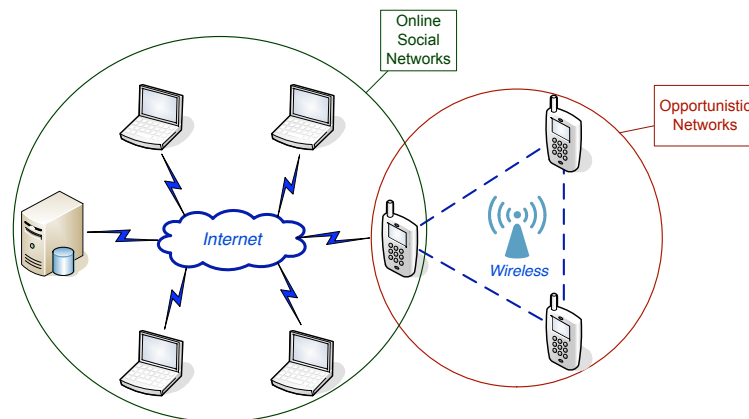


Figure A.8: Hybrid solution for ad-hoc USN environments

A. PRIVACY AND TECHNOLOGY CHALLENGES FOR UBIQUITOUS SOCIAL NETWORKING

to exchange personalized contents [171]. Specifically, local social networks aim at providing USN services between users with interpersonal affinities, while preserving users' privacy. In fact, users in the physical proximity exchange only sub-profiles consisting of personal information that is relevant, but not sensitive, for the specific circumstances of the encounter [171]. The first prototype of local social networks is called Spiderweb [166], which was implemented in Java 2 Micro Edition on Symbian OS [87]. In order to evaluate users' acceptance of local social networks, the Spiderweb services (presented in [166] and this video¹³) were tested according to Technology Acceptance Model [56] by investigating the perceived ease of use and usefulness of the application. Both tests returned satisfactory results, with users finding Spiderweb to be an interesting and innovative application, which is also very easy to use. Moreover, the majority of the respondents claimed that they would be potential users of those USN services, despite the indicated limitation of the time consumption of the Bluetooth discovery. Particularly they emphasized and appreciated the fact that social networking was no longer in front of a stationary PC, but these technologies were disappearing in the background.

Another hybrid architecture solution is the MobiClique [155] mobile social network, which focuses on dissemination of contexts such as user profiles, private and public messages. Unlike ScatterFriend [72], MobiClique does not implement scatternets, but it relies only on direct connections between the nodes for data dissemination. In case the recipient of the message is not in the direct range of the sender, MobiClique supports a forwarding mechanism, based on the mobility of intermediary nodes: users physically carry the message while being on the move.

¹³<http://www.youtube.com/watch?v=DgeVNv10CIM>

In order to limit the number of messages to be carried by the intermediary nodes, some restrictions exist, such as messages must be carried either by members of the target group of interest or by friends of the recipient [155].

A.4 Privacy design guidelines

Even when ubiquitous computing was just a vision, privacy threats were already identified as the greatest barrier to the long-term success [99, 191]. Nowadays, sensors are capable of acquiring not only environmental data, but obtaining users' personal information as well. Thus the technological development is moving towards people-centric era, where humans are the main focus of sensing. In people-centric sensing, users are parts of mobile sensor networks, where mobile devices are conceptually tied to individuals. Consequently, new challenges arise for privacy of USN, which specifically include safe collecting, storing, processing and disseminating users' personal information, related to daily users' activities [106].

The management systems of data privacy must acknowledge that privacy is not considered anymore as having "the right to be let alone" [19], but it is now understood as having "the right to select what personal information about me is to be disclosed and to whom" [196]. Therefore, the main challenge is shifting from hiding personal data to ensuring successful management of disclosure of users' personal information.

In [100], the authors discussed two different aspects for the management of users' privacy: *data protection* and *personal privacy*. The former typically refers to the management of identifiable personal information by third party entities, which

A. PRIVACY AND TECHNOLOGY CHALLENGES FOR UBIQUITOUS SOCIAL NETWORKING

regulates how, when and for which reasons data can be collected, processed and disseminated [196, 197]. On the contrary, personal privacy is related to the selection of personal data to be shared with other individuals or organisations. This concept cannot be seen as a static notion, where users set rules and enforce them, but rather it should be considered as a dynamic process, representing continuous negotiation and management of the boundaries that shape personal data disclosure [4, 5, 149].

Importantly, Iachello and Hong discussed that data protection and personal privacy should not be considered as adversary concepts. In the case of USN technology, both concepts are crucial for ensuring its long-term success. For example, a USN application should include data protection solutions for securing safe collection and dissemination of users' personal information. As well, it needs solutions regarding personal privacy for helping users to accurately select the right information to be disclosed to the right person under the right circumstances [98]. In the following we describe existing privacy protection laws and review previous works that adapted these laws to privacy system design in regard to data protection. Further, we focus on privacy design guidelines for ensuring users' personal privacy, based on both understanding of the privacy implications for participation in USN as well as the possibilities to conduct socially meaningful action through USN.

A.4.1 Data protection

The protection of users' data privacy has been already taken into consideration in the legal as well as academic worlds. The legal regulations draw the primary framework for privacy preservation in USN. However, when interpreting them, the usability and applicability must be also considered. In fact, unsuccessful design

of application of legal regulations could result in threats to the success of USN environments, as the user could be overwhelmed by unnecessary privacy protection measures.

Over the past quarter of century, the information practices, i.e. the ways entities collect and use personal information, have been extensively studied by government agencies in United States and European Union. Two main privacy protection laws have been enacted: European Union's Directive 95/46/EC [62] and US Privacy Act of 1974 [82]. Based on the established legal framework, the core data protection principles can be elicited [120, 202, 204]. Thus, in the following we are going to group these principles into three categories as well as proposing potential design implications for USN environments.

Notice and explicit consent

The notice is legally interpreted as prohibition of collection or storing any secret records of any personal information; thus the entity must notify the individuals about their personal information policies and practices. However, as enforced by the European Union's Directive 95/46/EC, data collection is not anymore approved only by announcement and declaration. Actually, the entity is legally permitted to collect and process the personal data only if the individual has given his or her explicit consent on a case-by-case basis [62, 69, 82].

In terms of USN design, the notice and consent legal laws constitute recommendations that privacy models would demand the attention of the individual under privacy threatening circumstances. Specifically, before acquiring, processing and sharing personal data in USN environments, the data owner must actively decide

A. PRIVACY AND TECHNOLOGY CHALLENGES FOR UBIQUITOUS SOCIAL NETWORKING

whether and which data about her should be collected [120, 204].

However, rigid rules ensuring choice and notice principles might not be possible for efficient and reliable implementations of USN environments. They would probably result in too many interactions between the users and systems for either approving users' data disclosure decisions or informing when the collection and dissemination of their personal information occur. While these principles are crucial for ensuring better users' privacy, they must as well be designed to embrace Weiser's vision of calm technology [191, 193]. For example, designing USN environments that would adopt e-privacy policies for allowing users to specify particular circumstances or potential sensitive data that would require either a notification of disclosure or an approval before the actual disclosure might considerably reduce such interaction between the users and the system.

Purpose, access and use limitation

Before the actual data acquisition, the entity must define and disclose its goals of collecting the data and further act on basis of those purposes. The entity is legally bound to collect only data, which is essential for the identified purpose and it can be retained only as long as necessary. Additionally, it must ensure that the personal data is accurate, relevant and fairly up-to-date in respect to the defined goals. Finally, the entity must enable the individual to access their personal data and, if necessary, allow to modify and revoke the contents of this data [62, 69, 82].

As a design requirement, USN environments should guarantee their inhabitants that the collection, processing and dissemination of their personal data is used only for a well defined and transparent purpose, i.e. there should be no "in-advance"

storage of personal data. Moreover, the USN environments should restrain themselves from attempting to collect any other personal information that is not directly relevant for the defined goal. Finally, USN environments should remove inhabitants' personal data as soon as the ultimate goal has been achieved [120, 204].

Notably, if these principles are well applied, it might save many resources that would be otherwise utilized to protect, collect and manage large amount of sensitive personal data. Moreover, the application of these principles would increase privacy protection benefits, as it would result in collection of only the data, which is essential for defined targets. This principle is of crucial importance for potential future privacy threats, because a set of data, given up freely today, might create major user's privacy concerns in the future. In order to avoid such privacy protection failures, the users must be ensured that their personal data will be erased once the purpose is achieved [120].

Data security

In order to secure individuals' personal data in the digital world, laws are enforcing adequate security measures to be adopted in accordance to the sensitivity of data collected. Entities are directly accountable for ensuring data security through their internal mechanisms as well as methods of personal data management. Moreover, different means for ensuring compliance of the security standards have been determined by the legal regulations. Firstly, the compliance mechanisms can be enforced by self-regulation, where a certain code of fair information practices is adopted by the participants of the industry. Also, the legal measures can be taken by the government authorities in order to ascertain the protection of data privacy

A. PRIVACY AND TECHNOLOGY CHALLENGES FOR UBIQUITOUS SOCIAL NETWORKING

[62, 69, 82].

In regard to design principles, Langheinrich suggested cryptography as one of the main mechanisms in order to provide confidentiality in information storing and distribution. However, cryptography requires availability of large amount of resources, which are not always accessible in USN environments. Thus, the suggested solution was to adjust the extent of cryptography proportionally to the sensitivity of the data. Other possible design solutions, which would overcome recourse limitations of USN environments, are the following [120]:

- Anonymity and pseudoanonymity: to remove or drastically minimize the connection between the data owner and the information;
- Location and proximity: to limit further dissemination of information after the initial data disclosure.

First, user should have the opportunity of not being identifiable in USN environments. Specifically, they should have the freedom of remaining anonymous when desired. On the other hand, some situations would present necessity of authentication (e.g in case of personalization of services). For such cases, Langheinrich suggests pseudoanonymity techniques as a possible design solution. The pseudoanonymity could be realized through linking the user to an ID, which would represent him in specific circumstances in order to be recognized as long as he uses the same ID. Thus, a user could build a history of that specific ID and consequently, in these cases, personalization of services would be feasible.

Finally, the proximity and locality design solution tied distribution of the disclosed information to the owner's (or witness's) presence and physical location re-

spectively. The former allows only the owner and the witnesses of data disclosure to further disseminate the user's personal information. However, this solution could be considered as unacceptable in some situations, e.g. when the owner does not approve further distribution of the data. Thus, the notion of physical location arises: data is only available to the location at which it is collected and thus only the users who are physically present in that location can collect such information [120].

A.4.2 Personal privacy

The legal regulations, outlined in section A.4.1, define the principles for designing trustable USN environments. Although there is no law enforcing such principles, it could be recommended that the legislators incorporate them into their existing privacy laws. Further, the essence of sociability and networking in such environments demands selection of relevant, but not sensitive, information to be disclosed to others. Thus, the design of privacy management of USN are challenged to provide solutions for reflecting on people's natural privacy handling, by ensuring management of users' personal privacy - the processes by which people selectively share personal information, such as email address, career skills and abilities, to other individuals or organisations [100, 122].

Previous studies have already proposed relevant design guidelines to prevent potential privacy threats that could discourage users from providing their personal information in ubiquitous computing environments [12, 104, 120, 122, 149]. These guidelines target at empowering users to make deliberate personal data disclosure decisions, which is also a crucial goal for USN environments. On this matter, Lederer et al [122] suggested that the socio-technical gap, introduced by Ackerman

A. PRIVACY AND TECHNOLOGY CHALLENGES FOR UBIQUITOUS SOCIAL NETWORKING

[1], should be addressed. The socio-technical gap refers to the division between “what we know we must support socially and what we can support technically”. If an intermediary point of the socio-technical gap is not found, the user would be either overwhelmed or disempowered, which would both result in uninformed and impulsive data disclosure choices. To find this balance, Lederer et al [122] proposed that the design of such environments should be focused on enabling users to both understand privacy implications of data disclosure as well as allowing them to perform natural social actions. Consequently, they provided privacy guidelines that target at reconciling Palen and Dourish’s theoretical insights with Bellotti and Sellen’s technical solutions [12, 149].

The theoretical insights of Palen and Dourish were inspired by the work of Altman, who describes privacy as a dynamic process, representing continuous negotiation and management of the boundaries that shape data disclosure [4, 5]. Palen and Dourish identified three dynamic boundaries for negotiation of users’ personal data disclosure. Firstly, the privacy and publicity boundary separates personal information into the disclosed and retained data sets. Then, the identity boundary defines the role, represented by the user based on the time, place and situation contexts. Finally, the temporal boundary regards the past, present and expected future of the users. The authors concluded that data disclosure decisions are taken by continuously negotiating the internal conflicts between the elements of the three identified boundaries. Bellotti and Sellen, instead, focused on more practical solutions. They firstly introduced potential privacy threats, such as disembodiment and dissociation. The former one refers to the danger that users would not be able to present themselves to others as they do in face-to-face interactions. Dissociation

refers to the threat that the results of actions are shown while the actions themselves are invisible. Consequently, they proposed control and feedback principles in order to avoid disembodiment and dissociation privacy threats by allowing the user to decide what to disclose and whom to disclose, as well as ensure subsequent feedback about their data disclosure decisions.

Moreover, in the privacy guidelines introduced in [122], the authors try also to honor the fair information practices, outlined by Langheinrich as well as attempt to encourage minimum information asymmetry between the parties, i.e. data owner, data collector and data user [104]. Information asymmetry is considered to be the imbalance in the amount of data flow between data owner and data collectors/users. The authors suggested to either decrease the flow of information from data owners to data collectors and users or otherwise increase the flow of information back to the data owner. Particularly, based on the legal regulations of US Privacy Act of 1974 [82] and European Union's Directive 95/46/EC [62], Langheinrich identified several main areas of innovation, already reviewed in Section A.4.1 and recalled in the following: notifying the user appropriately; taking into account the user's choice and seeking for consent; enforcing limitation of scope within the concepts of proximity and locality; enabling anonymity and pseudonymity when necessary; providing adequate security and appropriate data access.

As a result, the privacy guidelines proposed by Lederer et al depict five pitfalls to be avoided in the design of privacy management systems for such environments:

1. Obscuring potential information flow: USN should not obscure the nature and extent of data disclosure. Users should easily comprehend, for example, what kind of information is disclosed and to whom, how the information is shared,

A. PRIVACY AND TECHNOLOGY CHALLENGES FOR UBIQUITOUS SOCIAL NETWORKING

the presence of third-party observers and the potential for unintentional disclosure. Avoiding this pitfall would allow users to understand the scope of the privacy implications;

2. Obscuring actual information flow: USN should not obscure the actual disclosure of information. The disclosure should be obvious to the user as it occurs, however without overwhelming his attention. When immediate notice is not feasible, then it must be ensured with a reasonable delay. Avoiding this pitfall would allow users to understand what information is being disclosed to whom;
3. Emphasizing configuration over action: USN should not require exaggerated manual configuration to manage personal privacy. Instead, users' privacy should be managed as a natural consequence of their normal engagement with the environments. Avoiding this pitfall would enable users to control their privacy without requiring tremendous configuration;
4. Lacking coarse-grained control: USN should not be implemented without the possibility for their users to halt and resume data disclosure when desired. Avoiding this pitfall would empower the user to effectively control their participation in such environments;
5. Inhibiting established practice: USN should not inhibit users from transferring established social practice to emerging technologies. For example, USN should enable disclosure of ambiguous information as well as should ensure plausible deniability. Avoiding this pitfall would allow the users to participate in such environments without compromising their ordinary social behavior.

A.5 Privacy management models

The privacy protection models of USN should be designed to follow the design guidelines of privacy management systems presented in Section A.4. Particularly, the suitable privacy models for USN environments should follow the legal requirements, while still facilitating the data disclosure and as well following the natural data privacy handling of the individuals.

A privacy protection model, which followed the outlined legal framework of data privacy, was implemented in [204]. Particularly, the author provided a solution where environments are owned by an entity, which sets a framework of information policies, established according to the legal requirements. Moreover, each user is given an opportunity to personalize his privacy policy, by specifying what he is interested to share (observe) and under what terms. Thus, before any actual interaction with the environment, the user is asked to present his own set of information policies to be verified by the entity. If the policy is found to be compatible, the user is permitted to interact with the environment, otherwise a negotiation is attempted.

While primarily addressing privacy in terms of legal regulations, Yee's model requires more attention to the dynamic nature of personal data disclosure. Other privacy protection solutions [103, 111, 144] presented more extensive considerations of personal data disclosure by incorporating possibilities for dynamic sharing choices. Specifically, these models were implemented by enabling users to share varying personal data subsets, depending on the location of the inquiry.

However, Sadeh et al proved that users encountered difficulties while selecting their privacy preferences according to the single "location" factor [165]. Lederer

A. PRIVACY AND TECHNOLOGY CHALLENGES FOR UBIQUITOUS SOCIAL NETWORKING

et al conducted a survey, which investigated the important determinants for data disclosure decisions. Particularly, the research was focused on the inquirer and the current situation parameters. The inquirer is considered to be an individual whom the user is interacting with and the situation is defined according to the circumstances at that time. The authors determined the identity of the inquirer to be the most important value for influencing the users' privacy choices, followed by the situation as parameter of secondary significance [123].

Based on these findings several privacy models have been designed for disclosure of personal information. Notably, two main approaches were applied for managing disclosure of personal information in USN environments: predefined privacy preferences and ad-hoc privacy control. The former attempts to predict all the potential circumstances and associated data sharing decisions priori to the actual data disclosure. Further, these models were upgraded to ad hoc privacy control solutions, which target at limiting the predefined privacy preferences and thus support *in situ* data disclosure decisions, i.e. taken at the moment of actual disclosure. Consequently, in the following we review an example of predefined privacy management models and later we concentrate on methods that apply ad hoc privacy control of users' personal information.

A.5.1 Predefined privacy preferences

The management of data disclosure firstly focused on predefined privacy preferences. Priori to any data disclosure, users were asked to indicate "who" can access "what" and "when". The privacy model Faces [124], as well as other solutions [42, 103, 111, 144], apply predefined privacy preferences for management of personal information.

A.5 Privacy management models

Faces was inspired by the work of Goffman, who observed that an individual is inclined to present himself to a certain audience by undertaking a particular role or face and he will attempt to maintain that chosen face throughout the time [80]. Faces supports four predefined levels of privacy protection, ranging from “undisclosed” that defines absolute confidentiality to “precise”, which allows openness of entire user’s personal information. The process of Faces privacy model is shown in Figure A.9.

As the first step, users set the privacy rules through a desktop application, by creating 3-tuple of inquirers, situations and faces (Figure A.9-A). Afterwards, when the user meets one of the predefined circumstances, the Faces repository, where all the predefined preferences are stored, is queried. Based on the input of current situation and inquirer, the repository returns the corresponding face to be presented for the inquirer (Figure A.9-B). Additionally, in case of unknown inquirers

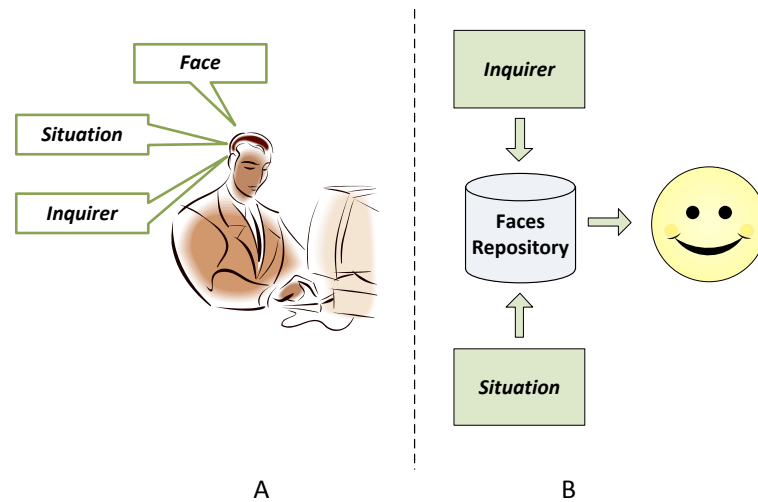


Figure A.9: Faces privacy model process

A. PRIVACY AND TECHNOLOGY CHALLENGES FOR UBIQUITOUS SOCIAL NETWORKING

or situations, a default face is returned.

Faces was tested by first users, who indicated that the predefined privacy preferences was not an accurate solution for guiding data disclosure in USN environments. Specifically, the users encounter difficulties in predicting their own data disclosure preferences precisely, as sometimes they wished to adjust their decisions while meeting the actual circumstances. Even more importantly, some of the users were not able to remember their predefined preferences, which further highlights the complexity of the privacy model. Thus, the test results imply that privacy models applying predefined preferences could not prevent invasion of privacy when the actual/real preferences do not meet the rules indicated by the user priori to data disclosure [122].

A.5.2 Ad hoc privacy control

The Faces privacy model [124] does not take into account the dynamic nature of data privacy, as it relies on preconfigured static privacy preferences. Inhabitants of USN environments might encounter situations where data disclosure decisions are not accurately predictable. Thus, in those circumstances, data disclosure decisions can only be made ad hoc - at the moment of the actual disclosure [25, 26, 27, 103]. In order to address these issues, Lederer et al upgraded the Faces concept into the Precision Dial model [122]. In comparison to Faces, Precision Dial removed the preconfigured privacy preferences and added quick manual selection of one of the four privacy protection levels, introduced in the Faces privacy model. While encountering different circumstances, the user has the opportunity to manually adjust his privacy settings when needed. However, despite achieving the primary goal of

enabling ad hoc data disclosure decisions, Precision Dial might require considerable amount of users' attention and intervention. In fact, users are required to manually adjust their precision settings, similar to the practice of adjusting ringer volume of mobile devices [122]. In case the user would forget to update the current privacy protection level when encountering other circumstances, it might result in unintentional data disclosure and the user could either be subject to invasion of privacy or lose potential networking possibilities.

Similarly, Bünning provided a solution for reducing predefined privacy preferences for data disclosure management [26]. The author suggested a privacy model called Disclosure Decision Model (DDM) that focuses on relieving the users from frequent data disclosure decisions. DDM [24] can be considered as an agent that manages information disclosure on behalf of the user through a process shown in Figure A.10. The first step of the DDM model is taken by the users, who manually set disclosure rules that are generalized privacy protection guidelines for particular circumstances. These rules are the primary determinant for the data disclosure decisions (step 2). However, if the general predetermined rules do not directly apply for a particular situation, the DDM is consulted for that particular disclosure choice (step 3). In case of user's disagreement with the automated disclosure decision of a specific data set, the design allows a manual veto possibility, which highlights a wrong disclosure choice, taken by DDM (step 4). Consequently, the DDM sends to the user a rule template, which enables the user to manually deal with the current circumstances (step 5). Thereafter, the fulfilled rule template will be added to the existing set of rules for improving future prediction of data disclosure decisions.

In order to investigate different data mining algorithms to be applied for devel-

A. PRIVACY AND TECHNOLOGY CHALLENGES FOR UBIQUITOUS SOCIAL NETWORKING

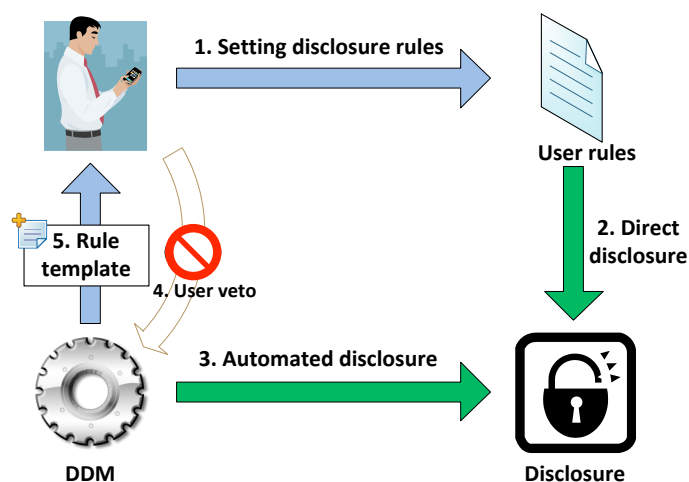


Figure A.10: Disclosure decision model process

oping DDM, Bünning asked test users to set their hypothetical privacy protection guidelines. Subsequently, an environment was implemented and used as a test environment for simulating potential service providers, requesting users' personal information [27]. Among data mining algorithms tested such as Rule Learner, Decision Tree, s-Nearest Neighbour, SVM and naive Bayes classifier, test results have indicated that naive Bayes classifier [126] was the most suitable algorithm, as it presented highest results with an approximate accuracy of 95% and potential for further increasing performance [25].

Furthermore, in [171], the authors suggested a privacy model called Diverged Personalities (DiP), which was originally designed for the local social networks environments, already discussed in Section A.3. Similarly to DDM, DiP as well focuses on reflecting the user's natural privacy handling upon disclosure of personal data without overwhelming the user's attention. However, in contrast to the previous privacy model, DiP does not rely only on observing and learning users' decisions,

but also takes into account the variation of personal data sensitivity under different circumstances.

In the DiP privacy model, the unified user profile is diverged into different users' personalities to be presented under different circumstances. The most suitable personality for each circumstance is generated by the process shown in Figure A.11. The central component of the Diverged Personalities is the Personality Logic. The Personality Logic receives as input the unified user profile, which is composed of a collection of various personal data about the user. Moreover, the Personality Logic also processes the inquirer social information (e.g. stranger, co-worker, familiar stranger, friend, etc) and context information (e.g. current activity, location, time etc). Based on these inputs, algorithms automatically provide the best personality to be shared, which can be compared to a business card that people often exchange for interpersonal benefits. The selected business card must be composed of relevant, however not sensitive, personal information according to the current circumstances [170]. When selecting the best personality to be disclosed, the Personality Logic

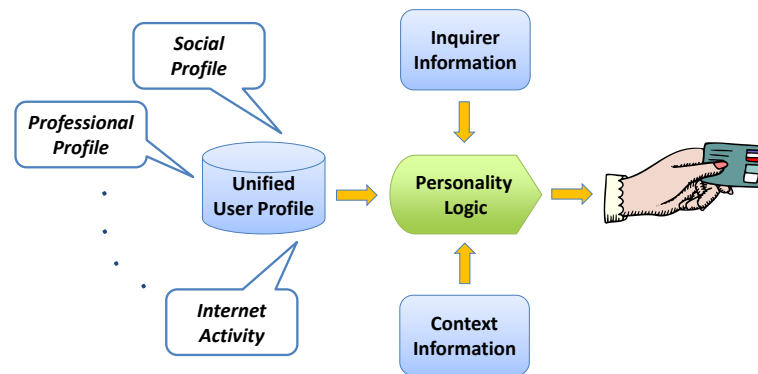


Figure A.11: Diverged personalities model process

A. PRIVACY AND TECHNOLOGY CHALLENGES FOR UBIQUITOUS SOCIAL NETWORKING

takes into account influential factors, identified to impact users' data disclosure decisions, such as identity of the inquirer [57, 123, 199], current environment [169], activity [108, 167, 169, 170], mood [45, 169], location familiarity [167, 169, 170], purpose of disclosure [45, 167, 169], number of previous meetings and mutual friends with the inquirer [167, 169].

In [169], the author conducted a mixed methods research for investigating the relationship between the above influential factors and ad hoc data disclosure decisions. Participants' ad hoc data disclosure decisions were collected by exploiting an USN prototype. The collected data, i.e. at least 1650 data disclosure decisions per each participant, was analyzed by applying the binary logistic regression statistical model. Thus, it was possible to simulate the DiP privacy model behaviour and present results about the overall proportion of cases that the binary logistic regression statistical model classified correctly. The achieved overall prediction success was observed to be approximately 90% with a peak accuracy of 93%, and a potential for further increasing performance.

However, despite the good prediction results achieved by the DDM and DiP privacy models, in [25, 169] the authors recommended to limit the autonomy of ad hoc privacy control mechanisms, in case of highly sensitive data. In such situations, it was emphasized to provide only suggested data disclosure choices, while waiting for user's approval before any actual disclosure.

A.6 Discussion

This paper identifies privacy as the main challenge for the development of ubiquitous social networking environments that target at promoting social interactions among their inhabitants. Particularly this challenge arises because privacy must be ensured during acquisition, management and disclosure of users' personal information. We described existing privacy protection laws and reviewed previous works that adapted these laws to privacy system design principles. Further, we presented existing privacy design guidelines for ensuring both understanding of privacy implications of participation in such environments as well as possibilities to conduct socially meaningful actions. Afterwards, we introduced existing privacy models based on predefined privacy preferences as well as ad-hoc privacy control approaches and suggested the adoption of the latter approach for ubiquitous social networking environments due to better support of the dynamic and intuitive aspects of protection of data privacy.

While considering privacy as a major challenge for the further development of USN environments, the article also identified and discussed the other two challenges that are related to ensuring privacy-aware social networking. In fact, we reviewed methods and technologies for acquisition of the relevant context in order to promote sociability between participants, such as users' identities and relationships, users' activities as well as personal users' information from online social networks. Moreover, we described diverse approaches for enhancing social networking on different architectures, i.e. centralized, decentralized and hybrid, with the main focus on promoting networking not only among acquaintances, but also between strangers

A. PRIVACY AND TECHNOLOGY CHALLENGES FOR UBIQUITOUS SOCIAL NETWORKING

with interpersonal affinities. The key findings of the three identified challenges are summarised in Table A.1.

Based on the results, presented in this article, we identified two different aspects for the management of users' privacy in ubiquitous social networking, i.e. data protection and personal privacy. We acknowledged that both concepts should be taken into consideration for the long term success of the ubiquitous social networking technology. However, it must be noted that managing users' data privacy according to data protection guidelines would lead to setting rigid rules and policies to prevent potential malicious attacks. For applications that focus on leveraging interpersonal affinities in physical environments, rules that regulate users' data disclosure under different circumstances are not straightforward and they cannot be algorithmically modeled using rigid privacy policies. Instead, we agreed with the position of Iachello and Hong, who claimed that the design of these environments should focus on enabling a constant and arbitrary selection of disclosed data for allowing users to easily present a desired profile to be shared under different situations [100].

Thus, we suggest further research to investigate solutions for ensuring users' personal privacy. Personal privacy in ubiquitous social networking should be analysed in two different directions, as the selection of users' personal information to be disclosed in such environments occurs in two different moments. The first selection regards whether users' personal data should or should not be included in the unified user profile. Such kind of data disclosure are principally made before actual use of the systems and would be highly influenced by designing services that would increase users' comfort through sharing of their personal information. The design guidelines, proposed by Lederer et al, were indisputably considered rele-

vant, as they lead to the design of ubiquitous social networking environments that would empower users to make informed data disclosure decisions. However, other important design factors are needed to be identified when taking into account the particular focus on promoting privacy-aware social networking. For instance, we acknowledged the importance of informing users about future potential implications of their data disclosure, e.g. informing them on whether their data is shared with third parties [122]. On the other hand, we believe that designers should investigate solutions for also preventing such implications, e.g. not embracing third parties, in order to increase users' comfort through sharing of their personal data.

Further, other design characteristics include advantages and disadvantages of software architectures, reviewed in this article. For instance, a pure peer-to-peer approach might present advantages that would motivate users' to disclose their personal information, as disclosure of personal data would only occur between end users by exactly replicating an exchange of paper business cards. However, the design of such decentralized architecture as well leads to a crucial disadvantage. If users can only wirelessly exchange their personal profiles in the close proximity, they cannot thereafter modify them. On the contrary, a centralized architecture might be ideal for enabling users to modify their personal data, even after actual disclosure. In fact, the third-party entities might store users' personal data disclosure decisions to be accessed, updated or even removed at any time. This kind of solution would certainly require embracing disclosure to third-party components, which might discourage users from providing their personal information, due to concern about potential negative future implications. Thus, designers are challenged to investigate solutions that would represent an optimal trade-off between the benefits

A. PRIVACY AND TECHNOLOGY CHALLENGES FOR UBIQUITOUS SOCIAL NETWORKING

and risks that the different software architectures might impose.

The second selection of personal data regards which information should be included into users' sub-profiles to be disclosed under the different circumstances of their encounters. Automated ad hoc privacy control models were identified to be more suitable, because they were found to better reflect on people's natural privacy handling, when compared to predefined privacy preferences mechanisms. However, these privacy mechanisms present a crucial challenge that needs to be further investigated. We refer to the disagreement between the users and the ad hoc privacy control about data disclosure decisions, which might jeopardise the users' personal privacy. In fact, an ideal ad hoc privacy control model should manage users' personal data privacy as individuals do in ordinary human interactions, where they intuitively evaluate various determinants and unconsciously choose what personal information to disclose during face-to-face interactions.

As discussed in this article, the inquirer determinant was found to be the primary predictor for users' data disclosure decisions in ubiquitous social computing environments over the current situations factor, e.g. location, activity. However, due to focus on disclosure of personal information to strangers, we strongly recommend ubiquitous social networking researchers to advance the attention to the current situations parameter. Thus, we strongly suggest a further analysis to gain an extensive comprehension of variation of human data sensitivity that affects information disclosure under different circumstances, including many expected and unexpected variables. In-dept investigations based on longitudinal studies about the determinants, already found to influence data disclosure decisions and identification of new predictors are needed for ensuring accurate selection of sub-profiles

to be disclosed under different circumstances of the users' encounters in ubiquitous social networking environments.

Finally, in order to accurately analyse the current users' circumstances, ubiquitous social networking environments must be capable of acquiring relevant and accurate context, which needs to be processed by the ad hoc privacy control model. The results presented in this article recommend further research on acquisition of detailed social and environmental context for helping privacy management systems to better understand and consequently evaluate current users' circumstances. We specifically refer to inference of detailed social activities in highly complex situations, such as collaborative actions within groups of two or more participants.

Acknowledgments

This work is supported by Nokia and developed as a part of the Converged Advanced Mobile Media Platforms (CAMMP) project¹⁴. The authors also acknowledge the anonymous reviewers for their valuable comments.

¹⁴<http://www.cammp.dk>

Table A.1: Key findings of the three identified crucial challenges for ubiquitous social networking

| Challenges | Parameters | Key Findings |
|------------------------|---|---|
| Context Acquisition | <ul style="list-style-type: none"> - User identities - User relationships - User activities - User profiles | <ul style="list-style-type: none"> • Bluetooth is the most widely used technology for detection of users' identities and relationships • Integration of ubiquitous devices with online social networks increases the amount of personal information included into user profiles • USN can benefit from large amount of preferences and information stored in online social networks • Infer of collaborative social activities involving groups of two or more users based on wearable sensors is a promising future research direction |
| Software architectures | <ul style="list-style-type: none"> - Centralized - Decentralized - Hybrid | <ul style="list-style-type: none"> • Pure peer-to-peer architecture motivates users to share their personal data as disclosure occurs directly between end-users without going through a central unit or server • Peer-to-peer architecture limits users to only exchange their personal profiles in close proximity with no possibility to modify it afterwards • Centralized architecture provides opportunity for users to modify their sharing preferences even after disclosure • Centralized architecture may require users to embrace disclosure to third-party entities, which could discourage users from disclosing their personal information • Hybrid architecture facilitates USN applications to collect personal information from user profiles in online social networks • Hybrid architecture, similar to centralized architecture, may require users to embrace disclosure to third-party entities, which could discourage users from disclosing their personal information |
| Privacy | <ul style="list-style-type: none"> - Data protection - Personal Privacy | <ul style="list-style-type: none"> • USN services should include data protection solutions to secure collection and dissemination of users' personal information • Personal privacy solutions are needed to help users select the right information to be disclosed to the right person under the right circumstances |
| | <ul style="list-style-type: none"> - Predefined preferences - Ad hoc privacy control | <ul style="list-style-type: none"> • Ad hoc privacy control is preferred to predefined privacy preferences for applications promoting social interactions between users during physical meetings given its better protection of dynamic and intuitive aspects of users' personal data • Predefined privacy preferences models are not considered suitable for protecting personal privacy in USN environments because of the complexity for users to predict all potential situations and associated data sharing decisions prior to actual data disclosure • Unintentional invasion of privacy may occur when rules indicated by predefined data disclosure preferences do not accurately reflect the actual users' sharing preferences • Further research is needed to minimize the disagreement on data disclosure between users and ad hoc privacy control models |

Appendix B

Ubiquitous social networking: concept and evaluation

Sapuppo Antonio Center for Communication, Media and Information Technologies - Aalborg University, Sydhavnsgade 17, Copenhagen 2450, Denmark - Email: antonio@cmi.aau.dk

Sensor Letters, Vol. 10, No. 8, pp. 1632-1644, 2012.

Abstract

Despite the great success of online social networks, there is still no automated way to facilitate communication between people in the physical environment. Ubiquitous social networking services target at transferring online social networking benefits to the physical world, by facilitating advantageous relationships during physical meetings between people who do not know each other, but probably they should. In this paper, we present a potential solution for establishing ubiquitous social networking services by integrating online social networks with opportunistic networks. This solution, called local social networks, focuses on uncovering relevant connections between people nearby, by providing a platform for automatic exchange of user personal information in order to discover interpersonal affinities. Firstly, we define and discuss the concept, advantages, preliminary architecture and potential future applications of local social networks as well as introduce the first prototype, named Spiderweb. Afterwards, we present results of a qualitative investigation that researched whether 16 active online social networks users would accept ubiquitous social networking services. The results revealed that all the participants perceived the usefulness of these services and 14 of them would be willing to accept all the necessary requirements for the establishment of local social networks and thus be potential users.

Keywords: Ubiquitous Computing; Social Networking; Privacy; Information Disclosure; Mobile Computing.

B.1 Introduction

The creation of Internet provided an innovative communication infrastructure that reduced the distance between people living in different parts of the world. Soon on the basis of this technology new services have been developed, which improved the communication between people. Such initiatives as Orkut¹⁵, MySpace¹⁶ and Facebook¹⁷, called online social networks (in the following referred to as OSNs), share a common characteristic: they enable people to create a virtual social network where users can stay in touch with friends from the whole world, share pictures, talk, chat, send messages and look for new acquaintances [48, 206].

Despite the wide spread of OSNs, the flexibility and sociability of these networks can be questioned. Firstly, the access to OSNs services is not available upon user's demand, as it occurs exclusively while using a desktop computer [65]. Further, the human communication is still highly embedded in the physical contact and closeness, provided by the physical environment. Unfortunately, OSNs do not facilitate social communication in the physical environment. Thus, people with shared interests and backgrounds fail to leverage interpersonal affinities for personal benefits [65, 83, 155].

Recently, the flexibility restriction of OSNs has been solved by enabling the OSNs services on mobiles. The real advantage of mobile social networks, if compared to classic OSNs, is that mobile terminals elevate the freedom of movement while using the applications [159]. Moreover, mobiles are not seen only as entry points to existing centralized OSNs, but thanks to their wireless technologies they also

¹⁵<http://www.orkut.com>

¹⁶<http://www.myspace.com>

¹⁷<http://www.facebook.com>

B. UBIQUITOUS SOCIAL NETWORKING: CONCEPT AND EVALUATION

enable Opportunistic Networks (in the following referred to as ONs). In ONs, nodes are wirelessly connected and have the possibility to identify each other as well as exchange contents in a short communication range [95, 102, 153]. When user personal information is incorporated to ONs, these networks can be perceived as an important tool for addressing sociability issues in the physical world, as they enable the establishment of ubiquitous social networking services [10, 59, 65, 153, 183].

Ubiquitous social networking services attempt to address sociability issues by providing a controlled automated communication system, applied in everyday physical world. These services help users to develop possible advantageous relationships such as friendships, partnerships, business relations by uncovering hidden connections that people share with others nearby and thus facilitating initialization of face-to-face interactions. As a result, the value of social networking is significantly enhanced and benefits are available immediately upon demand [65, 83, 155, 183].

In this paper, we present a potential solution for the establishment of ubiquitous social networking services, called local social networks, which incorporates the users' personal information into ONs by integrating OSNs with ONs. Thanks to this integration, local social networks are capable of combining online and ubiquitous social networking services and provide them to the users through a single platform. In order to analyze the acceptance of local social networks, we ran a qualitative investigation with 16 active online social networks users. Particularly, we researched the perceived usefulness of ubiquitous social networking services as well as participants' acceptance of the crucial requirements for the establishment of these services.

The rest of the paper is structured as follows: in the next section we describes

and defines in details the concept of local social networks, presents the preliminary architecture, first prototype as well as suggestions for future potential applications of ubiquitous social networking services. In Section B.3, we review the background and design of the qualitative investigation and information about the participants. We present results of the conducted qualitative tests in Section B.4. Final conclusions and recommendations for future work are drawn in the last section of the paper.

B.2 Local social networks

Local Social Networks (in the following referred to as LSNs) focus on promoting ubiquitous social networking services in order to facilitate face-to-face interactions between people during physical meetings. This solution attempts to address sociability issues by providing a platform for automatic exchange of user profiles in order to discover interpersonal affinities and consequently create new beneficial relationships between users, who do not know each other, but probably should.

In the following, firstly, we present an example scenario, illustrating the ubiquitous social networking process, followed by the definition and preliminary architecture of LSNs. Further, we introduce the first prototype and potential application areas of LSNs.

B.2.1 Example scenario

To better explain the LSN concept, we present an example scenario of ubiquitous social networking services in Figure B.1. A user named Bob, who is marked in blue, is located in a public place, such as a canteen of an ordinary work place.

B. UBIQUITOUS SOCIAL NETWORKING: CONCEPT AND EVALUATION

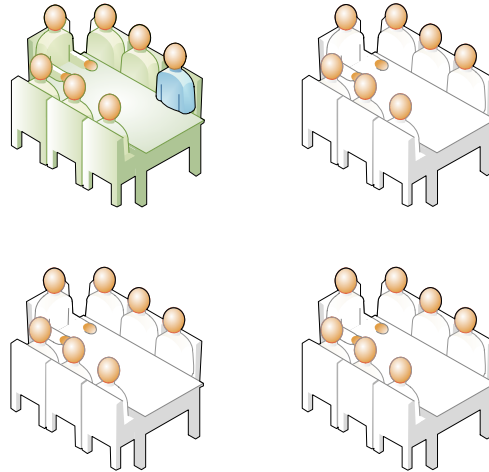


Figure B.1: Ubiquitous Social Networking Example Scenario

Bob is surrounded by people that he knows, marked in green, and people that are strangers to him, marked in white. Even if Bob does not interact with all people in the canteen, his mobile phone does it for him by exchanging personal information with other LSNs peers in his proximity, as shown in Figure B.2-A. Due to the exchange of personal information, LSNs develop an understanding about who are the people nearby as well as their respective preferences. Thus, these services are capable of highlighting relevant social paths between users that would be hidden otherwise, as shown in Figure B.2-B. When LSNs find profile similarities between Bob and other LSNs users, highlighted in yellow in Figure B.2-B, they are notified about each others' presence and, therefore, have the opportunity to immediately initiate a face-to-face communication. However, in LSNs the exchange of personal information is automatic and it does not interfere with the current Bob's activity. The relevant information, which is useful for networking with other users, can be retrieved and used even at a later time.

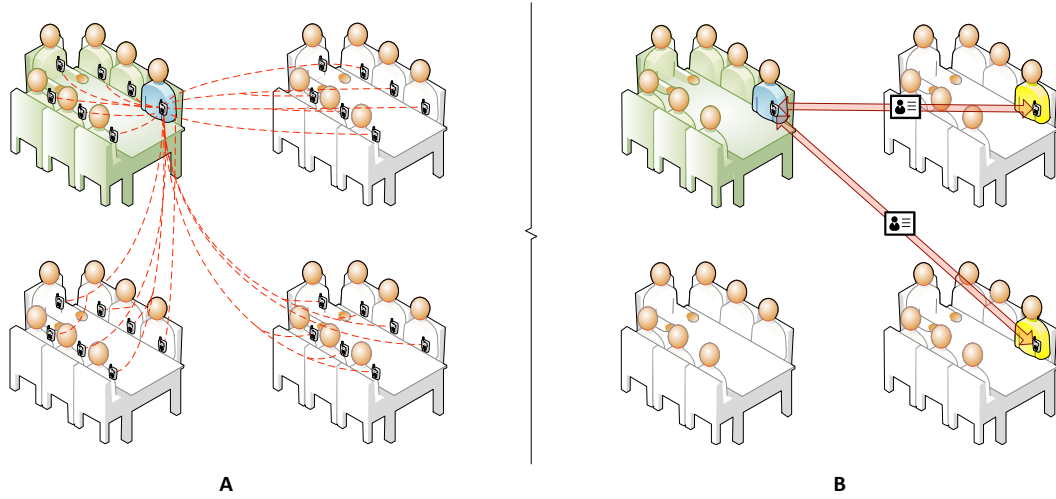


Figure B.2: Ubiquitous Social Networking Example Scenario

As it can be derived from above example scenario, when Bob is taking part in these networks, all the other LSNs peers become aware that he is somewhere around. Moreover, due to the exchange of users' personal information, Bob is able to access others' personal data, as well as his data is revealed to the other peers of the network. Finally, as a result of immediate notifications about relevant profile similarities, Bob can initiate profitable conversations with the users, marked in yellow in Figure B.2-B, and vice versa, i.e these users might decide to start a face-to-face interaction with Bob. Consequently, following this example scenario, we can identify three crucial requirements for the establishment of local social networks:

- Announcement of users' presence: users must accept to inform other nearby LSNs users about their whereabouts;
- Disclosure of personal data: users must accept to share their personal information with others;

B. UBIQUITOUS SOCIAL NETWORKING: CONCEPT AND EVALUATION

- Potential initiation of face-to-face interactions: users must accept possible immediate face-to-face interactions with other LSNs users, when notified about potential profile similarities.

B.2.2 Definition

After introducing the LSNs concept, we concisely define local social networks as follows:

A local social network is a wireless network of opportunistically connected sociable nodes

In other words, LSN is a distributed network architecture in which nodes are linked to online social networks profiles and wirelessly interconnected to exchange personalized contents. The communication range between the sociable nodes is direct and limited to the walking distance.

B.2.3 Preliminary architecture

Figure B.3 shows the preliminary architecture of local social networks, which is based on the integration of OSNs with ONs. The left side of Figure B.3 presents the OSN architecture that is following the classical client/server model. The server is connected to a database that contains all the information about the users of the application: user profiles, messages, contact invitations, relationships between the users, etc. All the other elements are an assortment of clients. A client is able to interact with the server by starting a communication through an IP bearer technology. The right side of Figure B.3 presents the ON architecture, which is

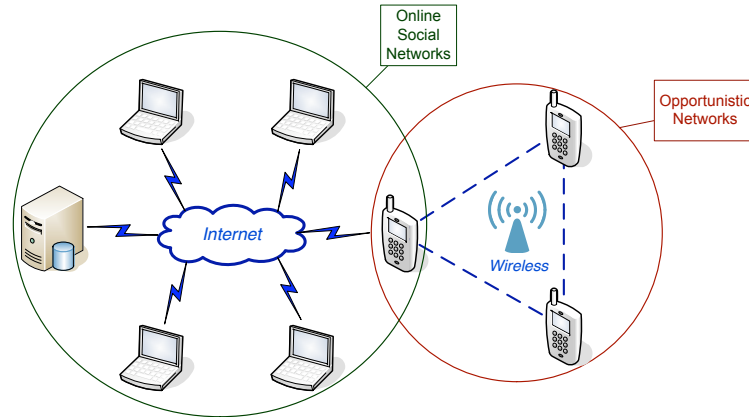


Figure B.3: Preliminary architecture of Local Social Networks

based on the peer-to-peer wireless communication and consequently does not present any central server.

The main difference between OSNs and ONs is evident: the range where social networks are established. Regarding OSNs, the amount of data is usually vast and therefore it may be difficult to find the needed information. In case of ONs the amount of data is restricted to the range of the wireless technology adopted. This range has to be short enough to ensure that users are in the proximity of each other. At the same time it has to be long enough for users to scan without being noticed [153]. Figure B.4-A shows the wireless range of the user, however in reality the communication range is not an ideal circle due to communication signal interferences with the surroundings [95]. Within the wireless range, users are able to instantly discover each other and exchange personal contents (e.g. profiles, messages, etc). When the device at the center of the circle discovers other users in its proximity, a direct connection between these two users is possible, as shown in Figure B.4-B. On the contrary, the device outside the wireless range is not discovered

B. UBIQUITOUS SOCIAL NETWORKING: CONCEPT AND EVALUATION

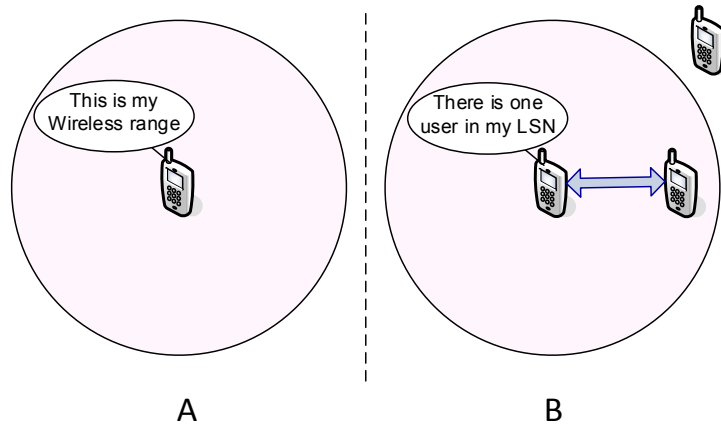


Figure B.4: The range of Local Social Networks

and, consequently, these devices do not know about each other's existence and communication between them is not possible unless they move into each others' wireless range.

Importantly, a LSNs node behaves as OSN node and ON node at the same time, as illustrated in Figure B.3, where a node is placed in the intersection space between the OSNs and ONs. The LSN node architecture is presented in Figure B.5. The bottom layer of the LSN node architecture manages the communication matters. While an IP Bearer technology can be used to access OSN services, the Exchange Data communication is used to enable direct communication in a short range of the selected wireless technology. The Node Discovery and Environment Discovery are adopted to define the current surroundings of the user, respectively information about other nodes in the user's LSN and relevant context data to interpret current users' location, activity, etc.

The second layer of LSN node architecture is composed of collection of services that can be offered by LSNs, which are enabled through the bottom communication

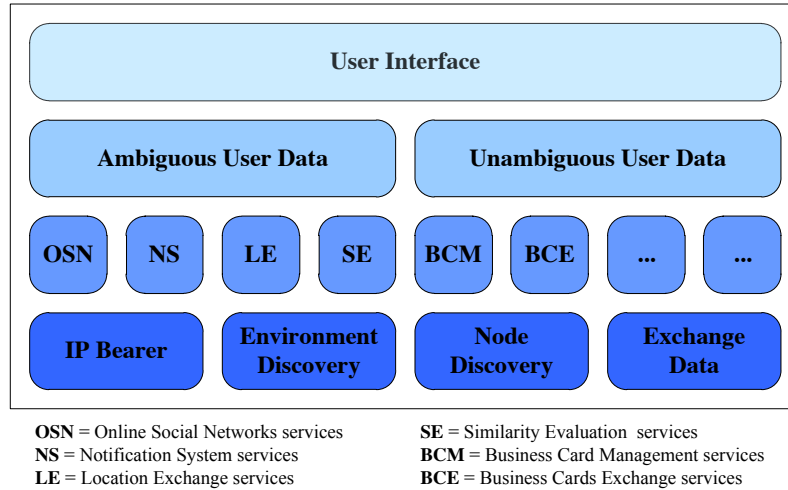


Figure B.5: Local Social Networks node architecture

layer. They are a collection of OSNs services (e.g. find new acquaintances, search for people on the basis of certain criteria, chat, view profile of other users, etc) and others, which promote sociability during physical meetings. Particularly, the latter services, such as Business Card Management and Exchange services, enable effective control and distribution of personalized user profiles in the users' vicinity. Moreover, the Similarity Evaluation services enable LSNs to compare the user profiles, calculate the similarity scores between the encountering users and consequently trigger the Notification System services when needed. The Notification System alerts users about the exchange of personal information with others, who present relevant profile similarities for networking.

The profile of the user is placed in the third architecture level of Figure B.5, which is divided into ambiguous and unambiguous personal data. The Ambiguous User Data represents a set of information that may be subject to continuous changes, such as user preferences (e.g. food taste); the Unambiguous User Data is related to

B. UBIQUITOUS SOCIAL NETWORKING: CONCEPT AND EVALUATION

a set of information, which does not change often (e.g. home address) and thus it is considered to be of high sensitivity. Consequently, different strategies for preserving user's privacy can be applied based on this data classification. Finally, the interface is placed in the top level of the LSN node architecture. The User Interface has access to all the other layers of the LSN node in order to accomplish its purpose.

B.2.4 The Spiderweb prototype

The first LSN prototype is the Spiderweb mobile social network application [166], which was selected to be implemented in Java 2 Micro Edition (J2ME) in order to be compatible with all mobile operating systems supporting J2ME [87]. The application follows the software architecture schema described in Figure B.3, thus Spiderweb integrates OSNs with ONs. Spiderweb partly relies on the Internet connectivity and offers several services, already well-known from online social networks sites. In fact, users are able to create their profiles and invite others to their social networks. They are also able to search for people based on certain filtering criteria, exchange messages, let other users know their current position and keep in touch with their friends. Additionally, Spiderweb uses Bluetooth connectivity to allow devices to identify each other and establish direct connections between them by using the short range communication. Within the Bluetooth range, Spiderweb users share a subset of the full user profile, called Business Card (in the following referred to as BC), which is stored in the local memory of the device and is synchronized with the relevant fields of the OSN profile. Due to exchange of BCs, Spiderweb users are capable of uncovering hidden connections that they share with other people nearby.

Two screenshots of the Spiderweb application are shown in Figure B.6. On the

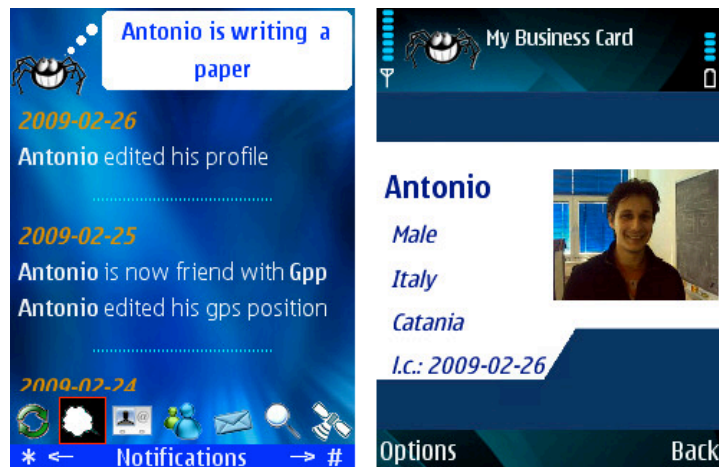


Figure B.6: Two screenshots of Spiderweb

right side of Figure B.6 an example of a BC is presented. The BC is composed of personal information that is accepted by the Spiderweb users to be shared with others in their proximity. On the left side of Figure B.6, the notification screen is presented. Notifications regard OSNs and ubiquitous social networking services. In relation to OSN services, users are notified when they change their current status, update their profiles or GPS positions as well as receive text and picture messages from other users or establish new friendships. Moreover, in respect to ubiquitous social networking services, the user is notified if others with specific characteristic that the user is looking for are in the proximity (e.g. likes rock music or is a friend in the OSN). From the notification screen, the user can access both OSN and ubiquitous social networking services, utilizing the icons on the bottom of the screen.

Even if Spiderweb is the first prototype of LSNs, many functionalities of LSNs preliminary architecture, presented in section B.2.3, are not fully applied. In fact,

B. UBIQUITOUS SOCIAL NETWORKING: CONCEPT AND EVALUATION

some components and services related to the automated creation of users' BCs have not been implemented yet (e.g. Environment discovery, Business Card Management services, etc). Future implementation of these components and application of these services are of crucial importance for improving the Spiderweb application and contributing to the long-term success of ubiquitous social networking services.

B.2.5 Application areas

In the following we present several application areas, where LSNs services might significantly improve people's social being and connectedness:

- Professional: Ubiquitous social networking services might provide significant improvements to users' professional life. Firstly, they might lead to new personal professional opportunities, such as connecting employers with potential employees and vice versa. Further, LSNs services would help to initiate collaborative teams within organizations. For example, they would assist in connecting people who are working on similar material as well as finding others who have abilities to solve another employee's current problem [65].
- Dating: OSNs that focus on dating services became very common in the recent years. Even Facebook, which is not primarily targeting at promoting dating services, significantly increased its popularity when it enabled users to discover the relation status of others. In comparison to these OSNs, LSNs would definitely provide a new way to facilitate such kind of services between the users. In fact, ubiquitous social networking services would not be limited to desktop applications, but they would help to connect nearby users with

similar social interests and thus provide opportunities to immediately initiate potential face-to-face social interactions in their everyday lives [10].

- Events: LSNs might also present significant potential in various events, such as conferences, company events, exhibitions, etc. These situations usually comprise large amounts of participants, who potentially share similar professional or social interests. However, due to time limitations, people do not have enough time to network with all the participants and thus fail to exploit potential networking benefits. As an example, many professional networking connections could be established during large academic conferences with the help of LSNs that would take into account similar research interests [65, 155].

B.3 Investigation methodology and design

In this section we present the methodology and design of a qualitative investigation that aims at analyzing the acceptance of ubiquitous social networking services among active online social networks users. Our investigation aims at answering the following question:

Would participants accept ubiquitous social networking services?

In order to answer this question, we considered to apply the Technology Acceptance Model (TAM) that comprises analysis of perceived ease of use and usefulness of the application [56]. The perceived ease of use was already investigated in regard to the Spiderweb local social networks application. The results of the investigation were very satisfactory, as the 70% of the participants found Spiderweb to be "very

B. UBIQUITOUS SOCIAL NETWORKING: CONCEPT AND EVALUATION

easy to use”. This indicated that Spiderweb users would not require much effort to become skillful in using the application [166]. Consequently, in this qualitative investigation, we focus the attention on analyzing the perceived usefulness of ubiquitous social networking services. Moreover, in order to get more insight into the participants’ acceptance of ubiquitous social networking services, other factors must be taken into consideration. We refer to investigation of participants’ acceptance of all the necessary prerequisites for the establishment of LSNs, presented in the example scenario in Section B.2.1.

In the following, we describe the background and design of the qualitative investigation, followed by the information about the participants.

B.3.1 Background

A qualitative investigation was conducted for evaluating the acceptance of ubiquitous social networking services. Qualitative interviews were preferred alternatively to other investigation methods, such as handing out questionnaires or establishing a focus group interview. This method was chosen because of the following two reasons: (i) participants’ unfamiliarity with the ubiquitous social networking subject and (ii) potential misinterpretation of the research questions due to their complexity and ambiguity. Moreover, we decided to run semi-structured interviews to better understand the motivation behind the participants responses and ensure that general areas of information are collected from each participant, however still allowing adaptability of the interview process [50, 118, 136].

In order to help participants to get more familiar with the ubiquitous social networking concept, firstly, we presented different scenarios from everyday lives, where

B.3 Investigation methodology and design

these services might be applied. Secondly, we introduced all the available Spiderweb services, which are presented in [166] as well as in this video¹⁸, and discussed with participants potential networking benefits and threats. Lastly, participants had an opportunity to utilize a mobile application that simulates the LSNs behavior for 11 days. At least 3 times per day, the mobile application was randomly asking the participants to upload their personal data disclosure decisions for the specific circumstances, encountered at the moment of the request. The selection of data types to be disclosed was provided in accordance to data categorization in popular OSNs sites (e.g. gender, age, favorite music, etc.). This categorization was already used in a previous investigation about disclosure of personal information in ubiquitous social computing environments and the detailed description of the provided data types can be found in [170]. Participants were aware that potential networking benefits would be directly proportional to the amount of shared information, thus their ad hoc data disclosure decisions were representing a compromise between privacy risks and potential networking benefits.

Notably, we preferred to provide a new mobile application, designed specifically for this investigation, rather than utilizing the Spiderweb or other existing ubiquitous social networking applications, such as [10, 65, 153, 183], because these applications are still not widely spread yet and participants would encounter difficulties in finding opportunities to disclose their personal information to real users. Afterwards, we interviewed the participants for a duration of 30-45 minutes. The interviews were audio taped and transcribed at later time. Questions were, firstly, related to perceived usefulness of ubiquitous social networking services, which inves-

¹⁸<http://www.youtube.com/watch?v=DgeVNv10CIM>

B. UBIQUITOUS SOCIAL NETWORKING: CONCEPT AND EVALUATION

tigated the degree to which participants believe that using a particular technology would enhance their networking performance [56]. Further, we analyzed the acceptance of the three crucial prerequisites for the establishment of ubiquitous social networking services, derived from the example scenario in Section B.2.1, i.e. announcement of user's presence, disclosure of personal information and potential initiation of face-to-face interactions.

B.3.2 Participants

The selection of participants was limited to OSNs users. We determined this category to be the most relevant because of their advanced experience in social networking, even if the perception towards the networking services might vary between virtual and physical worlds.

Respondents were asked to provide information about their demographic characteristics and asked to indicate their privacy preferences on visibility of their own personal data (e.g. user profile, pictures, posts) in their main OSN site. Based on these answers, we were able to observe patterns among data disclosure attitudes and divide the participants into three privacy clusters, following the Westin/Harris privacy segmentation model [198]:

- **Fundamentalists:** these respondents were extremely concerned about sharing their personal data with any other online social networks users (friends or strangers);
- **Pragmatists:** these participants also cared about loss of privacy due to the disclosure of their personal information. However, they often had specific

concerns and particular strategies for addressing them. For example, this category of respondents generally preferred sharing personal information only among their friends;

- Unconcerned: these respondents were trusting online social networks sites and believing that the privacy of their data was not jeopardized. Thus, they were willing to share their personal data not only with people who were their friends, but as well with users who were complete strangers to them.

In total we recruited 16 participants with the following privacy and demographic characteristics:

- Gender: 10 of the participants were male, while 6 of them were females;
- Age: 7 of the respondents were younger than 26 years, 7 of them were between 26 and 35 years old and 2 participants were older than 35 years;
- Occupation: 8 of the participants were working and 8 of them were studying at the time of the survey;
- Privacy: 9 of the respondents were pragmatists, 4 of them were fundamentalists and 3 participants were unconcerned.

B.4 Investigation results

In this section we present results of the qualitative investigation that analyzes participants' acceptance of ubiquitous social networking. As introduced in Section

B. UBIQUITOUS SOCIAL NETWORKING: CONCEPT AND EVALUATION

B.3, we firstly focus on researching perceived usefulness of LSNs services. Afterwards, we investigated the participants' acceptance of the crucial prerequisites for the establishment of ubiquitous social networking.

B.4.1 Perceived usefulness

The qualitative interviews were firstly attempting to get insight into perceived usefulness of LSNs by focusing on the following subquestions and motivations supporting corresponding participants' answers:

1. Would respondents perceive that ubiquitous social networking services improve their everyday communication?
2. For what purposes would the participants perceive the ubiquitous social networking services to be useful?

Firstly, as shown in Figure B.7, all the participants acknowledged the potential of LSNs services for improving their everyday communication, because they believed

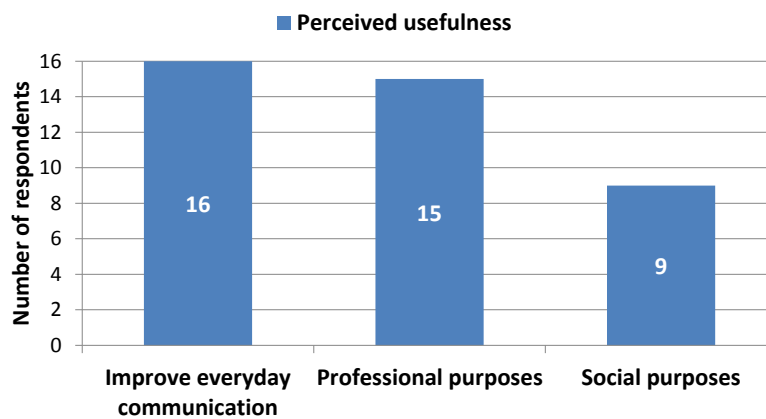


Figure B.7: Perceived usefulness of ubiquitous social networking services

that these services would help them to connect with other people nearby, who have similar social and professional interests and goals. Furthermore, the participants considered that LSNs services relevantly support ordinary human interactions and behavior that people usually maintain in their everyday lives. Some of the respondents highlighted that these services would be very useful for people with any distinctive interests, as one of them said:

"Many teenagers wear t-shirts representing their favorite music (e.g. metal, rock, etc). The reason why they do so is to attract the attention of other people with the same music preference in their surroundings. Now think about having a complex interest - how can you represent it in a t-shirt? In such cases LSNs would be really helpful"

Secondly, as illustrated in Figure B.7, potential professional networking benefits, received in exchange to disclosure of personal information, were notably considered to be the best motivation for using LSNs services by 15 out of 16 respondents. They indicated that LSNs would improve their professional lives, because such services were considered to be relevant to speed up the process of initiating beneficial professional relationships. For instance, a participant, who was just back from a big exhibition in London, claimed that these services would significantly increase the networking efficiency of that exhibition:

"Companies were presenting us what they were working on and invited to submit our curriculum vitae. However, the exhibition was very big and many companies were participating, thus I did not manage to get informed about all the job opportunities and I could not network with

B. UBIQUITOUS SOCIAL NETWORKING: CONCEPT AND EVALUATION

the representatives of all the companies. I believe that I lost relevant professional opportunities. In such situations, those services would significantly improve my networking by highlighting which companies are worth to interact with and to present my curriculum vitae”

While the majority of the respondents found LSNs services to be very useful in regard to professional life, some of them (i.e. 7 out of 16) seemed to be cautious when considering to share their personal information for improving their social relations, as shown in Figure B.7. These participants found social interactions to be more sensitive compared to professional ones and they were skeptical about sharing their social life with other people whom they do not know. However, some of them believed that they might start using LSNs also for social purposes after developing an initial familiarity with these services. However, despite their overall carefulness, one of respondents emphasized circumstances where he would be interested in using LSNs also for improving his social interactions:

”Last summer, I visited Los Angeles with my friends. I believe that such services would have been very useful in many circumstances during my holiday, as we spent a lot of time in different social environments and we were in the right mood for starting new social interactions with people around us”

B.4.2 Acceptance of prerequisites

When inquired about the acceptance of ubiquitous social networking prerequisites, only a few of the participants had serious concerns about accepting all the needed

requirements for the establishment of these services. In the following subsections, we present the individual results in relation to acceptance of the three identified prerequisites.

Announcement of users' presence

In order to enable the LSNs users to announce their presence to the others, two main concepts are relevant to be considered: sharing of location information and proximity information. Both of them are related to sharing of the user's current position. However, the location information is generally intended to be shared among acquaintances and in an unlimited range, while proximity information is meant to be disclosed to only all people nearby.

Sharing of the location information is one of the main characteristics of mobile social networks, which enables location based services. Generally, these services help users to connect to friends, be alerted when they are close and discover places around them by sharing users' GPS positions [206]. However, mobile social networks users have presented serious privacy concerns, related to disclosure of location information, even if it is applied only among acquaintances [37]. Sharing of the proximity information, instead, presents added value for solving privacy issues, related to disclosure of location information, because LSNs users are exclusively discovered within the range of the adopted wireless technology, as shown in Figure B.4. In fact, LSN users are notified about the presence of others only when they are in the vicinity. Once the people move away, the information about the users' presence is not available anymore, unless they re-enter into each other's wireless range. As a result, the privacy threats, related to disclosure of current position,

B. UBIQUITOUS SOCIAL NETWORKING: CONCEPT AND EVALUATION

are decreased and the application of this concept might lead to positive trade-off between potential benefits and threats.

In relation to ubiquitous social networking, while accepting to share the proximity information is considered to be a mandatory requirement for participation in LSNs, disclosure of location information is an optional feature to access a wider range of services. For example, Spiderweb users can disclose both location and proximity information. However, they can switch the location disclosure feature off at any time, and still be able to exploit ubiquitous social networking services by sharing their proximity information.

In order to gain insight into participants' willingness to share their current position, firstly, we asked them whether they would like to disclose their location information and afterwards we also investigated the willingness to share the proximity information. As shown in Figure B.8, the majority of respondents (i.e. 10 out of 16) preferred to maintain their current location private and provided comments, similar to the following:

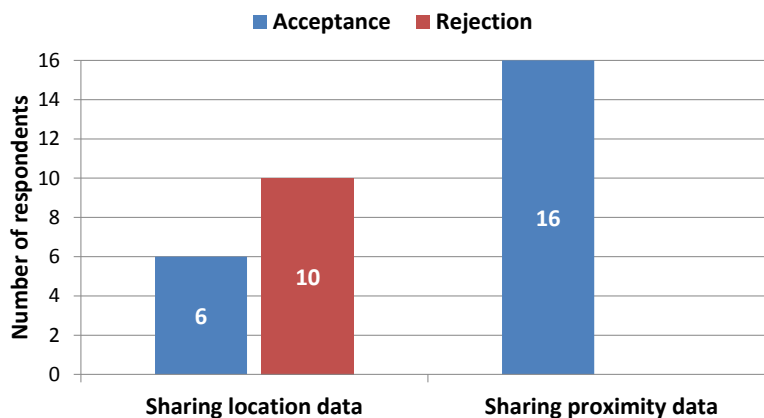


Figure B.8: Acceptance of announcement of user's presence

”Location is very sensitive information. I do not think it is important for other people to know where I am. Even in case of people whom I know very well, I still would not disclose the exact address of my position, but I might disclose less detailed information, such as country or city where I am now”

However, the rest of the participants (i.e. 6 out of 16) accepted to share their location information because they could perceived several advantages of this feature. For example, one of them claimed:

”Two years ago I would probably tell you that I would never share my current location, but now I have started using location based services, such as Google Latitude¹⁹, and I really enjoy the benefits of these services. My friends can see where I am and join me when they are not far away”

After discussing the willingness to share the location information, we further inquired respondents whether they would allow the disclosure of proximity information. As shown in Figure B.8, all the participants were very positive about this new concept and would permit the disclosure of their proximity information without any concerns. They acknowledged the advantages of announcing their presence only to people nearby over sharing of their location information. It could be assumed that this preference is motivated by the perception that disclosure of proximity information does not lead to invasion of privacy, as many participants provided comments, similar to the following:

¹⁹<https://www.google.com/latitude>

B. UBIQUITOUS SOCIAL NETWORKING: CONCEPT AND EVALUATION

"I think that disclosure of proximity information is a very good idea. In this case, I would not mind to announce my presence to exclusively people nearby, because if they look around, they can just see me anyway"

Furthermore, some of the respondents indicated that they would prefer disclosure of proximity information over location information, because the latter comprises unnecessary sharing of user's current position to third parties. These participants claimed that they would not utilize such services, if they had to disclose their current position to third parties, because they were worried about losing control over this *sensitive* personal data.

Disclosure of personal data

Sharing of personal information, such as user's preferences and contact information, is indisputably the crucial foundation of ubiquitous social networking services. Thus, we asked participants whether they had any concerns about disclosure of their personal information to other nearby users, who would be strangers for them.

As a result, all the participants accepted to disclose their personal information in order to gain potential networking benefits in exchange. Many of them discussed that disclosing their personal information through ubiquitous social networking services would not be much different from having face-to-face interactions with strangers themselves, which as well comprise sharing of personal data. Furthermore, the majority of the participants also emphasized that they would feel more motivated to share personal information, if LSNs ensured that the data sharing occurred only for specific purposes, as one of them noted:

"I would definitely disclose very detailed information about my work activities if I perceived potential for getting a better job"

Furthermore, participants discussed that they would appreciate that LSNs services would empower their users to always keep control over their personal data, even after actual disclosure, as it would increase perceived trust and provide better usability of such services. Many of the participants provided comments, similar to the following:

"If I had met a potential employer, I would like to make sure that this person got all the information about my career skills, abilities and expectations. Thus, I would appreciate a possibility to additionally share this information, if it was not disclosed during the initial data disclosure"

Also, one of the participants, who seemed to be strongly influenced by the public opinion, discussed how keeping control over his shared data would allow him to disclose newly discovered personal preferences, without being worried about future social implications:

"If I added to my business card that I am a fan of a famous runner, but at later time it would be discovered that this person was not honest in his sport achievements and thus lose his good public image, I would definitely be very embarrassed to keep him in my profile. The same would apply for other personal preferences, such as new movies or books, which are subjects to public opinion. I would not disclose such preferences, if I did not have the opportunity to modify them at later time"

B. UBIQUITOUS SOCIAL NETWORKING: CONCEPT AND EVALUATION

Finally, we inquired participants about three different methods for sharing of users' preferences in LSNs and asked them to choose their preferred approach. The three proposed solutions are following described:

1. Static profile: this solution discloses the same profile in all the encountered circumstances. It is a well-known approach, already utilized in the majority of existing ubiquitous social networking applications, such as [10, 65, 153, 183];
2. Predefined privacy preferences: this solution attempts to predict all the potential situations and associated data sharing decisions a priori to the actual data disclosure. This approach was already adopted in [103, 111, 124, 144, 176];
3. Ad hoc privacy control: this solution provides opportunities to take data sharing decisions *in situ* - at the moment of actual disclosure. This approach automatically manages information disclosure on the users' behalf in order to relieve them from frequent data disclosure decisions [26, 27, 103].

After introducing to the participants different scenarios, based on the first two approaches (i.e. sharing of a static profile and predefined privacy preferences), we showed them results of a simulation of ad hoc privacy control mechanism, which takes into account both previous data disclosure decisions and relevant influential factors (e.g. location, activity, mood, mutual friends) that were proven to impact users' data disclosure decisions [45, 123, 167, 169]. For each participant, we presented his/her corresponding prediction result, obtained by processing his/her data disclosure decisions, collected while using the LSN prototype. Specifically, we applied the binary logistic regression statistical model and achieved an approximate

accuracy of 90%, with peaks of 93%, and potential for further increasing performance.

As shown in Figure B.9, the majority of the participants (i.e. 14 out of 16) would trust ad hoc privacy control, as they highlighted the advantages of this solution over other techniques, i.e. sharing of a static profile and predefined sharing preferences, to manage their data disclosure. Firstly, the participants indicated that they preferred sharing different profiles under different circumstances, e.g. they did not want to share data related to private activities in work environments. Secondly, the participants claimed that it would be difficult for them to define in advance what to disclose per each circumstance and they expected to encounter situations where data disclosure decisions would not be accurately predictable in advance.

Moreover, respondents were confident that by utilizing automated ad hoc privacy control, LSNs would be capable of managing their personal data disclosure decisions in accordance to the real users' data disclosure preferences. In fact, after running the provided mobile application for a few days, participants experienced that they were already using a pattern on what to disclose in similar circumstances.

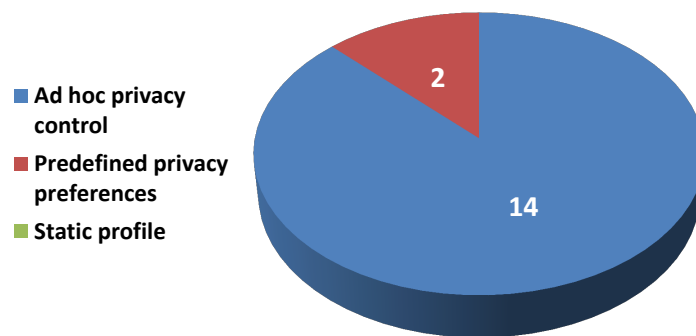


Figure B.9: Acceptance of mechanisms for disclosure of user's personal information

B. UBIQUITOUS SOCIAL NETWORKING: CONCEPT AND EVALUATION

Also, they did not express particular concerns about potential wrong data disclosure decision taken by LSNs. In case of unintended disclosure of not sensitive personal data, they discussed that many times they did not share some of their personal information, because they did not find any reasons for disclosing it, rather than for preserving their privacy. In these cases, the disagreement of data sharing decisions between the participants and ad hoc privacy control would only arise due to different evaluation of relevance, rather than data sensitivity for the current circumstances.

In case of inquiry for highly sensitive data, despite the good prediction results of the binary logistic regression model, some respondents (i.e. 4) would prefer to limit the autonomy of ad hoc privacy control. These participants provided comments similar to the following:

"In the majority of the cases I would have no concerns about allowing LSNs to manage my personal information. However there might be either highly sensitive personal data (e.g. religion) or some specific circumstances that are very important to me (e.g. I am attending a job interview) in which I would feel uncomfortable to allow a machine to take decisions on my behalf. In such situations, if possible, I would rather prefer to manage the disclosure of my personal information myself"

These results confirm the findings of previous studies, which as well investigated prediction of users' information disclosure, based on previous data disclosure decisions, utilizing data mining algorithms [24, 25, 27]. Even if presenting significant prediction results, Bünning et al claimed that automated data disclosure should limit its autonomy in case of inquiry for highly sensitive data. In such situations,

it was advised to provide only suggested data disclosure choices, while waiting for user's approval before any actual disclosure [25].

Potential initiation of face-to-face interactions

The last prerequisite for the establishment of ubiquitous social networking services that we analyzed in our qualitative investigation is potential initiation of face-to-face interactions. This requirement is directly dependent on another relevant, however not crucial, prerequisite: immediate notifications. In fact, notifying about the presence of other nearby LSNs users with relevant profile similarities provide the possibility to initiate immediate face-to-face interactions. Thus, before investigating participants' acceptance of potential face-to-face interactions, we firstly analyzed whether participants would prefer to receive immediate notifications over the possibility to retrieve the information, relevant for networking, at later time. We believe that participants had enough insight for answering this question, because of the experience gained when running the provided mobile application, simulating the LSN behavior. In fact, the respondents were alerted by the application at least 3 times per day, which can be considered as a realistic replication of LSNs notification system.

As shown in Figure B.10, only 4 out of 16 respondents preferred to access the collected information at later time. They emphasized that in many circumstances, when they received the notification from the provided mobile application, they would have preferred to postpone their attention for later time, as they did not want to be frequently interrupted. Following these considerations, relevant challenges for the implementation of the notification system were raised, as very frequent

B. UBIQUITOUS SOCIAL NETWORKING: CONCEPT AND EVALUATION

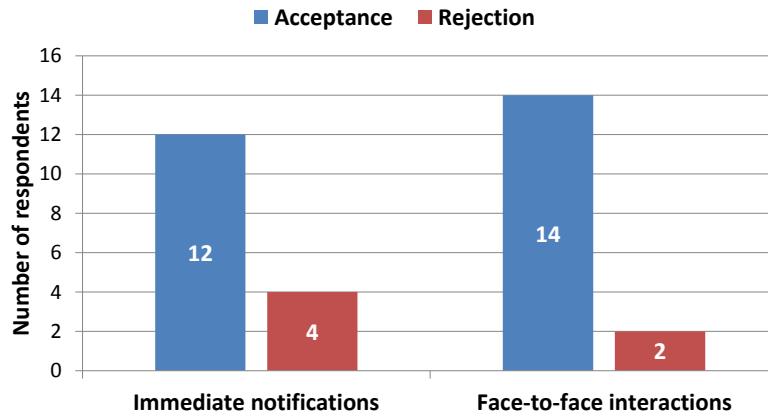


Figure B.10: Acceptance of immediate notifications about profile similarities and potential face-to-face interactions

alerts might encourage users to ignore notifications or even disable such feature. Firstly, the design of LSNs should drive users to provide detailed information in their profiles, in order to optimize discovery of profile similarities and thus avoid too frequent notifications. For example, liking *sport* would not be the same as stating to be a fan of *hockey*. Secondly, the participants discussed that the design of LSNs should enable prioritization of users' profile similarities, as one of them claimed:

"I would definitely consider to utilize the notification system if I was able to decide what to be alerted about. For instance, If I was unemployed, I would prefer to be alerted only about professional networking possibilities and to retrieve information about other types of profile similarities at later time"

Moreover, some of the respondents, as well acknowledged additional advantages of accessing the collected information about profile similarities with other users and

thus they would like to have this feature as a supplement for the notification system. Firstly, for important matters, participants discussed that they might need time to think and prepare before initiating a face-to-face interaction. In fact, they would prefer to contact the person via email, before having a conversation. Secondly, in case of lack of time for an immediate face-to-face interaction, such option would still allow to access networking benefits, as one of the respondents noted:

”Many times I write down phone numbers of people that I meet at work or social environments, but I rarely contact them, because after few days I forget the reason for having these numbers. LSNs would give me the opportunity to retrieve the relevant personal information, related to the phone numbers. Consequently, I would probably initiate a communication with them, as I would also know the motivation for contacting these people”

On the other hand, even if acknowledging that storing users’ business cards might provide relevant advantages for ubiquitous social networking, the majority of the participants, i.e. 12 out of 16, would prefer to be immediately notified when potential networking benefits arose, as shown in Figure B.10. They believed that immediate notifications is a crucial feature for ubiquitous social networking, because if the moment of interacting with other people is delayed, it loses the importance and interest to them. Participants also emphasized that they might not find the time for checking the collected information and contacting those people afterwards. Furthermore, they discussed that the benefits of ubiquitous social networking over online social networks arise due to application of the notification system in LSNs.

B. UBIQUITOUS SOCIAL NETWORKING: CONCEPT AND EVALUATION

In fact, many of them provided comments, similar to the following:

“Without the notification system, LSNs services would not be so different from classical online social networks in which a barrier is always placed between people who communicate. When notified, I can socially interact without any barrier, because what I need is just there”

While being notified is an important, but not mandatory prerequisite, potential initiation of face-to-face interactions is a requirement that users must accept when utilizing ubiquitous social networking services. Even if the users prefer to retrieve the collected relevant business cards at later time, they cannot avoid the possibility that another user would prefer to be immediately notified and, consequently, would attempt to initiate a face-to-face interaction. As potential face-to-face interactions are unavoidable in ubiquitous social networking services, during our qualitative interviews, we investigated the willingness of participants to accept this requirement. As shown in Figure B.10, the majority of the participants, i.e. 14 out of 16, would accept such prerequisite as long as they had a coarse-grained control over the ubiquitous social networking services:

“I do not see a reason for not accepting to be approached by other LSNs users, as having potential face-to-face interactions is the motivation for using such services. But, it is crucial for me to have full control over these services and be able to switch them off when desired”

However, a few respondents, i.e. 2 out of 16, claimed that they would not be potential users of LSNs if they had to accept such prerequisite, because they were too

much concerned about potential undesired face-to-face interactions. These participants were worried that someone would unnecessarily disturb them, just because of the information that they had shared. These respondents claimed that they would utilize ubiquitous social networking only if these services enabled an invisible mode option and disclose their information after user's approval, which implied manual evaluations of the trade-offs between potential networking benefits and privacy risks. However, such evaluation of trade-offs in USN would present increased complexity and require too much user's attention and intervention. Application of an invisible mode option should be carefully considered in the development of ubiquitous social networking, as it might not lead to a calm USN technology, where users could effortlessly exploit these services.

B.5 Conclusions

In this paper we presented a new communication system, called local social networks, as a potential solution for the establishment of ubiquitous social networking services. These services aim at uncovering hidden connections between people in order to leverage interpersonal affinities for networking benefits during physical meetings. We described in details the concept and the preliminary architecture of local social networks, which is based on the integration of online social networks and opportunistic networks. Moreover, we introduced the first prototype, called Spiderweb, and potential future application areas of local social networks, i.e. professional, dating and events. Afterwards, we presented results of a qualitative investigation that focused on understanding whether active online social networks users would

B. UBIQUITOUS SOCIAL NETWORKING: CONCEPT AND EVALUATION

accept ubiquitous social networking services. None of the participants were using ubiquitous social networking services at the time of the survey and 14 out of 16 of them claimed that they would be potential users of local social networks. They appreciated the possibility to be connected with other people and especially with those who share distinctive interests and goals. Participants indicated professional and events as the most relevant potential application areas for ubiquitous social networking services, however they would probably need time to get used to these services before they would also utilize them for facilitating as well their social interactions.

Moreover, we noticed that the participants, who preferred not to utilize ubiquitous social networking services, were younger than 26 years old and studying at the time of the investigation. It could be expected that these participants did not perceive any potential networking benefits in professional life because they had not started one yet. In regard to social life, they were concerned about accepting one of the three prerequisites for the establishment of ubiquitous social networking services, i.e. potential initiation of face-to-face interactions. Specifically, they were worried that their data disclosure would lead to unpleasant and undesired face-to-face interactions. However, all the other respondents acknowledged the usefulness of being immediately notified about discovered profile similarities with other nearby users and accepted the possibility to initiate a beneficial face-to-face interaction with them as long as they had coarse grained control over these services. Finally, we did not observe any crucial concerns about the other two prerequisites for the establishment of ubiquitous social networking services, i.e. announcement of users' presence and disclosure of personal data.

While the majority of respondents had serious concerns about accepting to disclose their location information, all of them accepted to announce their presence to all other nearby users by disclosing their proximity information. Respondents also appreciated the possibility to utilize automated ad hoc privacy control, which would relieve them from frequent data disclosure decisions. However, in case of highly sensitive personal information or specific circumstances (e.g. attending a job interview), a few users preferred to confirm LSNs data disclosure decisions before any actual disclosure. Moreover, they also indicated that ad hoc privacy control should provide possibilities to modify their personal data, even after actual disclosure, in order to increase perceived trust and provide better usability of LSNs.

The results of this qualitative investigation draw the attention to relevant development areas for ensuring the long-term success of ubiquitous social networking services. Firstly, further research is encouraged on variation of human data sensitivity under different circumstances in order to minimize wrong data disclosure decisions that would lead to potential unpleasant face-to-face interactions, as a result of ubiquitous social networking services. Secondly, additional insight into creation of more trustable and functional ubiquitous social networking is needed in order to provide opportunities for the users to effortlessly exploit ubiquitous social networking services, while still remaining in control of data disclosure when desired.

Acknowledgements

This work is supported by Nokia and developed as a part of the Converged Advanced Mobile Media Platforms (CAMMP) project²⁰, funded by the Danish Advanced Technology Foundation. The author is extremely grateful to the participants of the qualitative investigation who took the time to take part in this study. Without their participation and feedback, this work would not have been possible. Finally, the author thanks the anonymous reviewers of the journal for their valuable comments on the paper.

²⁰<http://www.cammp.dk>

Appendix C

Designing for privacy in ubiquitous social networking

Sapuppo Antonio Center for Communication, Media and Information Technologies - Aalborg University, Sydhavnsgade 17, Copenhagen 2450, Denmark - antonio@cmi.aau.dk

João Figueiras Telecommunications Research Center Vienna (FTW), Donau-City-Straße 1, A-1220 Vienna, Austria - figueiras@ftw.at

Accepted for publication in International Journal of Ad Hoc and Ubiquitous Computing.

Abstract

Improving human communication during face-to-face meetings is nowadays possible by transferring online social networking benefits to the physical world. This is enabled by the ubiquitous social networking services that became available by means of wirelessly interconnected smart devices, automatically exchanging personal user data. The main goal of these services is to facilitate the initialization of relationships between people who do not know each other, but they probably should. Given that sharing of personal information is an intrinsic part of ubiquitous social networking, these services are subject to crucial privacy threats. Inspired by the usability and privacy limitations of existing design solutions, we identify, describe and qualitatively evaluate four drawbacks to be avoided when designing ubiquitous social networking applications. By addressing these drawbacks, services become more functional and more oriented to ensure the end users' privacy, thus contributing to the long-term success of this technology.

Keywords: Privacy; Ubiquitous Computing; Information Disclosure; Social Networking; Design Guidelines.

C.1 Introduction

During the last decade, online social networks have quickly improved the communication between people by enabling their users to stay in touch with friends from the whole world, share pictures, talk, chat, send messages and look for new acquaintances. Since these online services have been introduced, many users have integrated them into their daily practice [18]. This great success has inspired researchers and practitioners to investigate additional mechanisms for improving human communication and to enhance social networking in the physical world [65, 153, 183]. Firstly, access to online social networks became available upon user's demands by enabling such services on mobile terminals. Then, due to their wireless technologies, mobiles are now capable of being wirelessly interconnected, resulting in new services such as discovering and connecting to neighboring devices without user intervention. This enables the creation of opportunistic networks that permit automatic data sharing in peer-to-peer ad hoc communication links. The combination of online social networks and opportunistic networks resulted in a new paradigm, named ubiquitous social computing (socUbiComp). In this context, socUbiComp emerges as an evolution of ubiquitous computing where a social dimension is introduced to respond to the sociability of the users and to increase awareness, knowledge and intelligence of such environments.

In socUbiComp environments, users are able to enrich the physical world interactions with the benefits from online social networks services. Many applications and prototypes, e.g. [10, 65, 153, 171, 183], have been created in the recent years. They target at developing eventual advantageous relationships such as friendships,

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

partnerships and business relations by uncovering hidden connections that people share with the others in their close proximity. Particularly, they facilitate the initialization of face-to-face interactions between users, who do not know each other, but probably should. As a result, the value of social networking is significantly enhanced by providing immediate benefits upon demand [65, 171]. Application areas of ubiquitous social networking are numerous and they range from professional, where these services might lead to new opportunities such as connecting employers with potential employees, to big events, such as conferences, company events and exhibitions that usually comprise large amounts of participants who potentially share similar professional or social interests [65, 168].

As disclosing personal information is an intrinsic part of ubiquitous social networking, these services are subject to crucial privacy threats [171]. Despite ongoing legal [62, 82] and academic [12, 104, 120, 122, 149] discussions about disclosure of personal information, the current design of socUbiComp environments does not provide adequate end users' privacy protection. In [122], the authors attempted to reconcile existing privacy design guidelines and suggested to empower the user to make more informed data disclosure decisions. However, when analyzing the design of existing ubiquitous social networking applications, other privacy limitations as well as relevant usability issues can be identified. Firstly, users tend to be averse to disclose their personal information to third parties [6, 84, 91, 130, 168] as they might be concerned about potential future implications. Secondly, users might not disclose their personal information if the socUbiComp does not take into account the variation of personal data sensitivity under different circumstances [169, 170, 171]. Thirdly, users might be concerned about loss of permanent control over their data

disclosure decisions [167, 168], if the individual participation principle is not enforced in the design of socUbiComp [62, 82, 120]. Lastly, with respect to usability, users might be displeased when too much user intervention is required for the accomplishment of ubiquitous social networking [191, 193]. If improperly addressed, these usability and privacy concerns could discourage users from disclosing their personal information and consequently threaten the further development of ubiquitous social networking applications.

In this paper, we present three relevant contributions for the development of privacy-aware social networking services in socUbiComp environments.

- Firstly, we identify, describe and introduce four drawbacks that should be avoided by designers of ubiquitous social networking. These drawbacks are inspired by common privacy and usability limitations existing on ubiquitous social networking applications and they do not aim at providing total security. We assume a non-malicious infrastructure aiming at preventing incidental data disclosure, where personal information is unintentionally revealed, with or without previous inquiry.
- Secondly, we present the design of a privacy-aware ubiquitous social networking platform engineered both to avoid the identified four drawbacks as well as to comply with other relevant privacy guidelines, also reviewed in this paper. The proposed platform aims at: (i) maximizing potential networking benefits, (ii) establishing connection with minimal effort of end users and (iii) preserving users' privacy.
- Lastly, we ran a qualitative investigation with 15 participants to investigate

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

whether their perception of personal privacy protection is enhanced when the design of ubiquitous social networking avoids the four identified drawbacks.

The rest of the paper is structured as follows: in the next section we review existing privacy guidelines for ubiquitous computing environments. In Section C.3, we present existing ubiquitous social networking solutions and analyze them as case studies for the reviewed privacy guidelines. In Section C.4 we introduce the four drawbacks to be avoided when designing privacy-aware ubiquitous social networking services and in the Section C.5 we describe the design of our proposed privacy-aware social networking platform. We review the design and background of the qualitative investigation, information about the participants and present the results of the qualitative tests in Section C.6. Conclusions and recommendations for future work are drawn in the last section of the paper.

C.2 Privacy design guidelines

Previous studies have already proposed relevant design guidelines to prevent potential privacy threats that could discourage users from providing their personal information in ubiquitous computing (ubicomp) environments [12, 104, 120, 122, 149]. These guidelines target at empowering users to make deliberate personal data disclosure decisions, which is also a crucial goal for socUbicomp environments. In this context, we consider them to be directly applicable in the design of privacy-aware platforms for ubiquitous social networking services. Thus, in this section, we use the original term *ubicomp*, however, it could be interchanged with the term *socUbicomp* without loss of adequacy.

In order to achieve the target of allowing ubicomp users to take informed data disclosure decisions, Lederer et al [122] suggested that the socio-technical gap, introduced by Ackerman [1], should be addressed. The socio-technical gap refers to the division between “what we know we must support socially and what we can support technically”. If an intermediary point of the socio-technical gap is not found, the user would be either overwhelmed or disempowered, which would both result in uninformed and impulsive data disclosure choices. To find this balance, Lederer et al [122] proposed that the design of ubicomp environments should be focused on enabling users to both understand privacy implications of data disclosure as well as allowing them to perform natural social actions. Consequently, they provided privacy guidelines that target at reconciling Palen and Dourish’s theoretical insights [149] with Bellotti and Sellen’s [12] technical solutions.

The theoretical insights of Palen and Dourish [149] were inspired by the work of Altman [4, 5], who describes privacy as a dynamic process, representing continuous negotiation and management of the boundaries that shape data disclosure. Palen and Dourish [149] identified three dynamic boundaries for negotiation of users’ personal data disclosure. Firstly, the privacy and publicity boundary separates personal information into the disclosed and retained data sets. Then, the identity boundary defines the role, represented by the user based on the time, place and situation contexts. Finally, the temporal boundary regards the past, present and expected future of the users. The authors concluded that data disclosure decisions are taken by continuously negotiating the internal conflicts between the elements of the three identified boundaries. Bellotti and Sellen [12], instead, focused on more practical solutions. They firstly introduced potential privacy threats for ubicomp

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

environments, such as disembodiment and dissociation. The former one refers to the danger that users would not be able to present themselves to others as they do in face-to-face interactions. Dissociation refers to the threat that the results of actions are shown while the actions themselves are invisible. Consequently, they proposed control and feedback principles in order to avoid disembodiment and dissociation privacy threats by allowing the user to decide what to disclose and whom to disclose, as well as ensure subsequent feedback about their data disclosure decisions.

Moreover, in their privacy guidelines, Lederer et al [122] try also to honor the fair information practices, outlined by Langheinrich [120] as well as attempt to encourage minimum information asymmetry between the parties (i.e. data owner, data collector and data user), as proposed by Jiang et al [104]. Particularly, based on the legal regulations of US Privacy Act of 1974 [82] and European Union's Directive 95/46/EC [62], Langheinrich [120] identified several main areas of innovation and system design for privacy protection in ubicomp: notifying the user appropriately; taking into account the user's choice and seeking for consent; enforcing limitation of scope within the concepts of proximity and locality; enabling anonymity and pseudonymity when necessary; providing adequate security and appropriate data access. Further, in order to reduce information asymmetry, Jiang et al [104] proposed to either decrease the flow of information from data owners to data collectors and users or otherwise to increase the flow of information back to the data owner.

As a result, the privacy guidelines proposed by Lederer et al [122] depict five pitfalls to be avoided in the design of privacy management systems for ubicomp environments:

1. Obscuring potential information flow: ubicomp should not obscure the nature

and extent of data disclosure. Users should easily comprehend, for example, what kind of information is disclosed and to whom, how the information is shared, the presence of third-party observers and the potential for unintentional disclosure. Avoiding this pitfall would allow users to understand the scope of the privacy implications in ubicomp environments.

2. Obscuring actual information flow: ubicomp should not obscure the actual disclosure of information. The disclosure should be obvious to the user as it occurs, however without overwhelming his attention. When immediate notice is not feasible, then it must be ensured with a reasonable delay. Avoiding this pitfall would allow users to understand what information is being disclosed to whom.
3. Emphasizing configuration over action: ubicomp should not require exaggerated manual configuration to manage personal privacy. Instead, users' privacy should be managed as a natural consequence of their normal engagement with the environments. Avoiding this pitfall would enable users to control their privacy without requiring tremendous configuration.
4. Lacking coarse-grained control: ubicomp should not forgo a binary choice for halting and resuming data disclosure. Avoiding this pitfall would empower the user to effectively control their participation in ubicomp environments.
5. Inhibiting established practice: ubicomp should not inhibit users from transferring established social practice to emerging technologies. For example, ubicomp should enable disclosure of ambiguous information as well as ensure

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

plausible deniability. Avoiding this pitfall would allow the users to participate in ubicomp environments without compromising their ordinary social behavior.

C.3 Ubiquitous social networking designs

In this section we present existing designs of ubiquitous social networking applications and highlight important usability and privacy limitations that might affect users' satisfaction as well as protection of their data privacy. Firstly, we introduce and discuss the design of three ubiquitous social networking applications that fall into at least one of the five pitfalls. Afterwards, we present a design solution, which follows the reviewed privacy guidelines. Finally, we identify additional privacy limitations that were not addressed in the privacy guidelines, presented in Section C.2.

C.3.1 Negative case studies for the privacy pitfalls

The first design solution that we review is implemented for the Nokia Sensor [153] mobile application. Nokia Sensor relies on a decentralized architecture and exploits dynamic mobile connectivity in terms of peer-to-peer communications, as shown in Figure C.1.

Nokia Sensor is designed to discover and identify other devices in the users' proximity and exchange users' profiles as well as other personal contents (e.g. pictures, messages, etc.), stored in the local memory of the mobile phones. Although this design solution provides opportunities for getting in touch with people nearby,

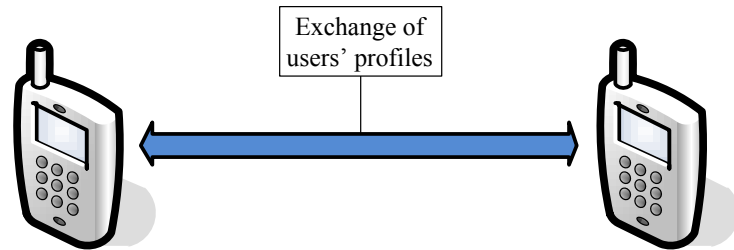


Figure C.1: Nokia Sensor design solution

it requires too much user intervention and comprises important privacy threats. Firstly, users must manually invoke methods for discovering other nearby people and importantly it also requires manual profile comparison for finding users with similar interests. Secondly, Nokia Sensor design is subject to potential privacy concerns, because users must disclose their profiles to everyone in their proximity, as a prerequisite for participation in those socUbicomp environments.

Beale [10] proposed a design solution, implemented in the Bluedating application, which overcomes some of the mentioned Nokia Sensor limitations. In fact, Bluedating releases the users from frequent interactions with their mobile devices by enabling an automatic method for discovery of other users nearby. Furthermore, the application does not require manual profile comparison for finding users with similar interests, because it automatically calculates users' similarities and it notifies the users when others with similar interests are in their proximity. In order to find profile similarities, Bluedating requires users to fulfill two profiles: the personal profile and the one that they are interested in for networking. While the former stores information that a user is willing to share, the latter is composed of information that a user is looking for, which can be compared to a wish list of the user. As illustrated in Figure C.2, when Bluedating users encounter each other,

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

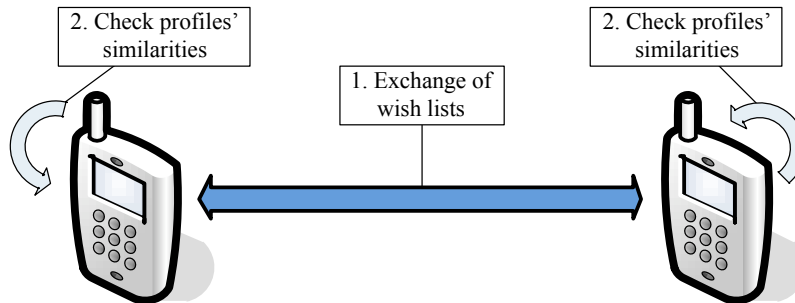


Figure C.2: Bluedating design solution

they exchange their wish lists (step 1) in order to calculate the similarities between their personal profile and the received wish list of the person nearby (step 2). If a match between the two profiles exists, both users are notified. For Bluedating no personal user profile is exchanged as a result of the discovered similarities, therefore the users must manually initiate the connection themselves. Even if this design solution overcomes many of the limitations incorporated in Nokia Sensor, it still presents important privacy and usability concerns. Firstly, Bluedating is subject to potential invasion of privacy, because users must accept to disclose their wish lists to all the people in their proximity. Secondly, especially in crowded places, it still requires considerable user intervention, because the application only informs about potential profile similarities, however does not initiate any connection between the users.

Tamarit et al [183] presented a mobile application, named BlueFriend, which utilizes a different approach for promoting ubiquitous social networking. In order to find similarities between the users, the application splits personal information of the user in two main categories: public and private profiles. While the former includes information generally not sensitive (e.g. users' preferences) the latter con-

C.3 Ubiquitous social networking designs

sists of data that users prefer to keep private (e.g. contact details). As shown in Figure C.3, when BlueFriend users encounter each other, they firstly exchange their public profiles in order to find similarities between them (step 1). Thereafter, if the matching index is found to be above a threshold value defined by both users (step 2), they also exchange their private profiles (step 3). Indisputably, this design solution succeeds on addressing the previously discussed usability limitations, because it requires considerably less user intervention than Nokia Sensor and Bluedating. In fact, BlueFriend provides automated user discovery as well as initialization of connections between people with similar interests. Moreover, BlueFriend avoids many of the discussed privacy threats by broadcasting only anonymous users' preferences to all others in their vicinity, unless relevant profile similarities have been discovered. However, such design solution still presents important privacy limitations as users cannot customize their data disclosure decisions, because they must share a static profile. Thus, they are incapable of manipulating their data sharing preferences in order to present different subsets of personal information, when desired. For example, if users are asked to disclose where they live, they do not have the

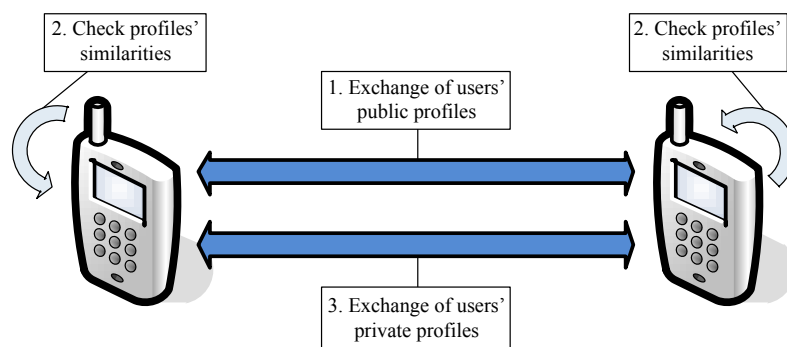


Figure C.3: BlueFriend design solution

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

option for choosing either to share their neighborhood or the exact address.

Due to the highlighted privacy limitations, Nokia Sensor, Bluedating and BlueFriend designs do not follow the privacy guidelines, already introduced in the previous section. Considering the existing proposals discussed in this section as case studies for the privacy pitfalls, it can be noticed that they fall into the majority of them, and thus they do not provide the possibility for users to take informed data disclosure decisions. As shown in Table C.1, Nokia Sensor and Bluedating designs obscure the potential (Pitfall 1) and actual (Pitfall 2) information flows, because they broadcast personal information to the other nearby users. Nokia Sensor, in particular, finds users' similarities by disclosing full user profiles, while Bluedating shares wish lists. Importantly, there is no mechanism for selecting a subset of potential recipients of the mentioned data disclosure, thus obscuring the potential information flow. Furthermore, no notification or log of the data disclosure is provided to the user, thus also obscuring the actual information flow. Contrarily, BlueFriend avoids Pitfall 1, because it is deliberately scoped to disclose private information only to other users that present relevant profile similarities. Moreover, this design also conveys the actual information flow, as users are aware of who has access to their identifiable information, due to the mutual exchange of private data, oc-

Table C.1: Evaluation of the socUbiComp applications in regard to the privacy pitfalls, described in Section C.2

| Ubiquitous social networking application | Pitfalls | | | | |
|--|----------|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| Nokia Sensor | X | X | | | X |
| Bluedating | X | X | | | X |
| BlueFriend | | | | | X |

curing after the discovery of relevant profile similarities between anonymous users' preferences. All the previously reviewed solutions do not fall into Pitfall 3 and Pitfall 4. They require only minimal configuration (Pitfall 3) for maintaining privacy, e.g. only the necessity to indicate predefined data disclosure decisions. According to the system design, once users have fulfilled their static profiles, no additional privacy configurations are needed. Nokia Sensor, Bluedating and BlueFriend also provide coarse-grained controls (Pitfall 4) for halting and resuming information flow, e.g. users can utilize application exit or mobile phone power buttons to effectively control their participation in these socUbiComp environments. Finally, all these approaches fall into Pitfall 5, as they do not help users to express themselves with a natural and socially meaningful behavior, forcing users to convey a static profile, thus preventing them to arbitrarily customize their disclosed preferences.

C.3.2 Positive case study for the privacy pitfalls

A design solution that might overcome the five pitfalls is the Serendipity application [65]. The software architecture of the Serendipity is shown in Figure C.4. Differently from the other reviewed solutions, Serendipity is designed to utilize a central unit for finding similarities between users in socUbiComp environments. When Serendipity users randomly meet, they only exchange their Bluetooth identification (step 1). This information is sent to the central server (step 2), which contains all the Serendipity users' profiles along with the matchmaking preferences. The server evaluates similarities between encountering users (step 3), and if the similarity score is higher than a threshold value, identified by both users, the server notifies them about their presence, related affinities and contact information (step 4).

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

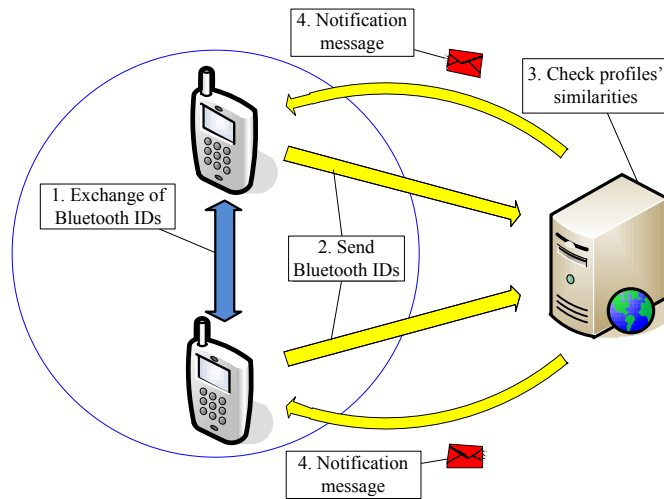


Figure C.4: Serendipity design solution

Similarly to BlueFriend, the Serendipity solution does not obscure the potential (Pitfall 1) and actual (Pitfall 2) information flows, because information disclosure occurs only when a profile match is found between the two users, who are clearly notified about it. Like Nokia Sensor, Bluedating and BlueFriend design solutions, Serendipity also avoids Pitfall 3 and Pitfall 4, because it does not require excessive manual configuration to manage personal privacy and it provides means for halting and resuming users' participation in those environments. Finally, Serendipity also overcomes Pitfall 5, because it allows users to express themselves with a natural and socially meaningful behavior by enabling a dynamic sharing of profiles that vary depending on the different similarities with the encountered users.

C.3.3 Additional privacy limitations

As discussed, the Serendipity design successfully avoids the five pitfalls, described in Section C.2, and users are actually empowered to make informed data disclosure decisions. However, other important privacy limitations can be identified, by carefully analyzing the Serendipity design solution. Firstly, we refer to unnecessary data disclosure to third parties, which might provoke crucial privacy threats and consequently discourage users from participation in socUbiComp environments [6, 84, 91, 130]. Note that by considering third parties, the design shall also address the topic of trust. Although it is a very important topic in the current context, we do not discuss it in this paper not to risk extensive considerations that would distract the reader from our main topic, privacy. Moreover, Serendipity similarly to the other approaches does not balance privacy concerns with potential networking benefits. Its design assumes that the sensitivity of the shared personal information is decreased due to the discovered similarities between the encountering users. However, this solution does not take into account situations where users might feel uncomfortable to disclose some of their personal information in some specific circumstances (e.g. data related to private activities in work environments), even if they share the same preferences with the encountered user. As a result, Serendipity would probably lead to either potential invasion of privacy or threatening of potential networking benefits. If the users would fulfill a detailed profile for maximizing potential networking benefits, it could result in disclosure of personal information, which is sensitive under certain circumstances. For example, a user may include his sexual orientation in his profile, because he considered this data to be not sen-

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

sitive in non-working environments. However, sharing the user sexual orientation would invade his personal privacy, if that occurred in work environments where this information was considered to be highly sensitive for him. Contrarily, in case the users would provide only a limited profile in order to avoid to include personal information, which is too sensitive to be disclosed under some circumstances, it could result in loss of potential networking benefits. Considering the above example, if the user decides not to include his sexual orientation in his profile, because it was considered to be highly sensitive in work environments, then the user would also lose potential networking benefits in non-working environments, where such personal data was not considered to be sensitive. Finally, Serendipity, as well as all the aforementioned design solutions, also lacks attention to an important privacy principle, called individual participation that is already incorporated into all major privacy laws worldwide, such as [62, 82]. The individual participation is an essential privacy protection principle for personal data disclosure regarding the right of the user to always be able to see and correct any data disclosure decisions. For all the discussed solutions, users do not have possibilities to modify data after actual disclosure, and consequently, they permanently lose control over their disclosed personal information.

C.4 Privacy drawbacks to be avoided by designers

Among all the reviewed design solutions for enabling ubiquitous social networking services, Serendipity can be considered as a positive case study that might overcome the five pitfalls. It achieves the primary goal of enabling users to make

C.4 Privacy drawbacks to be avoided by designers

informed data disclosure decisions. However, in the previous section we have identified and discussed additional design limitations in Serendipity as well as in the other reviewed approaches, which were not taken into consideration in the existing privacy design guidelines, described in Section C.2. Thus, we propose four additional drawbacks that should be avoided when designing ubiquitous social networking platforms. These new design guidelines do not target at replacing the five pitfalls, proposed by Lederer et al [122], but they are considered as additional guidelines for the design of socUbiComp environments. By avoiding the drawbacks introduced in this section, the designers would potentially allow more functional ubiquitous social networking, oriented to prioritizing better protect of end users' privacy. The four drawbacks are:

1. Ignoring the variation of human data sensitivity: Ubiquitous social networking should not disclose personal information without taking into consideration the human data sensitivity of the current circumstances. Instead, it should target at miming ordinary face-to-face communications, in which individuals intuitively evaluate various determinants and unconsciously choose what personal information to share [170, 171]. Consequently, different sets of personal information should be disclosed upon different circumstances. Avoiding this drawback would allow to prevent invasion of users' privacy as well as motivate users not to detain their personal information, considered as too sensitive to be shared in some circumstances.
2. Embracing disclosure to third parties: Ubiquitous social networking should avoid disclosure of users' personal information to third-party entities. In-

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

stead, disclosure should occur towards other users, whose profiles might lead to potential mutual interests and networking benefits. Avoiding this drawback would minimize important privacy concerns about data disclosure to third parties that have already been raised in regard to online social networks [6, 84, 91, 130, 168]. Thus, users would be ensured that their personal information is disclosed only when they might receive potential networking benefits in exchange.

3. Requiring too much user intervention: Ubiquitous social networking should not require too much user intervention. Connections between users must be created with minimal efforts of end users, thus allowing technologies to operate seamlessly in the background [191]. Avoiding this drawback would embrace Weiser's vision of a calm technology [193] and consequently would enable the development of more user-friendly and service-oriented ubiquitous social networking environments.
4. Lacking user's personal data control: Ubiquitous social networking should not lead to loss of permanent control over personal data. As enforced by the legal regulations [62, 82], users should always have opportunities to modify any piece of information even after actual data disclosure [146, 167, 168]. Moreover, if desired, ubiquitous social networking should empower users to keep their data disclosure decisions up to date to all the relevant peers without any additional effort, by synchronizing the modified information to all the users, who have access to that specific updated data. Avoiding this drawback would empower users to effectively control their personal data at any time,

C.4 Privacy drawbacks to be avoided by designers

which might also increase potential networking benefits to an even greater extent by allowing user information to be continuously updated.

C.4.1 Negative case studies for the privacy drawbacks

When taking into consideration the design of ubiquitous social networking applications, described in Section C.3, as case studies for the privacy drawbacks, we can notice that all the solutions fall into at least two of the four identified drawbacks.

As shown in Table C.2, none of the applications take into consideration the variation of human data sensitivity (Drawback 1), because they do not have any privacy control mechanism for customizing personal data disclosure, based on the different encountered circumstances. For example, Nokia Sensor, Bluedating and BlueFriend disclose the same personal information in all the circumstances, while Serendipity only customizes users' data disclosure based on related preferences between the encountering users, rather than taking into consideration the variation of human data sensitivity in different circumstances. Secondly, Serendipity is the only mobile application that falls into the Drawback 2, because its design requires disclosure of personal information to a third-party component, in order to calculate similarities between the encountered users. On the contrary, all the other design

Table C.2: Evaluation of the socUbiComp applications in regard to the four privacy drawbacks

| Ubiquitous social networking application | Drawbacks | | | |
|--|-----------|---|---|---|
| | 1 | 2 | 3 | 4 |
| Nokia Sensor | X | | X | X |
| Bluedating | X | | X | X |
| BlueFriend | X | | | X |
| Serendipity | X | X | | X |

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

solutions avoid this drawback by relying on decentralized architectures that disclose personal information to only end users. Thirdly, Nokia Sensor and Bluedating also require considerable amount of user intervention (Drawback 3). Each time that a Nokia Sensor user would like to discover others nearby, they must manually invoke a Bluetooth discovery and carry out a manual profile comparison, while Bluedating user are forced to immediately initiate face-to-face relationships with the encountered users at the moment of discovery of profile similarities. Contrarily, Serendipity and BlueFriend do not require too much user intervention as they provide automatic discovery and profile comparison between the encountering users as well as means for initiating the relationship between users even at later time. Lastly, all the approaches fall into Drawback 4 as they do not provide any further control over users' personal data after actual disclosure.

C.4.2 Interdependencies among the four drawbacks

Privacy designers are advised to carefully evaluate how to avoid these drawbacks, because by successfully heeding one might result in the risk of falling into another. Designers of ubiquitous social networking should find solutions that avoid the proposed drawbacks by also taking into consideration the interdependencies between them. When that is not possible, designers are challenged to find solutions that would represent an optimal trade-off between the risks that the drawbacks might impose. In the following, we discuss two examples of interdependency between the drawbacks.

Drawback 1 vs Drawback 3 The first interdependency that we discuss is between ignoring variation of human data sensitivity and requiring too much user intervention. In order to evaluate what personal information is sensitive or not for the current circumstances, ubiquitous social networking users might be requested to disclose their data at the moment of the actual disclosure. However, such kind of approach would indisputably need a considerable amount of users' attention and intervention. As an example, a privacy model called Precision Dial [122] supports four predefined levels of privacy protection, ranging from "undisclosed" that defines absolute confidentiality to "precise", which allows openness of entire user's personal information. While encountering different circumstances, the user has the opportunity to manually adjust his privacy settings when needed. Despite achieving the goal of avoiding Drawback 1, Precision Dial still demands considerable amount of users' attention and intervention, as users are continuously required to adjust their precision settings, similar to the practice of adjusting ringer volume of mobile devices (falling into Drawback 3). Contrarily, all the approaches that can be considered as user-friendly ubiquitous social networking applications and thus avoiding Drawback 3 (e.g. Serendipity and BlueFriend) do not provide opportunities to customize users' personal data disclosure by taking into consideration the variation of human data sensitivity (falling into Drawback 1).

Drawback 2 vs Drawback 4 The second interdependency that we discuss is between embracing third parties and lacking user's personal data control. All the design solutions, described in Section C.3.1, do not disclose users' personal information to third parties. In fact, they rely on a pure peer-two-peer architecture and

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

the sharing of information exactly replicates an exchange of paper business cards (avoiding Drawback 2). However, when in close proximity, users can wirelessly exchange their personal profiles, but they cannot thereafter modify them (falling into Drawback 4). Contrarily, a centralized architecture might be ideal for enabling users to modify their personal data, even after actual disclosure. In fact, the third-party entities might store users' personal data disclosure decisions to be accessed, updated, synchronised with all the relevant encountered users, or even removed at any time (avoiding Drawback 4). However, this kind of solution would indisputably require to embrace disclosure to third-party components (falling into Drawback 2).

C.5 Privacy-aware platform design

In this section we present the design of a privacy-aware ubiquitous social networking (PAUSN) platform that focuses on promoting ubiquitous social networking services during physical meetings, while preserving the users' privacy. PAUSN is designed to overcome both privacy pitfalls and drawbacks, described respectively in Section C.2 and in Section C.4. In order to comply with the existing and proposed privacy guidelines, PAUSN utilizes a trusted third-party entity that receives encrypted (thus incomprehensible) profiles from the encountering users in order to calculate the profile similarities. By relying on trusted servers, PAUSN prevents other users from executing brute force discovery of encrypted data. Moreover, the third-party is capable of comparing the encrypted profiles and computing similarities between the two users, as profiles are ciphered with the same security key. When the similarity scores, defined by the users, exceed their corresponding threshold values, users

are notified and their personal information is disclosed to each other, however it still remains incomprehensible for the third-party. At any time, PAUSN users are empowered to modify their data disclosure decisions, thanks to its centralized architecture. The remainder of this section shall provide the necessary insight to better understand the concept of PAUSN.

C.5.1 User profile management

The main focus of PAUSN is to maximize potential networking possibilities while preserving users' personal privacy. In order to achieve this goal, PAUSN utilizes two users' profiles, i.e. Unified User Profile (UUP) and Business Card (BC). As shown in Figure C.5, the UUP is a collection of all personal data of the user that includes for instance online social networks profiles, Internet activities, etc. Contrarily, the BC is a subset of the UUP, composed of user's personal information that is relevant, but not sensitive, for the current circumstances. The circumstances of the users' encounters can be defined based on the user's location, mood, activity, identity of

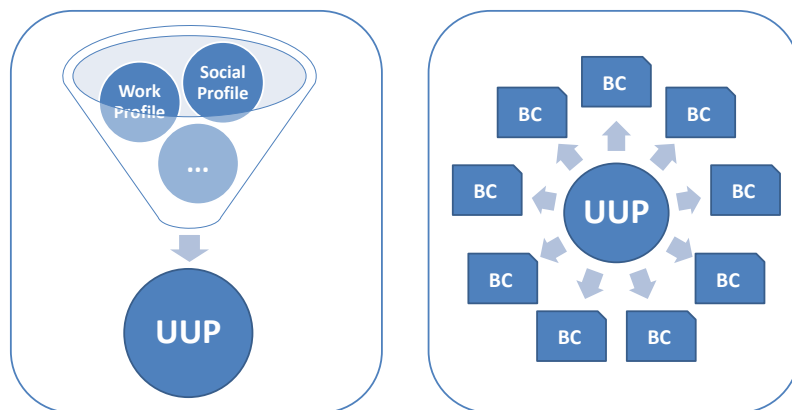


Figure C.5: Unified User Profile and Business Cards

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

the other users, etc.

In order to evaluate users' similarities, we considered different techniques, illustrated in Table C.3. These strategies attempt to achieve an optimal balance between potential networking benefits and privacy concerns. The first strategy considers the comparisons of users' BCs to minimize potential privacy concerns. When identifying relevant profile similarities between users' BCs, the USNs services would disclose the BCs to the users, which are composed of not sensitive personal data for the current circumstances. However, comparing only subsets of personal data would probably lead to the minimization of users' networking opportunities, as the discovery of profile similarities would be relevantly reduced.

The second strategy attempts to maximise potential networking benefits by comparing the users' UUPs and disclose them if relevant profile similarities were found. However, this option was also rejected, because it presents significant disadvantages. Given that the UUPs are complete users' personal profiles, they might include sensitive data for the current circumstances and disclosing such profiles might result in invasion of users' personal privacy. Moreover, this kind of profile management might be also subject to potential attacks to users' personal privacy by malicious

Table C.3: Comparison of four different strategies for identifying profile similarities between encountering users

| | Strategy | Sharing | Disadvantages | Advantages |
|---|-----------------|----------------|--|--|
| 1 | BC - BC | BCs | low networking benefits | low privacy threat |
| 2 | UUP - UUP | UUPs | high privacy threat | high networking benefits |
| 3 | UUP - UUP | BCs | misinterpretation of user similarities | low privacy threat high networking benefits |
| 4 | UUP - BC | BCs | | low privacy threat high networking benefits |

users, who target at sniffing others' personal information by including a rich set of data into their UUPs. In such case, a malicious user can opt for a whole set of preferences with the goal of receiving everybody's personal information.

In order to prevent such potential privacy concerns arising from the disclosure of users' UUPs, we also took into account the possibility to compare the two UUPs for still maximising potential networking benefits, but disclosing only users' BCs if relevant profile similarities between the two UUPs were discovered. This third strategy was also neglected, because a match found between users' UUPs might not be perceived in the users' BCs. For example, users might present high relevant profile similarities in their UUPs, however if such similarities were not included in their shared BCs the received BCs would not be useful to the users. In this case, the users would not be aware about their common similarities and they would not be encouraged to networking.

The fourth strategy, illustrated in Table C.3, compares the UUP of a user with the BC of the other (and vice versa). This technique increases potential networking benefits in comparison to the first strategy that discovers profile similarities only between the users' BCs. Further, it decreases potential accidental invasions of personal privacy in comparison to the second strategy. In fact, if relevant profiles similarities were found, the fourth strategy would share to the users only their BCs and not complete users' profile, as the second strategy does. Furthermore, when comparing to the third strategy, the last option might present disadvantages in relation to the maximization of potential networking benefits. We believe that this is a necessary compromise for establishing a more efficient and calm USN technology, as the fourth strategy connects users only if the discovered profile similarities were

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

also included into their BCs. Thus, differently from the third technique, users would only receive the BCs of others that contain matching profile similarities with their UUPs. This would significantly motivate them to start an initiation of relationship.

Finally, this kind of profile management also protects users' personal privacy against users, who maliciously include a rich set of data into their UUPs and BCs for sniffing others' personal information. Even if the malicious users might relevantly increase the opportunity to establish more connections when encountering others, they would not have access to users' sensitive data. The malicious users would only receive the others' BCs, which are composed of personal information considered relevant, but not sensitive, for the current circumstances. More details about the selection of personal data to be included into the users' BC is provided in Section C.5.3.

C.5.2 Communication flow

In order to better explain the design of PAUSN, we describe a scenario, shown in Figure C.6. In our example, we present two users, Bob and Alice. When Bob and Alice enter each other's wireless range, they exchange their public keys through direct ad hoc links, i.e. Bob sends to Alice his public key (Q_B) and Alice sends to Bob her public key (Q_A), as shown in Figure C.6-A. Afterwards, in Figure C.6-B, Bob encrypts his UUP using his Q_B and his BC using Alice's public key, i.e. Q_A . Accordingly, Alice encrypts her UUP using her public key, i.e. Q_A and her BC using Bob's public key, i.e. Q_B . Finally, the encrypted profiles are submitted to a trusted third-party (step 1) using their own broadband infrastructure link, i.e. Internet connection. Even if the third-party is not able to access the contents

C.5 Privacy-aware platform design

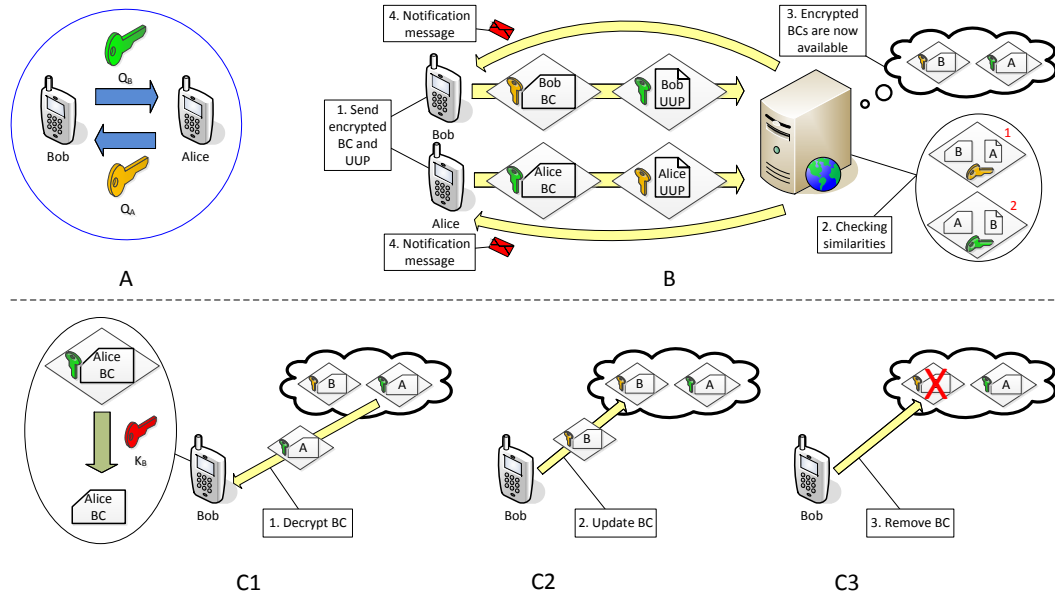


Figure C.6: PAUSN design

of the encrypted profiles, it checks for similarities by comparing Bob's UUP with Alice's BC, as both profiles are encrypted with the same Bob's public key, i.e. Q_B (see Section C.5.3 for details). The third-party also compares Bob's BC with Alice's UUP, both encrypted with the Alice's public key, i.e. Q_A (step 2). If no match is found, the third-party deletes the received encrypted profiles. Otherwise, if similarities are found, the third-party deletes the users' UUPs and stores only the encrypted BCs of Bob and Alice (step 3). Afterwards, the third-party sends a notification to both users (step 4).

As illustrated in Figure C.6-C1, Bob is now able to access Alice's BC from the third-party entity and decrypt it by using his private key, i.e. K_B , as Alice's BC was previously encrypted using Bob's public key, i.e. Q_B . Even if not shown, Alice is also able to access and decrypt Bob's BC using her private key. Moreover, the users

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

are always able to modify or revoke their data disclosure decisions. For example, if Bob is interested to networking with Alice, he can modify his BC by encrypting his new BC with Alice's public key (i.e. Q_A) and replacing it on the server, as shown in Figure C.6-C2. The contents of the new Bob's BC would still be inaccessible for the third-party and available only for Alice. Finally, as shown in Figure C.6-C3, Bob has the opportunity to revoke his BC. In this case, any information about Bob will not be available anymore, Alice would be optionally notified and the link between the two users will be deleted. When desired, any piece of information that is updated or removed by Bob can be effortlessly synchronised not only with Alice, but also with all the other users who have access to that specific information.

C.5.3 Analysis of privacy guidelines

The design of PAUSN attempts to overcome all pitfalls and drawbacks described in the previous sections. In the following, we firstly discuss the design solutions that enable PAUSN to overcome the four identified drawbacks. Afterwards, we describe how our proposed platform copes with the identified interdependencies between the drawbacks. Finally, we also analyze PAUSN as a case study for the privacy pitfalls, in order to evaluate whether it provides means for users to take informed data disclosure decisions.

Avoiding Drawback 1 - Ignoring the variation of human data sensitivity

PAUSN does not ignore the variation of human data sensitivity, because it is designed to disclose different business cards under different circumstances. The shared business card is intended to be composed of relevant, but not sensitive, personal

information. In order to achieve this goal, two main approaches can be considered: predefined privacy preferences and ad hoc privacy control. The former attempts to predict all the potential situations and associated data sharing decisions. An example of a predefined privacy preferences model is Faces [124] that, similarly to [103, 111, 144, 176], allows users to indicate *who* can access *what* and *when* before the actual data disclosure. Further, this approach was upgraded to ad hoc privacy control solution, due to the possibility to encounter situations where data disclosure decisions are not accurately predictable in advance [26, 27, 103]. The ad hoc privacy control solution provides opportunities to take data sharing decisions *in situ* - at the moment of actual disclosure [26, 27, 103]. Examples of ad hoc privacy control models are Precision Dial [122], Disclosure Decision Model [26, 27] and Diverged Personalities (DiP) [171].

Among these privacy mechanisms, DiP is ideally suited for the design of PAUSN, because it targets at reflecting on user's natural privacy handling upon disclosure of personal data without overwhelming the user's attention. In the DiP privacy model, the UUP is diverged into different users' personalities (or BCs) to be presented under different circumstances. The most suitable business card for each circumstance is generated by the process shown in Figure C.7. The central component of the DiP is the Personality Logic, which receives as input the user's UUP that is composed of a collection of various available user's profiles. Moreover, it also processes information about the inquirer and context data, which is used to define the current circumstances. This information is acquired thanks to the sensing components of the mobile devices, e.g. Bluetooth, accelerometer, GPS, microphone [139]. Based on these three inputs, algorithms automatically provide the best BC by taking

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

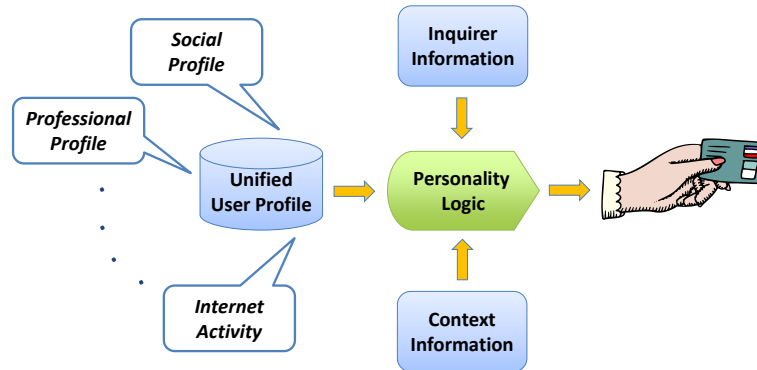


Figure C.7: Diverged personalities model process

into account both previous data disclosure decisions and influential factors, already identified to impact users' data disclosure decisions in previous investigations. For instance, we refer to the following influential factors, already identified to impact users' personal data disclosure decisions in ubiquitous social networking: identity of the inquirer [57, 123, 199], current environment [169], activity [108, 167], mood [45, 169], location familiarity [167, 169], purpose of disclosure [45, 167, 169], number of previous meetings and mutual friends with the inquirer [167, 169]. In [169], a simulation of the DiP privacy model on users' data disclosure prediction achieved significant results with an approximate accuracy of 90%, with peaks of 93%, and potential for further increasing performance.

Avoiding Drawback 2 - Embracing third parties

PAUSN does not disclose any personal user information to the third-party entities, because users send to them encrypted, thus incomprehensible, profiles. In order to encrypt users' profiles, i.e. UUPs and BCs, the public-key (or asymmetric) cryptographic scheme is adopted. This cryptography method uses two different

keys, referred to as public and private keys. The public key is used for encrypting data and it can be sent to anyone. Contrarily, the private key is used for decrypting the data and it is never revealed to anyone. Such mechanism typically achieves higher security by compromising the length of the encryption keys and subsequently its intrinsic computation complexity. This may result in significant challenges if attempting to apply it on mobile devices due to their limited battery, computation and communication capabilities. Thus, it is desirable that the encryption keys are kept short. This is the case for instance of the Elliptic Curve Cryptography (ECC), where a 160-bit key is considered to be as secure as 1024-bit key in Rivest Shamir Adleman (RSA) cryptography [113]. There are quite a few possible cryptographic schemes related to ECC [41, 88, 131, 188], which could be used for mobile devices, thus suitable for PAUSN.

In order to be compared, users' profiles must be partitioned before being encrypted. For this reason, the concept of atomic data parts is introduced. This concept regards the minimal indivisible segment of user profile data. As depicted in Figure C.8, the atomic data parts are structured in a pair of key and value, where the key identifies the data category and the value identifies the actual profile data. The corresponding values can either be limited (e.g. gender, age) or unrestricted (tallness, preferred music bands) in the number of options. When the number of options is low, profile comparisons may match, however, when the number of options is fairly large, it can result in unfeasible comparisons. Thus, we propose the options to be limited, lowered and grouped when possible. In order to achieve this clustering, it is suggested to encourage a vocabulary of options when implementing the social network. Figure C.8 implicitly shows an example of a vocabulary, where

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

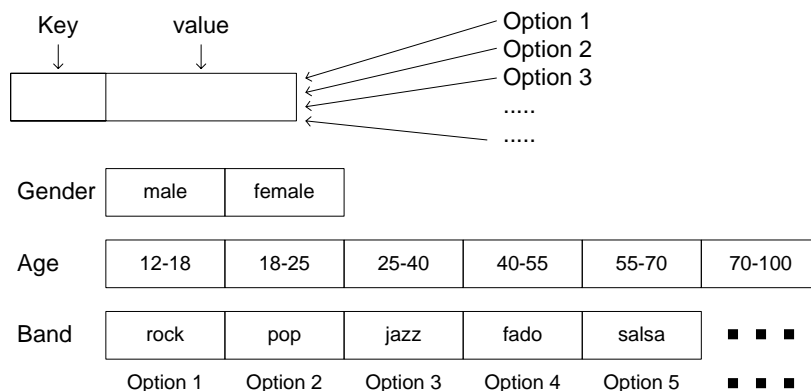


Figure C.8: Structure of the atomic parts

options for preferred bands, age range and gender are presented. Thus, the BC and the UUP comprise a list of atomic data parts where each encrypted key indexes a vector with all the encrypted entries for that specific key. For instance concerning the atomic part ‘gender’, the indexed vector would have a single entry, while for the atomic data part ‘band’, the indexed vector might be larger.

Given that the atomic data parts are all properly encrypted with the corresponding public key and that the entire data is organized into the corresponding BC or UPP, the comparison is possible on “per atomic data part basis”. The important aspect of this method is that the compared UPP (of the first user) and the BC (of the second user) are encrypted with the same public key, according to the method described in Section C.5.2. Thus, the data remains incomprehensible to the third-party, who can only perform direct comparison among atomic data parts.

Avoiding Drawback 3 - Requiring too much user intervention

In order to reduce the amount of user interaction needed to participate on the ubiquitous social networking services, the PAUSN platform automates the process of finding users with similar interests and the subsequent establishment of the connection. When a match between two users is found, the platform notifies both users and it automatically establishes the connection between them.

As previously mentioned in Section C.5.2, two users submit the encrypted UUP and BC to the server. At the server, the encrypted atomic data parts are extracted from the UUP of the first user and from the BC of the second user. A direct matching of keys and corresponding comparison of values is executed. This comparison is made by matching the several indexes of the atomic data parts of the BC of the second user and the UUP of the first user and subsequently compare the corresponding values. Given that the individual atomic data parts are encrypted, the only possible results are either a full or no match for each atomic comparison. By calculating the number of atomic matches against the total number of data parts, the similarity score can be calculated in terms of percentage of similarity. In order to better understand this mechanisms, we introduce an example of a possible algorithm to calculate the similarity score in a format of pseudocode.

```
1. function compare($uup, $bc) {
2.   $score = 0;
3.   for each ($bc as $key => $val) {
4.     if (array_key_exists($key, $uup)) {
5.       $is = array_intersect($val, $uup[$key]);
6.       $score += count($is)/count($val)/count($bc);
7.     }
```

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

```
8.   }
9.   return $score;
10.  }
11. function ismatch($uup, $bc, $threshold) {
12.   return compare($uup, $bc) > $threshold;
13. }
14. $bob_alice_match = ismatch($enc_uup_bob,
    $enc_bc_alice, $threshold_bob);
15. $alice_bob_match = ismatch($enc_uup_alice,
    $enc_bc_bob, $threshold_alice);
```

From an algorithmic perspective, the mechanism resembles a for-loop that scans each and every key within the BC (line 3) list and checks whether the key exists on the UUP list of data parts (line 4). If the key from the BC exists in the UUP, an intersection is performed between the values corresponding to that key within the BC and the values corresponding to that same key within the UUP (line 5). The number of intersections accounts for the sub-score to be given to that index. This sub-score is of course normalized to the number of values in the corresponding index and the number of atomic data parts existing in the BC (line 6). The cumulative sum of all the sub-scores subsequently accounts for the final score to be compared against the threshold defined by the users (line 12). The threshold shall be defined as a percentage of similarity above which the user commits to accept a connection. Finally, line 14 exemplifies the inquiry that Bob automatically sends to the server when he encounters Alice. The query aims at determining whether the similarities between the UUP of Bob and the BC of Alice are above the threshold defined by Bob. Alice does the same as Bob, as it is illustrated in line 15.

Avoiding Drawback 4 - Lacking user's personal data control

PAUSN does not lack personal user's data control, as it follows the individual participation principle. By relying on a centralized architecture, it empowers users to see and modify their personal data, even after actual disclosure. As previously mentioned, when the profile similarity score exceeds the threshold, predefined by both users, a connection between the users is established. From this moment on, both users are allowed to keep a copy of their BCs in the server. These BCs are encrypted by utilizing the public key of the other corresponding user, i.e. the first user encrypts his BC utilizing the public key of the second user and vice versa. This means that the user keeps in his possession the encryption public keys of each user whom he has a connection with. Note that the user must store these public keys as long as he wants to keep the corresponding connections alive. Thus, at any time, users can modify their BC, encrypt it with the public key of the user whom the BC is made available to and upload it again to the server. According to this mechanism, the second user, who is the only person holding the private key, can decrypt that copy of the first users' BC and have access to its contents. Additionally, if desired, any of the two users can simply break the connection by erasing their BCs from the server and deleting the public keys of the other user.

Interdependencies between the drawbacks

PAUSN was also designed to take into consideration the interdependencies between the drawbacks, introduced in Section C.4.2. Firstly, applications of ad hoc privacy control systems should provide an acceptable balance in relation to the interdepen-

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

dependency between Drawback 1 and Drawback 3. In fact, PAUSN takes into account the variation of human data sensitivity (avoiding Drawback 1), while relieving the users from frequent data disclosure decisions (avoiding Drawback 3). This is achieved by applying ad hoc privacy control mechanisms that manage personal privacy on users' behalf. However, it is important to notice that some user intervention should be expected, as users must be always empowered to adjust their data disclosure decisions, in case of disagreement with the automated disclosure decisions. Furthermore, another case that might require ad hoc privacy control to limit its autonomy is the inquiry for highly sensitive data to be disclosed. In such situations, these mechanisms should provide only suggested data disclosure choices, while waiting for user's approval before any actual disclosure [25, 168].

Secondly, PAUSN relies on a centralized architecture that stores only encrypted profiles, which should also provide an acceptable balance in relation to the interdependency between Drawback 2 and Drawback 4. In fact, the proposed platform only discloses personal information to end users (avoiding Drawback 2) and allows personal data control, even after actual disclosure (avoiding Drawback 4). However, this design solution introduces challenges for increasing the possibilities to find users with similar interests. PAUSN might limit discovery of profile similarities, as it might be very difficult to compare complex preferences between the encountering users, due to profile comparison on the encrypted domain. As an example, sophisticated profile matching methods, usually utilized in online dating sites, would be probably difficult to be applicable in our design solution.

Avoiding the five pitfalls

When considering our proposed design as a case study for the privacy pitfalls, described in Section C.2, PAUSN can be considered as a positive case study, because it overcomes all the five pitfalls. In fact, this design solution does not obscure the potential (Pitfall1) and actual (Pitfall 2) flows of disclosed information. The potential information flow is deliberately constrained to an intentional disclosure of personal information between users with interpersonal affinities. The actual information flow, instead, is evident through the immediate notifications from the third-party entity. Moreover, users are also aware about who can access their profiles, because the encrypted business cards are stored by the third party component after discovery of relevant profile similarities. PAUSN does not require extensive configuration to manage personal privacy (Pitfall 3), as it is designed to adopt ad hoc privacy control models that release the user from frequent data disclosure decisions. Similarly to the previously reviewed mobile applications, PAUSN provides a largely coarse-grained control for halting and resuming information flow thanks to application exit and mobile phone power buttons (Pitfall 4). The proposed design also supports existing practices of plausible deniability and ambiguous data disclosure (Pitfall 5). Ambiguous data disclosure is ensured, because PAUSN does not disclose static detailed users' profiles, as application of ad hoc privacy control model allows arbitrary customization of disclosed personal user data. Furthermore, plausible deniability is also supported because one user never knows the true reasons why another user discloses a specific subset of their unified user profile in detriment of other pieces of information. On one hand, the reason may be the different sub-

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

jective levels of the user's privacy perceptions or the different environments, time circumstances or enquirers that justify the disclosure of different pieces of information. On the other hand, the reason may be simply the fact the user is not willing to interact and wants to be left alone.

C.6 Evaluation

In this section, we firstly present the design and methodology of the qualitative investigation, which focuses on the analysis of users' perceptions towards protection of personal privacy in relation to the four identified drawbacks and their respective interdependencies. Afterwards we describe the information about the participants and the results of the investigation.

C.6.1 Background

In order to ensure the validity of answers, we helped participants to get more familiar with the ubiquitous social networking concept. We introduced them the existing applications, described in Section C.3. Further, we presented different scenarios from everyday lives, where these services might be applied, such as professional areas, dating and big events, e.g. conferences and exhibitions, as described in [65, 168]. Finally, we discussed with the participants the potential networking benefits in the identified application areas as well as possible privacy threats that might arise as a result of the information disclosure in ubiquitous social networking, e.g. potential undesired face-to-face interactions [168].

Qualitative interviews were preferred alternatively to other investigation meth-

ods, such as handing out questionnaires or establishing a focus group interview. This method was chosen because of the following two reasons: (i) lack of participants' extensive experience in utilizing these services and (ii) potential misinterpretation of the research questions due to their complexity and ambiguity. Moreover, we decided to run semi-structured interviews to better understand the motivation behind the participants responses and ensure that general areas of information are collected from each participant, however still allowing adaptability of the interview process [50, 118, 136]. We interviewed the participants for a duration of 45-60 minutes. The interviews were audio taped and transcribed at later time. Questions were related to the identified privacy drawbacks and their respective interdependencies. Specifically, per each drawback, we presented two different scenarios of ubiquitous social networking services, which either avoid or fall into the specific drawback. We investigated, whether participants expected that their personal privacy would be better protected when the drawbacks were avoided. Consequently, we asked them whether the features of the design that avoids the specific drawback are:

- Must have: participants would not disclose their data in ubiquitous social networking if their design solutions do not avoid the drawback;
- Nice to have: participants prefer services designed to avoid the drawback, however they would still share some of their information in ubiquitous social networking even in the other case, as the potential networking benefits overcome the potential privacy risks;
- Indifferent: participants do not consider as an advantage if the ubiquitous social networking services avoid the drawback.

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

Finally, we investigated the perceptions of the participants about the PAUSN design characteristics, which were included in order to cope with the interdependencies between the drawbacks. Thus, we inquired them the following questions:

- Are you comfortable with a need for some user intervention in order to balance the trade-off between drawback 1 and drawback 3?
- Are you comfortable with potential reduction of relevant people discovery, caused by the difficulty of applying complex profile matching methods in order to balance the trade-off between drawback 2 and drawback 4?

C.6.2 Participants

The method for selecting the participants of our qualitative investigation, followed the suggestions of Von Hippel, in relation to evaluation of new emerging technology, e.g. USN. The author recommended to contact and interview the most advanced users in the field of interest. The reason for recruiting this type of users, called lead users, is motivated due to their interests and predisposition to innovative product ideas, as they are capable of facing needs long time before others encounter them, and have already found solutions to address potential concerns [97].

In our investigation, the selection of participants was limited to online social networks users. We determined this category to be lead users of the USN technology, because of their advanced experience in disclosure of personal information in online social networks, even if the perception towards the networking services might vary between virtual and physical worlds. We believed that these kind of users have already encountered relevant privacy concerns in relation to the data

disclosure in online social networks and found strategies to balance those concerns with networking benefits. Moreover, we also restricted the selection of participants to the ones who claimed to be potential users of USN services, because they could perceive high potential networking benefits in exchange to their data disclosure.

The potential participants were asked to provide information about their demographic characteristics and to indicate their privacy preferences on visibility of their own personal data (e.g. user profile, pictures, posts) in their main online social networks site. Based on these answers, we were able to observe patterns among data disclosure attitudes and divide the participants into three privacy clusters, following the Westin/Harris privacy segmentation model [198]:

- **Fundamentalists:** these participants were extremely concerned about sharing their personal data with any other online social networks users (friends or strangers);
- **Pragmatists:** these participants also cared about loss of privacy due to the disclosure of their personal information. However, they often had specific concerns and particular strategies for addressing them. For example, this category of participants generally preferred sharing personal information only among their friends;
- **Unconcerned:** these participants were trusting online social networks sites and believing that the privacy of their data was not jeopardized. Thus, they were willing to share their personal data not only with people who were their friends, but as well with users who were complete strangers to them.

When selecting the participants, we aimed to achieve stratification between

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

participants' privacy clusters to ensure that specific characteristics of individuals are represented in the sample in accordance to the proportion in the entire population [75]. Consequently, in this study, we target at obtaining similar proportions of participants' privacy clusters in reference to our latest empirical investigation where a random sample was selected [167]. In total we recruited 15 participants with the following privacy and demographic characteristics:

- Gender: 9 of the participants were male, while 6 of them were females;
- Age: 7 of participants were between 26 and 35 years old, 5 of them were younger than 26 years and 3 participants were older than 35 years;
- Occupation: 8 of the participants were studying and 7 of them were working at the time of the survey;
- Privacy: 7 of the participants were pragmatists, 4 of them were fundamentalists and 4 participants were unconcerned.

C.6.3 Results

In the following section we present the results of the qualitative investigation in relation to the four identified drawbacks and the respective interdependencies.

Privacy drawbacks

Figure C.9 shows participants' preferences in relation to the features of the PAUSN design that are adopted in order to avoid the four privacy drawbacks.

All the participants agreed that ubiquitous social networking services must not require too much user intervention as well as they should disclose different user

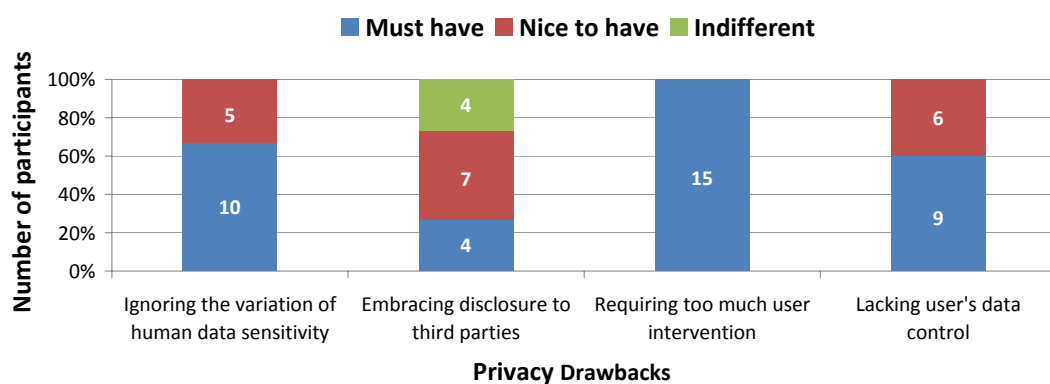


Figure C.9: Preferences on adopting design solutions for avoiding the privacy drawbacks

profiles according to the encountered circumstances. Furthermore, participants also acknowledged the importance of keeping control over their personal data, even after actual disclosure. Finally, they had contradictory opinions with respect to embracing disclosure to third parties. In the following we separately discuss the qualitative results.

Drawback 1 - Ignoring the variation of human data sensitivity In order to analyze the first drawback, we presented a scenario inspired by the BlueFriend mobile application, where users must always share a static profile for exploiting ubiquitous social networking services. Afterwards, we introduced another scenario, based on the PAUSN design, where users are able to disclose different profiles under different circumstances. As shown in Figure C.9, all the participants could foresee advantages when design of the services allows to share different profiles, by taking into consideration the variation of human data sensitivity for the current circumstances. The participants identified three major reasons for adopting

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

dynamic disclosure of personal information: (i) gaining more potential networking benefits, disclosing only (ii) relevant and (iii) not sensitive personal data, as one of the participants noted:

“If I had to choose a single profile to be shared, it would be very basic and limited. Thus, I would not benefit much from these services. I would prefer the dynamic sharing of different user profiles, because I would be disclosing only relevant information. For example, in a work environment, I don’t want to share information that is related to social activities and vice versa. This could as well jeopardize my privacy. In fact, if I would apply for a job that requires physical fitness, but at the same time I would disclose that I cannot join soccer matches because of a recent knee injury, such wrong information disclosure might be highly influential for my job application being rejected”

Drawback 2 - Embracing disclosure to third parties In order to analyze the second drawback, we firstly presented a scenario, inspired by Serendipity, which included disclosure of user personal information to a third party component. Afterwards, we introduced to the participants another scenario, based on the PAUSN design, which did not share personal information with third parties. As shown in Figure C.9, the majority of participants preferred not to disclose their personal information to the third parties. In fact, many of them provided a reference to the Facebook online social network and stated that they were intending to unsubscribe from this service in order to avoid disclosure of their personal information to this third party. Specifically, they were very worried that their personal data would

be given to other parties without their permission, which would result in negative future implications. One of them claimed:

“The main problem of third parties is that they gather an enormous amount of personal information, which can be used for different purposes. For example, one day I might get a marketing call because I disclosed my phone number and personal preferences on Facebook. This is an example of a minor invasion of privacy because I can always just hang up. However, it would already irritate me a lot. Who knows what other purposes this data might be useful for? I believe it could lead to a substantial invasion of privacy”

Moreover, all the participants that preferred this design property, agreed that they would share more personal information, e.g. phone number, in case it is only disclosed to end users. Contrarily, some of the participants, i.e. 4 out of 15 (who were notably categorized as privacy unconcerned), did not perceive any advantages of not disclosing their personal information to third parties:

“I do not think that there is any data that I would share with a stranger, but not to disclose to a third party. My information does not include any compromising data at the moment and I do not think that it will have negative implications in the future as well”

Drawback 3 - Requiring too much user intervention In order to analyze the third drawback, we firstly presented two user scenarios, based on the designs of Nokia Sensor and Bluedating mobile applications. We emphasized a manual profile

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

comparison in case of Nokia Sensor and lacking effortless establishment of connections between users in case of Bluedating, as already presented and discussed in Section C.3. Afterwards, we introduced another scenario, inspired by the design of PAUSN, which overcomes both limitations. As shown in Figure C.9, none of the participants would disclose their personal information in ubiquitous social networking services in case they required a considerable amount of user intervention:

“In order to work properly, this technology should require as little user intervention as possible. What about a potential mass adoption of ubiquitous social networking services? In this case, without an automatic profile comparison and possibilities to contact people also at later time, this technology would probably not be so useful and impossible to be adopted”

Drawback 4 - Lacking user’s personal data control In order to analyze the fourth drawback, we presented to the participants two different user scenarios. The first scenario relied on a decentralized architecture, such as Nokia Sensor, Bluedating and BlueFriend, and it did not empower users to control their data after actual disclosure. Contrarily, the second scenario was based on a centralized architecture, such as PAUSN, which enabled the individual participation principle and empowered the users, at any time, to add, revoke or modify any disclosed information. As shown in Figure C.9, all the participants could perceive significant advantages, when the design of ubiquitous social networking services avoids this drawback. In fact, 9 out of 15 participants would not disclose their personal data if they are not empowered to keep control over their data disclosure decisions:

“I would be very worried to share my information, in case I could not

keep control over it. I would not know how my information will be used and what will happen to it”

The majority of the participants indicated this design property to help them to feel more comfortable when exploiting ubiquitous social networking services:

“I am not sure how many times I will update my data disclosure decisions or revoke any of my sharing preferences, however it is of crucial importance for me to know that I can do it. Otherwise, I would not utilize these services”

Moreover, all the participants acknowledged that it would be also very useful to be empowered to modify their data disclosure decisions:

“Much of my personal information would change along time, such as phone number, address, career skills and abilities, etc. The opportunity to keep control over my data disclosure, not only enhances my privacy protection, but it also enables better social networking services as data is always up to date”.

Interdependencies between the drawbacks

At the end of the interview, we introduced to the participants the two interdependencies between the drawbacks, discussed in Section C.4.2, and presented how PAUSN design solution copes with these interdependencies. In relation to the interdependency between drawbacks 1 and 3, we presented to the participants a scenario, where users were required to intervene in case of (i) disagreement between the preferred and automated data disclosure decisions and (ii) needed approval for

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

disclosure of highly sensitive data. All the participants were very positive about these two options and they did not consider them as too much user intervention. Actually, many of them provided comments, similar to the following:

“I must have the opportunity to disagree with the the automated data disclosure. Without this option I would not use ubiquitous social networking, because I feel like losing control over my data”

The participants also agreed that they would not include in their profile highly sensitive data into their profiles, if the services would not ask for their approval before the actual disclosure of that highly sensitive data.

In relation to the interdependency between drawbacks 2 and 4, we presented to the participants a scenario, inspired by PAUSN, where the third-party component discovered relevant profiles similarities with other 4 users during a week. We emphasized that these similarities were found by comparing encrypted users’ profiles, which imply both (i) not disclosing personal information to the third-parties (avoiding Drawback 2) and (ii) keeping control over personal data, even after actual disclosure (avoiding Drawback 4). We discussed with participants that in case users would disclose not encrypted profiles to third-party entity, the number of discovered users with relevant profile similarities might increase to e.g. 9. As a result, all the participants, who expressed in the previous question that they preferred not to disclose personal information to third parties, presented comments, similar to the following:

“I would prefer to compromise the number of discovered users with relevant similarities, if the opposite would mean jeopardizing my privacy”

C.6.4 Investigation limitations

We acknowledged that conducting this qualitative investigation with a larger number of participants might lead to a more reliable analysis. Unfortunately, this was not possible due to limited resources. We attempted to address this issue by recruiting only lead users that presented relevant experience on the disclosure of personal information for gaining potential networking benefits in exchange, as discussed in Section C.6.2. This methodology was as well supported by Kujala and Kauppinen, who suggested to recruit lead users for the evaluation of new emerging technologies. In their research, the authors discovered that one lead user provided as much information and ideas as five *ordinary* users did [115].

Additionally, the data disclosure decisions, collected during our qualitative investigation, were based on predefined sharing preferences: participants were asked to predict their sharing preference in relation to the proposed user scenarios. We acknowledge that there might be a difference between what people say they want to share and what they actually do share in practice [100]. Thus, this analysis would have benefited from an implementation of the PAUSN application and a comparison with the other USN prototypes, introduced in Section C.3. Unfortunately, due to the limited number of participants, this approach was not feasible at the time of the investigation, because the reviewed prototypes and PAUSN are not disseminated enough so that it would be possible for users to find opportunities to disclose their personal information to other users in their everyday lives.

Finally, the presence of the interviewer might have biased the results of the qualitative investigation. Unfortunately, this is a well known problem, already

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

discussed in previous literature, and identified to be difficult to be avoided for such kind of investigations [163].

C.7 Conclusions

In this paper we focused on privacy as the main challenge for the development ubiquitous social networking services. We reviewed existing ubiquitous computing design guidelines that aim at providing means for the users to make informed data disclosure decisions. Afterwards, we described several ubiquitous social networking designs in terms of privacy protection. Notably, the majority of the design solutions were not following the reviewed privacy guidelines. Even when the platforms were providing means of informed data disclosure decisions, additional crucial usability and privacy limitations were identified. Therefore, we proposed four drawbacks to be taken into consideration when designing ubiquitous social networking. We refer specifically to (1) ignoring the variation of human data sensitivity, (2) embracing disclosure to third parties, (3) requiring too much user intervention and (4) lacking personal user's data control. These additional guidelines focus on creating more functional ubiquitous social computing environments, oriented to respecting the privacy of end users. Importantly, the drawbacks do not aim at ensuring total security, but instead we assumed a non-malicious infrastructure aiming at preventing incidental data disclosure, i.e. when personal data is unintentionally revealed, with or without previous inquiry.

Based on these findings, we proposed the design of a privacy-aware ubiquitous social networking (PAUSN) platform, which overcomes the identified drawbacks,

while allowing users to make informed data disclosure decisions. PAUSN focuses in particular on maximizing potential networking benefits while preserving users' personal privacy. To achieve this goal, PAUSN relies on a centralized architecture that enables disclosure of personal user information upon comparison of encrypted profiles. As a result, dynamic user sub-profiles, created by taking into account the human data sensitivity of the current circumstances, are effortlessly disclosed only to other end users, who hold a profile that might lead to potential networking benefits. Furthermore, thanks to its centralized architecture, PAUSN empowers users to modify the contents of their disclosed profiles at any time.

Finally, we evaluated the users' perceptions towards the protection of personal privacy in relation to the four identified drawbacks and their respective interdependencies. We presented to the 15 participants of our qualitative investigation two different scenarios of ubiquitous social networking services, which either avoid (based on the PAUSN design) or fall into (in case of the others applications) the specific drawback. All the participants agreed that the PAUSN design would allow more functional and privacy-oriented services and it also copes well with the interdependencies between the drawbacks. Moreover, the majority of them claimed that if the identified drawbacks were not avoided in the design of ubiquitous social networking, they would not disclose their personal information. However, among the four investigated drawbacks, we found contradictory results only in regard to embracing disclosure to third parties (Drawback 2). In fact, 4 out of 15 participants, who were actually classified as privacy unconcerned, would share their personal information even if it was disclosed to third parties.

In conclusion, we further encourage designers of ubiquitous social networking to

C. DESIGNING FOR PRIVACY IN UBIQUITOUS SOCIAL NETWORKING

take into account the four drawbacks, identified in this paper, additionally to the guidelines for enabling users to make informed data disclosure decisions, discussed in Section C.2. As future work, we suggest to complement our qualitative analysis with an additional investigation, based on users' sharing preferences made at the moment of the actual disclosure. This would comprise an implementation of the proposed privacy-aware ubiquitous social networking platform that should be adopted by a large number of users in order to collect users's ad hoc data disclosure decisions for further statistical analysis. Moreover, the proposed platform can be studied in a broader perspective in order to achieve higher levels of privacy not only on ubiquitous social networking, but on other Internet services, such as online social networks and cloud computing.

Acknowledgments

This work is partly supported by Nokia and developed as a part of the Converged Advanced Mobile Media Platforms (CAMMP) project²¹, funded by the Danish Advanced Technology Foundation (Højteknologifonden). The authors would like to thank Samant Khajuria for his valuable comments and suggestions.

²¹<http://www.cammp.dk>

Appendix D

Privacy analysis in mobile social networks: the influential factors for disclosure of personal data

Sapuppo Antonio Center for Communication, Media and Information Technologies - Aalborg University, Sydhavnsgade 17, Copenhagen 2450, Denmark - antonio@cmi.aau.dk

International Journal of Wireless and Mobile Computing, Vol. 5, No. 4, pp. 315-326, 2012.

Abstract

Nowadays, mobile social networks are capable of promoting social networking benefits during physical meetings, in order to leverage interpersonal affinities not only among acquaintances, but also between strangers. Due to their foundation on automated sharing of personal data in the physical surroundings of the user, these networks are subject to crucial privacy threats. Privacy management systems must be capable of accurate selection of data disclosure according to human data sensitivity evaluation. Therefore, it is crucial to research and comprehend individual's personal information disclosure decisions happening in ordinary human communication. Consequently, in this paper we provide insight into influential factors of human data disclosure decisions, by presenting and analyzing results of an empirical investigation comprising of two online surveys. We focus on the following influential factors: inquirer, purpose of disclosure, access & control of the disclosed information, location familiarity and current activity of the user. This research can serve as relevant input for the design of privacy management models in mobile social networks.

Keywords: Privacy; Information Disclosure; Mobile Computing; Social Networking; Social and Proximity Interactions; Ubiquitous Computing.

D.1 Introduction

The development of Internet reduced the distance between people living in different parts of the world by providing an innovative communication infrastructure. Soon on the basis of this technology new services have been developed, which improved the communication between people. Online social networks (in the following referred to as OSNs), such as Orkut, MySpace and Facebook, share a common characteristic: they enable people to create a virtual social network. By using OSNs services, users can stay in touch with friends from the whole world, share pictures, talk, chat, send messages and look for new acquaintances. The success of OSNs, the wide spread of mobile phones and the current development of numerous information and communication technologies allowed to create similar services also for mobile terminals [48, 206].

Notably, mobile devices are not just entry points to existing online social networks, but they also offer new networking services due to their advanced technological capabilities. In fact, thanks to the wireless technologies of mobile devices, they enable Opportunistic Networks (in the following referred to as ONs). In ONs, nodes are wirelessly interconnected and have the possibility to identify each other as well as share data in peer-to-peer networks with communication links created in ad hoc manner [128].

The integration of ONs with OSNs enables mobile social networks users to exploit social networking benefits in the physical world, rather than just in the virtual world. This integration has been previously introduced as Local Social Networks (in the following referred to as LSNs) [166, 171]. LSN is a distributed network ar-

D. PRIVACY ANALYSIS IN MOBILE SOCIAL NETWORKS: THE INFLUENTIAL FACTORS FOR DISCLOSURE OF PERSONAL DATA

chitecture in which nodes are linked to online social networks profiles and wirelessly interconnected to exchange personalized contents. LSNs target at developing possible advantageous relationships (e.g. friendships, partnerships, business relations) during physical meetings between people who do not know each other, but probably they should [65, 171].

When transferring OSNs benefits to the physical world, the privacy threats are indisputably increased due to support of face-to-face interactions between strangers during physical meetings. While the risk of unintentional information sharing is similar in virtual and physical worlds, the consequences of such disclosure are more crucial in the physical world. For example, when LSN users disclose their personal information, the shared data is tied to a physical person and immediately available for the recipient [171]. Thus, the information disclosure can be directly translated into physical contact and potentially undesired or unpleasant face-to-face interactions. In order to address these privacy concerns, privacy management systems should protect users' personal privacy as individuals do in ordinary human interactions and ensure accuracy of selective disclosure of personal information [25, 26, 99]. In fact, during face-to-face communication, people intuitively evaluate various determinants and unconsciously choose what personal information to share. In order to help privacy management systems to attempt to act as the real user would, it is necessary to gain an extensive comprehension of variation of human data sensitivity that affects information disclosure under different circumstances. The factors that might influence users personal data disclosure decisions must be depicted and evaluated for enabling privacy management systems to take automated data disclosure decisions.

In the past work, the identity of the inquirer was identified as the primary index for selection of data disclosure decisions [45, 57, 108, 123, 148]. However, mobile social networks, such as LSNs, advance the attention to other factors as crucial determinants for data disclosure, due to their primary focus on relationship initiation between strangers [171]. Consequently, in this paper we firstly identify the relevant influential factors that might impact users' personal data disclosure decisions in LSNs. Afterwards, we present results of an empirical investigation comprising 2 online surveys to evaluate the identified influential factors. We collected more than 100 responses in each of the surveys and we applied the Wilcoxon Signed Rank statistical test to examine whether the identified influential factors impact on users' personal data disclosure decisions. The results of our analysis can provide significant input for the design and development of privacy management systems for mobile social networks.

The rest of the paper is structured as follows: firstly, we introduce the potential influential factors for the disclosure of personal information in mobile social networks. In Section D.3, we present the design and methodology of the two surveys, which investigate the relevant influential factors in LSNs. Further, the information about the participants is provided in Section D.4. In Section D.5, we present and discuss the results of the empirical investigation. Final conclusions and recommendations for future work are drawn in Section D.6.

D.2 Human data disclosure

The core foundation of mobile social networks, such as LSNs, is based on automated sharing of users' personal data. Surely, the amount of disclosed information is directly proportional to networking benefits. The optimal outcome would be achieved by sharing as much as possible personal information (e.g. the full user profile). However, this would result in jeopardy of users' privacy and a compromise is necessary. It can be achieved by following the assumption that the sensitivity of the users' personal information is not stable; it may vary depending on different circumstances in which the user is involved [123, 171, 202]. Consequently, only information that is relevant, but not sensitive in specific circumstances should be disclosed at a time [25, 26, 111, 120, 204]. Therefore, no standard rules can be applied for all the cases of disclosure of users' personal data [4, 5, 149].

In previous studies [25, 103, 123, 171], the sensitivity of personal information was assumed to vary depending on the inquirer and the situation determinants. The inquirer is considered to be the individual that the user is interacting with and the situation is defined according to the circumstances at that time.

Lederer et al determined the identity of the inquirer to be the most important factor, influencing the users' data disclosure decisions, followed by the situation as parameter of secondary significance [123]. Based on these findings several privacy management models have been designed for disclosure of personal information: Faces [124], Precision Dial [122], Diverged Personalities [171] and Disclosure Decision Model [25, 26, 27].

In [57], the authors provided further insight into the inquirer influential fac-

tor by carrying out a survey to investigate the nature of relationships between the users as a crucial determinant. Their results showed that users differentiate choices of disclosure of personal information upon relationships with the inquirer. Additionally to Davis and Gutwin, other studies [108, 148] highlighted the relevance of users' clustering into a manageable categories of inquirers (e.g. friends, families, co-workers, etc) in social location disclosure applications.

Even if defining the inquirer as a crucial parameter, Consolvo et al emphasized that knowing the particular reason of data disclosure would significantly motivate users to share their personal information [45]. Other studies as well researched [29, 30] and applied [3, 20, 134, 185] the purpose of disclosure as a crucial determinant.

Additionally to the purpose of disclosure, Consolvo et al investigated the granularity of the disclosed information, which refers to the extent of details of shared data. The results showed that users tend not to differentiate granularity of disclosed information in order to protect their data privacy. In the majority of the cases, users either choose to disclose detailed information or they do not disclose anything at all. However, when they decide to disclose not detailed set of information, they do so because they assume that it is more useful for the inquirer, rather than for preserving their privacy [45].

Finally, anonymity can also be considered to be a relevant influential factor for sharing of personal information in mobile social networks. Being anonymous is defined as the state of not being identifiable within a set of subjects, due to removal of connections between the data owner and information. In [120], the author discussed that having the possibility to remain anonymous would significantly increase users' data privacy protection. Consequently, applications of anonymity might allow users

D. PRIVACY ANALYSIS IN MOBILE SOCIAL NETWORKS: THE INFLUENTIAL FACTORS FOR DISCLOSURE OF PERSONAL DATA

to feel safer and thus influence users' personal information decisions.

Additionally to the previously introduced influential factors, in this paper we draw the attention to other potential determinants that might impact human data disclosure decisions in mobile social networks, which are following defined:

- Location familiarity: it is considered to be the users' familiarity with his current location (e.g. home, parents' place, work environments, social environments, holiday environments, etc.);
- Current activity: it refers to the current action of the user (e.g. working, relaxing, shopping, etc.);
- Access & Control: it regards empowering users to add, remove or modify any information disclosed at any time, i.e. enabling to control other people's access to one's personal data even after the actual disclosure.

Importantly, access & control should be considered as an essential privacy protection principle for personal data disclosure [62, 82, 120]. This principle is of crucial importance for avoiding potential future privacy threats, because a set of data, given up freely today, might create major user's privacy concerns in the future. Moreover, mobile social networks are becoming increasingly complex, thus users might feel that they are losing control over their personal data after the actual disclosure.

To the best of our knowledge, access & control of the disclosed information, user's current activity and location familiarity influential factors were not previously empirically investigated in regard to the disclosure of personal information in mobile social networks. Moreover, we did not observe other research considering additional

influential factors for personal data disclosure in mobile social networks apart from the ones discussed in this section.

D.3 Design of the surveys

In order to gain insight into human data sensitivity, we asked surveys' participants to indicate personal information that they would like to share in different circumstances of their lives. The participants were informed that sharing of personal data is motivated by potential networking benefits, provided in return to disclosed information. Naturally, the benefits would be directly proportional to the amount of shared information, thus respondents were asked to compromise between privacy risks and potential benefits.

The different circumstances, presented to the respondents, were defined according to the influential factors, outlined in Section D.2. However, anonymity and granularity of disclosed information influential factors were not included in this analysis, as we focus on investigating information disclosure in LSNs. The granularity of the disclosed information is often applied in mobile social networks in relation to disclosure of social locations among acquaintances, e.g. extent of details of current location: country, city, neighborhood, exact address where I am now [8, 101, 178]. However, the main target of LSNs is to promote potential networking benefits between strangers by exploiting ONs. In ONs the disclosed information is restricted to the range of the wireless technology adopted. Particularly, users are notified about the presence of other LSN users only when they are in the proximity. When they move away, their location information is not available anymore,

D. PRIVACY ANALYSIS IN MOBILE SOCIAL NETWORKS: THE INFLUENTIAL FACTORS FOR DISCLOSURE OF PERSONAL DATA

unless they re-enter into each other's wireless range. Therefore, the granularity of the disclosed information was not further investigated in this paper. Moreover, anonymity influential factor was not included in this research because LSNs users must be identifiable, i.e. they must allow other users to link their profiles to real people. If anonymity would be applied in LSNs, it would result in significant losses of potential networking possibilities.

In the following we provide a detailed description of the design of the two surveys, which researched on the remaining influential factors, introduced in Section D.2.

D.3.1 Survey I

In the first survey we investigated whether the location familiarity and current activity of the user can be considered as relevant determinants for data disclosure decisions in mobile social networks. First, we researched whether the time that the user had previously spent in his current location could influence the amount of disclosed information. For example, we examined if the user would differentiate his data disclosure choices between places where has spent a lot of time (e.g. a bar of his home town) and unfamiliar locations (e.g. a bar during a holiday). Secondly, we analyzed whether the user's current activity might influence users data disclosure decisions. For example, while working the user might be more motivated to share data related to working activities (e.g. professional abilities) in comparison to data related to social interactions (e.g. music taste).

In order to study those influential factors, we grouped the most common life situations into five categories, and asked the participants to indicate, which information they would like to disclose when they are facing those situations:

- Family places: these environments can be considered to be places where the user or her family members live (e.g. parents' apartment, uncles' apartment, etc). Thus, it was assumed that users would encounter their family members as well as family members' acquaintances, who could also be strangers for them;
- Social environments: these environments refer to the places where the users spend their leisure time, e.g. restaurants, bars, theaters in his home city. Thus, it was assumed that they would encounter friends and strangers;
- Holiday: similarly to the social environments, holiday environments are considered to be social leisure places, however the users' encounters and activities are occurring outside their home city;
- Work environments: these environments can be considered to be the ordinary employment places of the users, such as university, office, etc. Thus, users would mainly encounter co-workers and strangers, associated to their employment activities;
- Work trip: similarly to work environments, during work trips the users were assumed to encounter colleagues and strangers, associated to their employment activities, however these encounters and activities were occurring outside their regular work place.

D.3.2 Survey II

In the second survey, we investigated whether inquirer, access & control and purpose of disclosure can be considered as relevant determinants for data sharing in

D. PRIVACY ANALYSIS IN MOBILE SOCIAL NETWORKS: THE INFLUENTIAL FACTORS FOR DISCLOSURE OF PERSONAL DATA

mobile social networks. In mobile social networks inquirers can be generally categorized into friends and strangers segments. In this investigation we target at the latter segment of inquirers, namely strangers, due to focus of LSNs. Consequently, we chose to investigate the following two concepts in the analysis of inquirer as influential factor. Firstly, we analyzed whether knowing the number of mutual friends between the inquirer and the user a priori any data sharing, might impact his data disclosure decisions. Further, we also researched whether being familiar strangers with the inquirer could be considered as a relevant determinant. Two people are identified as familiar strangers if they encounter each other regularly without interacting or forming an explicit relationship of social nature [138]. Moreover, we researched access & control as determinant factor by investigating whether clearly emphasizing access & control rights might influence the users' data disclosure decisions. Finally, we also analyzed whether users' disclosure decisions might be affected by knowing beforehand what potential benefits they could get for disclosing their personal information to strangers.

In order to research these influential factors, we asked respondents to select their personal information that they would like to disclose in different scenarios. It was emphasized that the exchange of personal data would be automated, thus it would not interfere with the user's current activity. The relevant information that could be applied for networking with other users could be retrieved and used even at a later time. All the scenarios, presented in this survey, were indicated to be occurring in a social environment. Particularly, respondents were asked to imagine to be in a bar of their home city, drinking a coffee with friends. The respondents decided what to disclose to different inquirers, who were strangers for them. A priori any data

disclosure decision, some information about the inquirer was known. Particularly, at least the basic information set about the inquirer, consisting of name, surname and portrait, was available in all the scenarios, which are following presented:

- Basic scenario: the respondents did not know so much about the inquirer. Particularly, only the basic information set was available a priori any data disclosure;
- Familiar strangers scenario: the respondents knew the basic information set and the number of previous encounters with the inquirer a priori any data disclosure. Notably, encountering does not necessarily imply interaction - they may have just passed by each other without noticing. Specifically, in this scenario the respondents had already encountered the inquirer 280 times;
- Mutual friends scenario: additionally to the basic information set, the respondents knew the number of mutual friends with the inquirer a priori any data disclosure. Specifically, in this scenario respondents had 15 mutual friends with the inquirer;
- Access & Control scenario: the respondents only knew the basic information set about the inquirer a priori any data disclosure. Moreover, it was explicitly emphasized that they can always edit/delete their disclosed personal information. Thus, respondents were empowered to control the inquirer's access to their disclosed personal data at any time in the future;
- Purpose scenario: additionally to the basic set of information, the respondents also knew other personal data regarding the inquirer a priori any data

D. PRIVACY ANALYSIS IN MOBILE SOCIAL NETWORKS: THE INFLUENTIAL FACTORS FOR DISCLOSURE OF PERSONAL DATA

disclosure. This data indicated that the inquirer was a project manager in a major company within the respondent's professional area. Moreover, the inquirer's professional targets were also available beforehand and particularly matching with the ones of the respondents.

D.4 Participants of the surveys

The two questionnaires were distributed to 500 potential respondents. The distribution of the questionnaires was limited to online social networks users. We determined this category to be the most relevant because of their advanced experience with personal data disclosure in online social networks sites, even if the perceptions of data disclosure might vary between virtual and physical worlds. Due to anonymity of the responses and different timeframes of the surveys, it cannot be ensured that the respondents of both surveys completely match, however a significant overlap is expected.

Respondents were asked to provide information about their demographics characteristics. We focused on three demographic features, namely gender, age and occupation, which were further applied for clustering purpose. Moreover, respondents were asked to indicate their privacy settings in their main OSN site, such as visibility of their user profile, pictures, posts to the other users. Based on these answers, we were able to observe patterns among data disclosure attitudes. Consequently, we classified the participants into three privacy clusters, following the Westin/Harris privacy segmentation model [198]:

- Fundamentalists: these respondents were extremely concerned about sharing

their personal data with any other online social networks users (friends or strangers);

- Pragmatists: they also cared about the disclosure of their personal information. However, they often had specific concerns and particular strategies for addressing them. For example, this category of respondents generally preferred sharing personal information only among their friends;
- Unconcerned: these respondents were trusting online social networks sites and believing that the privacy of their data was not jeopardized. Thus, they were willing to share their personal data not only with people who were their friends, but as well with users who were complete strangers to them.

In the following we present an overview of the demographic information as well as privacy clusters of the respondents in both surveys.

D.4.1 Respondents of the first survey

In total we received 121 complete answers for the first survey, which composed the sample. In the following we present the demographic characteristics of the respondents:

- Gender: 54.5% of the respondents were males and 45.5% were females;
- Age: 64.5% of the respondents were between 26 and 35 years old, 28.1% were younger than 26 years and 7.4% were older than 35 years;
- Occupation: 75.2% of the respondents were working and the 24.8% were studying at the time of the survey.

D. PRIVACY ANALYSIS IN MOBILE SOCIAL NETWORKS: THE INFLUENTIAL FACTORS FOR DISCLOSURE OF PERSONAL DATA

In regard to privacy clusters, the sample of the first survey was composed as follows:

- Fundamentalists: 10.7% of the respondents;
- Pragmatists: 74.4% of the respondents;
- Unconcerned: 14.9% of the respondents.

D.4.2 Respondents of the second survey

The sample of the second survey was composed of 101 answers. The demographic characteristics of the respondents are following presented:

- Gender: 67.3% of the respondents were males and 32.7% were females;
- Age: 57.4% of the respondents were between 26 and 35 years old, 33.7% were younger than 26 years and 8.9% were older than 35 years;
- Occupation: 58.4% of the respondents were working and the 41.6% were studying at the time of the survey.

Moreover, following we present the privacy clusters of the sample of the second survey:

- Fundamentalists: 17.8% of the respondents;
- Pragmatists: 64.4% of the respondents;
- Unconcerned: 17.8% of the respondents.

D.5 Survey results and discussion

In order to investigate the influential factors, defined in Section D.2, we relied on statistical characteristics and methods. First, we tested if the responses, grouped by different clusters, were normally distributed. We found out that many datasets of both surveys were not normally distributed. Consequently, we focused on analysis based on non-parametric statistical tests, due to expected higher precision of the results in comparison to the parametric tests [141]. Specifically, the Wilcoxon Signed Rank (in the following referred to as WSR) test was applied to examine the surveys' results by comparing two datasets and evaluate whether their population means differ [201]. When statistically comparing two samples, they are considered to be statistically different if the *p-value* is observed to be less than the critical significance level, commonly set to *0.05*. However, when analyzing more than 2 datasets, to evaluate our results we used the Bonferroni correction in order to avoid potential type I errors [195]. In these cases, the critical significance level is decreased to $0.05/n$, in which n is the total number of comparisons.

The following sections summarize the major results of our investigation. At the beginning we present and discuss results of the first survey followed by the ones of the second survey. Results are classified according to the privacy segmentation as well as the demographic characteristics, introduced in Section D.4.

D.5.1 Results of the survey I

In this section we investigate the impact of user's location familiarity and current activity influential factors for the disclosure of personal information in mobile social

D. PRIVACY ANALYSIS IN MOBILE SOCIAL NETWORKS: THE INFLUENTIAL FACTORS FOR DISCLOSURE OF PERSONAL DATA

networks.

Location familiarity

In order to evaluate the location familiarity influential factor, respondents were asked to choose which kind of personal information they would like to share in different locations, as described in Section D.3.1. The selection of personal data to be disclosed was limited to a dataset composed of 28 different types of personal information.

Table D.1 presents the standard deviations (σ), means (μ) and medians (\tilde{x}) of amounts of data shared in different user's locations. The mean and median results highlighted that respondents tend to share more personal information in familiar locations such as family places and work environments in comparison to less familiar places as work trip and holiday locations. This inclination can be explained by the fact that the users spend the majority of their time in these places and thus they develop an unconscious trust in more familiar environments.

Indisputably, work environments and work trip comprise similar conditions because in both circumstances the user is still in his professional environment. However, the user's familiarity with these locations is notably different and it motivates significantly lower data sharing preferences in work trip in comparison to work environments. Similar results were also observed when analyzing social environments and holiday locations, however with lower overall impact of the location familiarity influential factor.

On the contrary, locations that comprise different conditions (i.e. working and leisure), such as holiday and work trip, presented relevantly low differences between

D.5 Survey results and discussion

the amounts of shared data in all the clusters. This inclination can be explained by the fact that both locations can be considered to be unfamiliar to the user as he/she is outside of his/her ordinary environment.

Table D.1: Descriptive statistics of information disclosure in different locations

| | | | FP | WE | SE | H | WT |
|------------|-------|-------------|-------|-------|-------|-------|-------|
| Privacy | Fund. | σ | 5.64 | 3.26 | 6.47 | 5.43 | 5.66 |
| | | μ | 17.54 | 12.46 | 6.85 | 6.23 | 5.92 |
| | | \tilde{x} | 19 | 12 | 4 | 5 | 6 |
| | Prag. | σ | 7.42 | 5.88 | 6.50 | 6.34 | 6.20 |
| | | μ | 18.97 | 15.82 | 12.02 | 10.69 | 10.68 |
| | | \tilde{x} | 18.50 | 16 | 12 | 10.50 | 11 |
| | Unco. | σ | 4.19 | 4.16 | 6.07 | 5.34 | 5.58 |
| | | μ | 23.89 | 19.61 | 15.83 | 14.67 | 13.94 |
| | | \tilde{x} | 24.50 | 19 | 17.50 | 14.50 | 14.50 |
| Gender | Male | σ | 7.10 | 5.56 | 6.57 | 6.34 | 6.32 |
| | | μ | 19.83 | 16.48 | 13.23 | 11.89 | 11.38 |
| | | \tilde{x} | 20.50 | 16 | 14 | 12 | 11.50 |
| | Fema. | σ | 7.07 | 5.87 | 6.78 | 6.34 | 6.31 |
| | | μ | 19.20 | 15.47 | 10.60 | 9.49 | 9.78 |
| | | \tilde{x} | 21 | 15 | 10 | 8 | 10 |
| Age | < 26 | σ | 8.56 | 5.78 | 6.19 | 5.64 | 5.53 |
| | | μ | 18.24 | 14.88 | 12.12 | 11.67 | 10.85 |
| | | \tilde{x} | 19 | 14 | 12 | 11 | 10 |
| | 26-35 | σ | 6.41 | 5.55 | 6.67 | 6.35 | 6.11 |
| | | μ | 19.73 | 16.15 | 11.90 | 10.22 | 10.30 |
| | | \tilde{x} | 20 | 16 | 12 | 10 | 11 |
| | > 35 | σ | 6.10 | 6.07 | 9.99 | 9.39 | 10.40 |
| | | μ | 22.67 | 19.11 | 12.89 | 12.78 | 13 |
| | | \tilde{x} | 25 | 21 | 17 | 16 | 14 |
| Occupation | Stud. | σ | 7.51 | 5.91 | 6.31 | 5.80 | 6.50 |
| | | μ | 19.67 | 15.17 | 12.37 | 11 | 10.40 |
| | | \tilde{x} | 20 | 15 | 13 | 11.50 | 10 |
| | Empl. | σ | 6.96 | 5.64 | 6.94 | 6.65 | 6.32 |
| | | μ | 19.51 | 16.31 | 11.92 | 10.74 | 10.74 |
| | | \tilde{x} | 21 | 16 | 12 | 11 | 11 |

FP: Family Places; **WE:** Work Environments; **SE:** Social Environments; **H:** Holiday; **WT:** Work Trip.

D. PRIVACY ANALYSIS IN MOBILE SOCIAL NETWORKS: THE INFLUENTIAL FACTORS FOR DISCLOSURE OF PERSONAL DATA

Table D.2 shows results of statistical WSR test, comparing data disclosure in different locations. To account for multiple testing, we used the Bonferroni correction and considered significant only those p -values for which $P < 0.05/10 = 0.005$. As a result, we observed common statistically significant differences of data disclosure between all users' locations, except *Social Environments - Work Trip* and *Holiday - Work Trip*.

Similarly to results presented in Table D.1, all the clusters presented statistical differences between *Work Environments - Work Trip*, except of the respondents older than 35 years ($> 35.p = .007$). As well, many clusters also presented statistical differences between *Social Environments - Holiday* locations. Moreover, evidence towards equal amount of data sharing were observed in *Holiday - Work Trip*, as tests of statistical differences between those locations presented considerably high p -values.

Comparing the responses of different clusters, no relevant differences were observed between males and females, except in *Social environments - Holiday* ($Fema.p = .003$, $Male.p = .006$). Moreover, it can be noticed that pragmatists, employed as well as respondents between 26 and 35 years old were more affected by the familiarity of user' s location factor in comparison to the other relevant clusters. For example, data sharing results in *Work Environments - Social Environments* ($Pra.p = 26-35.p = Empl.p = .000$) and *Social environments - Holiday* ($Pra.p = 0.001$; $26-35.p = Empl.p = .000$) presented statistically significant differences in contrast to the other relevant clusters. Finally, no statistical differences were observed in any location comparison among the respondents older than 35 years.

Table D.2: Results of Wilcoxon Signed Rank test for information disclosure in different locations

| | FP-WE | FP-SE | FP-H | FP-WT | WE-SE | WE-H | WE-WT | SE-H | SE-WT | H-WT |
|---------|-------|-------|------|-------|-------|------|-------|------|-------|------|
| Privacy | Fund. | .004 | .001 | .001 | .009 | .003 | .002 | .324 | .301 | .611 |
| | Prag. | .000 | .000 | .000 | .000 | .000 | .000 | .001 | .050 | .603 |
| | Unco. | .002 | .000 | .001 | .000 | .032 | .003 | .000 | .082 | .507 |
| Gender | Male | .000 | .000 | .000 | .001 | .000 | .000 | .006 | .009 | .486 |
| | Fema. | .000 | .000 | .000 | .000 | .000 | .000 | .003 | .364 | .463 |
| Age | < 26 | .005 | .000 | .000 | .022 | .010 | .000 | .421 | .273 | .493 |
| | 26-35 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .012 | .554 |
| | > 35 | .051 | .008 | .008 | .011 | .008 | .007 | .483 | .999 | .892 |
| Occup. | Stud. | .000 | .000 | .000 | .016 | .000 | .000 | .108 | .119 | .699 |
| | Empl. | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .036 | .753 |

FP: Family Places; **WE:** Work Environments; **SE:** Social Environments; **H:** Holiday; **WT:** Work Trip.

D. PRIVACY ANALYSIS IN MOBILE SOCIAL NETWORKS: THE INFLUENTIAL FACTORS FOR DISCLOSURE OF PERSONAL DATA

Current activity

In order to evaluate the user's current activity influential factor, we focused on two different datasets: data related to work activities (DWA) and data related to social interactions (DSI). Both datasets were composed of nine different types of personal information, which were subsets of the full dataset of the first survey. For example, data related to working activities is employer, work phone number, career skills and abilities, while examples of data related to social interactions are relation status, food taste, interests, etc. In this analysis we compared these two datasets in their associated environments, i.e. social and work environments.

In Table D.3 we present the standard deviations (σ), means (μ) and medians (\tilde{x}) of the amounts of DWA and DSI, disclosed in both work and social environments. The results showed significantly different sharing preferences between the two analyzed datasets. In work environments, the mean and median values of DWA were considerably higher than DSI and notably close to the possible maximum amount of shared data, i.e. 9. Similar patterns were also observed in regard to social environments, in which DSI achieved higher sharing rate in comparison to DWA. However, the current activity influential factor presented a lower impact in social environments as the difference between sharing of DSI and DWA was considerably lower than the one in work environments.

These results were confirmed to be statistically significant by the WSR test. As shown in Table D.4-A, we firstly compared amounts of DWA and DSI shared in work environments and afterwards in social environments. In both circumstances, all the clusters presented statistically significant differences between DWA and DSI,

D.5 Survey results and discussion

Table D.3: Descriptive statistics of information disclosure during different users' activities

| | | | WE | | SE | |
|------------|-------|-------------|------|------|------|------|
| | | | DWA | DSI | DWA | DSI |
| Privacy | Fund. | σ | 1.33 | 1.30 | 1.94 | 2.73 |
| | | μ | 7.46 | 3.23 | 1.62 | 3.15 |
| | | \tilde{x} | 7 | 3 | 1 | 3 |
| | Prag. | σ | 2.00 | 2.33 | 2.18 | 2.48 |
| | | μ | 7.48 | 4.66 | 3.12 | 5.17 |
| | | \tilde{x} | 8 | 5 | 3 | 5.50 |
| | Unco. | σ | 0.75 | 1.98 | 2.34 | 2.24 |
| | | μ | 8.72 | 5.94 | 4.94 | 6.22 |
| | | \tilde{x} | 9 | 6 | 5.50 | 6.50 |
| Gender | Male | σ | 1.73 | 2.20 | 2.38 | 2.44 |
| | | μ | 7.61 | 5.03 | 3.67 | 5.45 |
| | | \tilde{x} | 8 | 5 | 3 | 6 |
| | Fema. | σ | 1.99 | 2.34 | 2.17 | 2.69 |
| | | μ | 7.73 | 4.29 | 2.71 | 4.69 |
| | | \tilde{x} | 9 | 5 | 3 | 4 |
| Age | < 26 | σ | 1.73 | 2.30 | 2.02 | 2.48 |
| | | μ | 7.48 | 4.36 | 3.30 | 5.09 |
| | | \tilde{x} | 8 | 4 | 3 | 5 |
| | 26-35 | σ | 1.96 | 2.25 | 2.20 | 2.55 |
| | | μ | 7.62 | 4.72 | 3.06 | 5.16 |
| | | \tilde{x} | 8 | 5 | 3 | 5 |
| | > 35 | σ | 0.71 | 2.45 | 3.91 | 3.32 |
| | | μ | 8.67 | 5.67 | 4.44 | 4.67 |
| | | \tilde{x} | 9 | 7 | 5 | 5 |
| Occupation | Stud. | σ | 1.78 | 2.29 | 2.10 | 2.27 |
| | | μ | 7.53 | 4.47 | 3.07 | 5.57 |
| | | \tilde{x} | 8 | 4 | 3 | 6 |
| | Empl. | σ | 1.88 | 2.29 | 2.40 | 2.66 |
| | | μ | 7.70 | 4.77 | 3.29 | 4.96 |
| | | \tilde{x} | 9 | 5 | 3 | 5 |

WE: Work Environments; **SE:** Social Environments;

DWA: Data related to Work Activities; **DSI:** Data related to Social Interactions.

D. PRIVACY ANALYSIS IN MOBILE SOCIAL NETWORKS: THE INFLUENTIAL FACTORS FOR DISCLOSURE OF PERSONAL DATA

Table D.4: Results of Wilcoxon Signed Rank test for information disclosure during different activities

| | | (A) DWA-DSI | | (B) WE-SE | |
|------------|-------|-------------|------|-----------|------|
| | | WE | SE | DWA | DSI |
| Privacy | Fund. | .001 | .024 | .001 | .559 |
| | Prag. | .000 | .000 | .000 | .101 |
| | Unco. | .001 | .050 | .000 | .647 |
| Gender | Male | .000 | .000 | .000 | .180 |
| | Fema. | .000 | .000 | .000 | .371 |
| Age | < 26 | .000 | .000 | .000 | .138 |
| | 26-35 | .000 | .000 | .000 | .200 |
| | > 35 | .007 | .748 | .011 | .104 |
| Occupation | Stud. | .000 | .000 | .000 | .013 |
| | Empl. | .000 | .000 | .000 | .536 |

WE: Work Environments; **SE:** Social Environments;
DWA: Data related to Work Activities; **DSI:** Data related to Social Interactions.

except respondents older than 35 years ($> 35.p = .748$) and the unconcerned privacy cluster ($Unco.p = .050$) in social environments. However, it must be noted that respondents older than 35 years presented a strong evidence of similarity, while unconcerned privacy clusters showed relevant differences, even if not statistically significant.

In order to complement the investigation of the current user's activity as crucial determinant, we also compared data disclosure between work and social environments by testing separately DWA and DSI, as shown in Table D.4-B. The results indisputably proved that the user's current activity factor had different impact on different data types, i.e. DWA and DSI. Notably, all the clusters differentiated DWA between work and social environments. However, the same tendency was not observed in regard to DSI. In fact, the only significant statistical difference was presented among the students ($Stud.p = .013$). No other significant differences among

the clusters were observed, despite an overall increasing influence of user' s current activity among the pragmatists privacy cluster and respondents older than 35 years ($Prag.p = .101$; $> 35.p = .104$). Consequently, results of Table D.4-B proved that differences between DWA and DSI in Table D.4-A were mainly caused by significant differentiation of DWA sharing preferences upon different activities.

D.5.2 Results of the survey II

In this section we investigate the impact of the following influential factors for the disclosure of personal information in mobile social networks: inquirer, access & control and purpose of disclosure.

Inquirer and Access & Control

In order to research on the inquirer influential factor we investigated the impact of being familiar strangers as well as having mutual friends with the inquirer. Afterwards, we analyzed the influence of explicit emphasis of access & control rights to the users. To evaluate the impact of these influential factors, we compared the responses in regard to the relevant scenarios, described in Section D.3.2. We focused on additional personal information, which was not shared in the basic scenario, however it was preferred to be disclosed in the familiar strangers, mutual friends or access & control scenarios.

Table D.5 presents the standard deviations (σ), means (μ) and medians (\tilde{x}) of the amounts of data shared in the analyzed scenarios. The median and mean results in familiar strangers, mutual friends and access & control scenarios presented higher data sharing preferences in comparison to the basic scenario, where no con-

D. PRIVACY ANALYSIS IN MOBILE SOCIAL NETWORKS: THE INFLUENTIAL FACTORS FOR DISCLOSURE OF PERSONAL DATA

Table D.5: Descriptive statistics of information disclosure in basic, familiar strangers mutual friends and access & control scenarios

| | | | Basic | FS | MF | AC |
|---------|-------|-------------|-------|-------|------|------|
| Privacy | Fund. | σ | 3.11 | 3.47 | 4.10 | 3.63 |
| | | μ | 2.17 | 3.17 | 4 | 3.50 |
| | | \tilde{x} | 0 | 3 | 3 | 4 |
| | Prag. | σ | 3.05 | 4 | 4.12 | 4.06 |
| | | μ | 4.72 | 6.69 | 7.08 | 7.22 |
| | | \tilde{x} | 5 | 7 | 7 | 7 |
| | Unco. | σ | 3.50 | 3.51 | 3.63 | 3.17 |
| | | μ | 7.17 | 10.11 | 9.89 | 9.56 |
| | | \tilde{x} | 6 | 10 | 10 | 9.50 |
| Gender | Male | σ | 3.52 | 4.26 | 4.53 | 4.45 |
| | | μ | 4.96 | 6.91 | 7.32 | 7 |
| | | \tilde{x} | 5 | 7 | 7 | 7 |
| | Fema. | σ | 3.32 | 4.48 | 4.03 | 3.82 |
| | | μ | 4.18 | 6.18 | 6.42 | 6.91 |
| | | \tilde{x} | 4 | 7 | 6 | 7 |
| Age | < 26 | σ | 3.43 | 4.66 | 4.26 | 4.48 |
| | | μ | 5.21 | 8.21 | 7.44 | 8.35 |
| | | \tilde{x} | 5 | 7.5 | 7 | 9 |
| | 26-35 | σ | 3.41 | 3.98 | 4.46 | 3.88 |
| | | μ | 4.17 | 5.67 | 6.67 | 6 |
| | | \tilde{x} | 3.50 | 6 | 7 | 6 |
| | > 35 | σ | 3.49 | 3.67 | 4.47 | 4.27 |
| | | μ | 6.22 | 7.33 | 7.78 | 8 |
| | | \tilde{x} | 6 | 8 | 9 | 7 |
| Occup. | Stud. | σ | 3.75 | 4.49 | 4.48 | 4.79 |
| | | μ | 5.05 | 7.43 | 7.29 | 7.74 |
| | | \tilde{x} | 5 | 7.50 | 7 | 8 |
| | Empl. | σ | 3.24 | 4.15 | 4.32 | 3.73 |
| | | μ | 4.46 | 6.14 | 6.85 | 6.42 |
| | | \tilde{x} | 5 | 6 | 7 | 7 |

FS: Familiar Strangers; **MF:** Mutual Friends; **AC:** Access & Control

nections with the inquirer were highlighted and no access & control rights were emphasized. Moreover, it was observed that unconcerned privacy cluster and respondents younger than 26 years were the most impacted by the influential factors.

D.5 Survey results and discussion

In fact, they presented higher mean and median differences between the basic and other scenarios in comparison to the other clusters.

Table D.6 presents WSR test results obtained by comparing the responses associated to the four scenarios. To account for multiple testing between different scenarios, we used the Bonferroni correction and considered significant only those p-values for which $P < 0.05/6 = 0.008$. In regard to the inquirer influential factor, all the clusters generally presented significant differences between sharing of personal information in the basic and mutual friends/familiar strangers scenarios. Particularly, statistically significant differences were not observed only among fundamentalists privacy cluster in *Basic - Familiar Strangers* ($Fund.p = .011$) as well as respondents older than 35 years in *Basic - Familiar Strangers* ($> 35.p = .039$) and *Basic - Mutual Friends* ($> 35.p = .026$). Moreover, when comparing mutual friends and familiar strangers scenarios, significant differences were observed only among working respondents ($Empl.p = .003$) and respondents between 26 and 35

Table D.6: Results of Wilcoxon Signed Rank test for information disclosure in basic, familiar strangers, mutual friends and access & control scenarios

| | | B-FS | B-MF | B-AC | FS-MF | FS-AC | MF-AC |
|------------|-------|------|------|------|-------|-------|-------|
| Privacy | Fund. | .011 | .003 | .026 | .107 | .905 | .402 |
| | Prag. | .000 | .000 | .000 | .064 | .092 | .655 |
| | Unco. | .001 | .001 | .001 | .908 | .302 | .430 |
| Gender | Male | .000 | .000 | .000 | .104 | .935 | .234 |
| | Fema. | .000 | .000 | .000 | .138 | .121 | .343 |
| Age | < 26 | .000 | .000 | .000 | .292 | .913 | .057 |
| | 26-35 | .000 | .000 | .000 | .000 | .322 | .031 |
| | > 35 | .039 | .026 | .026 | .279 | .245 | .786 |
| Occupation | Stud. | .000 | .000 | .000 | .903 | .680 | .228 |
| | Empl. | .000 | .000 | .000 | .003 | .301 | .188 |

B: Basic; **FS:** Familiar Strangers; **MF:** Mutual Friends; **AC:** Access & Control

D. PRIVACY ANALYSIS IN MOBILE SOCIAL NETWORKS: THE INFLUENTIAL FACTORS FOR DISCLOSURE OF PERSONAL DATA

years old ($26-35.p = .000$). No other significant differences were presented, despite an overall increase of the difference between *Familiar Strangers - Mutual Friends* among the pragmatists privacy cluster ($Prag.p = .064$).

Similarly to the previous results, *Basic - Access & Control* also presented statistically significant evidence of differences among the clusters, except of respondents older than 35 years and fundamentalists ($> 35.p = Fund.p = .026$). Moreover, no statistically significant differences were observed when comparing access & control with familiar strangers/mutual friends scenarios, despite an overall increase of differences in *Familiar Strangers - Access & Control* among pragmatists and females ($Prag.p = .092; Fema.p = .121$) and in *Mutual Friends - Access & Control* among respondents younger than 35 years ($< 26.p = .057; 26-35 = .031$).

Comparing the privacy and demographic clusters, the fundamentalists were the only privacy cluster not affected by familiar strangers and access & control factors. We did not observe any differences between clusters within gender and occupation segments. Moreover, respondents older than 35 years were not influenced by any factor presented in this section in contrast to the other age groups.

Finally, these results confirmed that when access & control rights were clearly emphasized, the respondents were motivated to disclose more personal information as they might feel not to lose control over their personal data, even after actual disclosure.

Purpose of disclosure

In order to evaluate the impact of the purpose of disclosure influential factor, we compared data sharing preferences in basic and purpose scenarios. As described

D.5 Survey results and discussion

in Section D.3.2, the purpose scenario focused on potential professional networking benefits, even if it occurred in a social environment. This test was limited to personal information related to work activities. Specifically, the chosen dataset was composed of 7 different types of user's personal information, which was a subset of

Table D.7: Descriptive statistics of information disclosure in basic and purpose scenarios

| | | | Basic | Purpose |
|------------|-------|-------------|-------|---------|
| Privacy | Fund. | σ | 1.98 | 1.94 |
| | | μ | 1.56 | 5.33 |
| | | \tilde{x} | 0.50 | 6 |
| | Prag. | σ | 1.78 | 1.81 |
| | | μ | 2.69 | 4.98 |
| | | \tilde{x} | 3 | 5 |
| | Unco. | σ | 2.03 | 2.40 |
| | | μ | 3.67 | 5 |
| | | \tilde{x} | 3 | 5.50 |
| Gender | Male | σ | 1.96 | 2.03 |
| | | μ | 2.71 | 5.01 |
| | | \tilde{x} | 3 | 5.50 |
| | Fema. | σ | 1.99 | 1.73 |
| | | μ | 2.48 | 5.12 |
| | | \tilde{x} | 3 | 5 |
| Age | < 26 | σ | 1.90 | 2.30 |
| | | μ | 2.71 | 4.85 |
| | | \tilde{x} | 3 | 6 |
| | 26-35 | σ | 1.95 | 1.71 |
| | | μ | 2.43 | 5.14 |
| | | \tilde{x} | 2 | 5 |
| | > 35 | σ | 2.18 | 1.92 |
| | | μ | 3.67 | 5.22 |
| | | \tilde{x} | 4 | 5 |
| Occupation | Stud. | σ | 2.07 | 2.14 |
| | | μ | 2.95 | 4.55 |
| | | \tilde{x} | 3 | 5 |
| | Empl. | σ | 1.87 | 1.69 |
| | | μ | 2.41 | 5.41 |
| | | \tilde{x} | 2 | 6 |

D. PRIVACY ANALYSIS IN MOBILE SOCIAL NETWORKS: THE INFLUENTIAL FACTORS FOR DISCLOSURE OF PERSONAL DATA

the full dataset of the second survey. Examples of data related to work activities are employer, career skills and abilities, education details, etc.

Table D.7 presents the standard deviations (σ), means (μ) and medians (\tilde{x}) of amounts of information shared in basic and purpose scenarios. In the purpose scenario, the values of mean and median were considerably higher than in basic scenario and very close to the maximum possible amount of shared data, i.e. 7. Comparing the responses of different demographic and privacy clusters, we did not observe important mean and median differences within the age and gender segments, however fundamentalists and employed respondents were the most impacted by the purpose of disclosure in comparison to the other relevant clusters.

Table D.8 illustrates the WSR test results, obtained by comparing the responses of the basic and purpose scenarios. It can be noticed that all the clusters presented statistically significant differences between the amounts of data, shared in basic and purpose scenarios.

Finally, results of Table D.7 and Table D.8 proved that all the users were willing

Table D.8: Results of Wilcoxon Signed Rank test for information disclosure in basic and purpose scenarios

| | | Basic-Purpose |
|------------|-------|---------------|
| Privacy | Fund. | .001 |
| | Prag. | .000 |
| | Unco. | .037 |
| Gender | Male | .000 |
| | Fema. | .000 |
| Age | < 26 | .000 |
| | 26-35 | .000 |
| | > 35 | .014 |
| Occupation | Stud. | .000 |
| | Empl. | .000 |

to share more personal information, when they had reasons to disclose their data, e.g. they had the possibility to predict potential professional networking benefits.

D.6 Conclusions

In this paper we provided insight into personal data sensitivity in order to contribute to design of privacy management systems for mobile social networks. Firstly, we outlined the relevant influential factors that might impact users' personal information disclosure decisions. Afterwards, we empirically investigated the most relevant determinants for data disclosure in mobile social networks, which promote networking not only among acquaintances, but also between strangers with interpersonal affinities in the physical world.

According to our analysis, the purpose of data disclosure was found to be the most determinant factor for privacy preferences among the ones tested, as it is statistically proven to be affecting all the respondent clusters. Moreover, emphasizing access & control rights was proven to help users to feel more secure to share their personal information. Thus, we strongly encourage privacy designers to take into account purpose of data disclosure factor as primary index into users' privacy preferences as well as apply and clearly emphasize access & control rights.

Further, following the results of our research we also suggest designers of privacy systems to consider the other influential factors, however as indexes of secondary importance. Particularly, the location familiarity factor was commonly approved by all the respondents who presented tendency to be more open to share their personal information in more familiar locations. Moreover, our analysis proved that knowing

D. PRIVACY ANALYSIS IN MOBILE SOCIAL NETWORKS: THE INFLUENTIAL FACTORS FOR DISCLOSURE OF PERSONAL DATA

beforehand information about the inquirer, such as number of mutual friends or previous encounters, relevantly impacted the information disclosure decisions. Finally, the investigation of the activity factor presented different impact in relation to different data types. This factor was observed to be significantly influential only on disclosure of data related to work activities. In fact, respondents did not significantly differentiate sharing of data, related to social interactions, between work and social environments.

In regard to the demographic and privacy clusters, our analysis did not show relevant differences between data sharing among male and female clusters. The pragmatists privacy cluster and respondents between 26 and 35 years old were overall the most affected by the influential factors, in comparison to the other relevant clusters. Furthermore, the fundamentalist privacy cluster was the most influenced by the purpose of disclosure determinant, even if they were generally only slightly impacted by the other factors. Finally, respondents older than 35 years did not present impact of any influential factors, except of the purpose of disclosure.

The results of our research strongly encourage further research on the influential factors, discussed in Section D.2, for the disclosure of personal information in mobile social networks. Particularly, a qualitative investigation would be a relevant supplement to the results of the quantitative research, presented in this paper. Moreover, the current mood of the user might also be considered as a determinant for data disclosure. This potential influential factor is suggested to be taken into account during qualitative investigations in the future work.

Acknowledgments

This work is supported by Nokia and developed as a part of the Converged Advanced Mobile Media Platforms (CAMMP) project²², funded by the Danish Advanced Technology Foundation. The author would like to acknowledge Boon-Chong Seet and Emil Heinze for their important suggestions during the design of the two surveys and Egle Juzokaite for the contribution to the statistical data analysis. Finally, the author thanks Lene Sørensen, Reza Tadayoni for their valuable comments on the paper.

²²<http://www.cammp.dk>

**D. PRIVACY ANALYSIS IN MOBILE SOCIAL NETWORKS: THE
INFLUENTIAL FACTORS FOR DISCLOSURE OF PERSONAL
DATA**

Appendix E

The influential factors for the variation of data sensitivity in ubiquitous social networking

Sapuppo Antonio Center for Communication, Media and Information Technologies - Aalborg University, Sydhavnsgade 17, Copenhagen 2450, Denmark - antonio@cmi.aau.dk

International Journal of Wireless and Mobile Computing, Vol. 6, No. 2, pp. 115-130, 2013.

Abstract

Ubiquitous social networking services offer new opportunities for developing advantageous relationships by uncovering hidden connections that people share with others nearby. As sharing of personal information is an intrinsic part of ubiquitous social networking, these services are subject to crucial privacy threats. In order to contribute to the design of privacy management systems for ubiquitous social networking, we present results of a mixed methods study that investigated the influential factors for the variation of human data sensitivity upon different circumstances. The results indicate that the users' information sensitivity is decreasing inversely proportionally to the relevance of data disclosure for initiation of relationships with others. We suggest privacy designers to take into account the purpose of disclosure and environment influential factors as primary indexes for data disclosure, because they were found to be the most determinant for evaluation of data relevance for networking. Other influential factors, i.e. activity, mood, location familiarity, number of previous encounters and mutual friends, were as well found to influence participants' data disclosure, but as factors of secondary importance.

Keywords: Privacy; Ubiquitous Computing; Information Disclosure; Social Networking.

E.1 Introduction

The development of computing originated with many people serving one computer and has gradually evolved into the currently existing possibility for many computers to serve one person anywhere around the world [192]. The latest wave of computing, called ubiquitous computing, shifts the central focus of users' attention away from the computers by embedding many seamless highly specialized devices within people's surroundings [191]. These devices are aware of their current environments and users and, consequently, they are able to improve humans' lives and support their everyday tasks [193].

For ubiquitous computing applications to intelligently and naturally support humans who are by nature social beings, it is essential to embody social intelligence, which can be defined as the ability of the environment to acquire and apply users' social context [61, 180, 184]. This led to the development of ubiquitous social computing, where a social dimension has been introduced in order to increase awareness, knowledge and intelligence of ubiquitous computing environments [205].

The establishment of ubiquitous social computing allows the possibility to transfer online social networking benefits to the physical world, by promoting ubiquitous social networking (in the following referred to as USN) services. These services target at developing possible advantageous relationships such as friendships, partnerships, business relations by uncovering hidden connections that people share with others nearby and thus facilitating initialization of face-to-face interactions between people who do not know each other, but probably should. As a result, the value of social networking is significantly enhanced and benefits are available

E. THE INFLUENTIAL FACTORS FOR THE VARIATION OF DATA SENSITIVITY IN UBIQUITOUS SOCIAL NETWORKING

immediately upon demand [65, 83, 155, 183].

When transferring online social networking benefits to the physical world, the privacy threats are indisputably increased due to support of face-to-face interactions between strangers during physical meetings [168]. While the risk of unintentional information sharing is similar in virtual and physical worlds, the consequences of such disclosure are more crucial in the physical environments. For example, when USN users disclose their personal information, the shared data is tied to a physical person and immediately available for the recipient [171]. Thus, the information disclosure can be directly translated into physical contact and potentially undesired or unpleasant face-to-face interactions [168]. To address these privacy concerns, privacy management systems should protect users' personal privacy as individuals do in ordinary human interactions [25, 26, 99]. In fact, during face-to-face communication, people intuitively evaluate various determinants and unconsciously choose what personal information to share. In order to help privacy management systems to attempt to act as the real user would and ensure accuracy of selective disclosure of personal information, it is necessary to gain an extensive comprehension of variation of human data sensitivity that affects information disclosure under different circumstances.

In previous studies, the identity of the inquirer was identified as the primary index for selection of data disclosure decisions in ubiquitous computing [45, 57, 108, 123, 148]. On the other hand, USN services advance the attention to other factors as crucial determinants for data disclosure, due to their primary focus on initiation of relationships between strangers [170, 171]. Several influential factors that impact personal data disclosure in USN, such as users' current activities and location famil-

ilarity, were identified during an empirical investigation, based on predefined data disclosure preferences [167]. This investigation provided statistically significant results, obtained by asking participants to predict their sharing preferences a priori the actual data disclosure. However, there might be a difference between what people say they want to share and what they actually do share in practice [13, 100]. Furthermore, data disclosure decisions taken at the moment of actual disclosure in ubiquitous social computing environments were found to be more accurate in comparison to predefined privacy preferences, as users might encounter circumstances where data disclosure decisions are not precisely predictable [25, 26, 27, 103, 122]. Consequently, it is important to investigate the previously identified influential factors for variation of human data sensitivity, by analyzing these factors based on in situ data disclosure privacy preferences, as well as gain an extensive understanding of people' attitudes and motivations that govern such data disclosure decisions in USN.

In order to achieve these goals, we applied a sequential two-phase mixed methods study for analysis of ad hoc data disclosure preferences in USN with active online social networks users. In the first phase, a quantitative research investigated the relationship between the identified influential factors and ad hoc data disclosure decisions. By exploiting a USN prototype, we collected participants' ad hoc data disclosure decisions and applied the binary logistic regression statistical model for examining whether the selected influential factors can be considered as predictors for users' data disclosure decisions in USN. Information, acquired during the first phase of the study, was explored further in the second phase, where qualitative interviews were used to gain in-depth understanding of different aspects and motivations of

E. THE INFLUENTIAL FACTORS FOR THE VARIATION OF DATA SENSITIVITY IN UBIQUITOUS SOCIAL NETWORKING

users' data disclosure in USN. The results of our mixed methods analysis can provide significant input for the design and development of privacy management systems for USN environments.

The rest of the paper is structured as follows: firstly, we discuss related work regarding influential factors for the disclosure of personal information. In Section E.3, we list and define each of the influential factors that might impact on users' data sharing preferences in ubiquitous social networking. In Section E.4, we present the design and methodology of the mixed methods study. Further, the information about the participants of our investigation is provided in Section E.5. In Section E.6, we present the major findings of the analysis. Final conclusions and recommendations for future work are drawn in Section E.7.

E.2 Related Work

In past works, different studies have questioned whether the sensitivity of personal data remains unchanged upon different circumstances. In [123], the authors found the sensitivity to vary depending on the inquirer and the situation determinants. The inquirer is considered to be the individual that the user is interacting with and the situation is defined according to the circumstances at that time. Lederer et al determined the identity of the inquirer to be the most important factor, influencing the users' data disclosure decisions, followed by the situation as parameter of secondary significance. Other studies provided further insight into the inquirer influential factor by emphasizing that users differentiate choices of disclosure of personal information upon relationships with the inquirer. In fact, they indicated that

self-reported closeness was a crucial factor for deciding whether to disclose their personal information to a specific inquirer [57, 199]. Moreover, other research also highlighted the need to cluster users into manageable categories of inquirers (e.g. friends, family members, co-workers, etc) for taking users' data disclosure decisions, in order to better preserve their data privacy [108, 148].

Even if defining the identity of the inquirer as a crucial parameter, in [45] the authors investigated other factors that might impact users' personal data disclosure decisions. Firstly, they analyzed the granularity of the disclosed information, which refers to the extent of details of shared data. The results showed that users tend not to differentiate granularity of disclosed information in order to protect their data privacy. In the majority of the cases, users either choose to disclose detailed information or they do not disclose anything at all. However, when they decide to disclose not detailed set of information, they assume that it is more useful for the inquirer, rather than for preserving their privacy. Secondly, the authors also indicated users' current mood and activities as relevant factors for personal data disclosure. The former implies that users differentiate their data disclosure upon their humor, e.g. participants were most willing to disclose their personal data when "depressed", in contrast to being "angry". In regard to users' current activity, Consolvo et al discussed that during some activities (e.g. exercising) users were more inclined to share their personal information rather than others, such as studying. Thirdly, the authors indicated that knowing the particular reason for data disclosure would also significantly motivate users to share their personal information [45]. The purpose of disclosure was also researched in other studies, which attempted to ensure that users' personal data is processed for only the intended reason [29, 30, 154, 185].

E. THE INFLUENTIAL FACTORS FOR THE VARIATION OF DATA SENSITIVITY IN UBIQUITOUS SOCIAL NETWORKING

Another relevant factor that might impact users' personal data disclosure decisions is anonymity. Being anonymous is defined as the state of not being identifiable within a set of subjects, due to removal of connections between the data owner and information. Having the possibility to remain anonymous would significantly increase data privacy protection and consequently might influence users' personal information disclosure decisions [120]. However, in case of necessity for the users' authentication, the pseudoanonymity approach could be applied. Pseudoanonymity is realized through linking the users to IDs, which represent them in specific circumstances and allow them to be recognized as long as they use the same ID [14, 120].

All the aforementioned studies, even if acknowledging the existence of other relevant influential factors, commonly indicated the identity of the inquirer as the most crucial determinant for data disclosure in ubiquitous computing environments. On the contrary, in [167] the author advanced the attention to analysis of other influential factors for data disclosure, due to specific focus on USN and consequent initialization of relationships between strangers. The findings strongly encourage privacy designers of USN to take into account the purpose of data disclosure factor as the primary index for decisions about data disclosure to strangers. Moreover, it strongly recommends to consider the access & control influential factor, which implies the right for users to be able to influence other people's access to one's personal data, even after the actual disclosure. Further, it also suggests to consider other influential factors for the disclosure of personal information, however as indexes of secondary importance, i.e. familiarity with the current location, current activity and other information about inquirer, such as the number of previous encounters and mutual friends [167]. The mutual friends influential factor was also confirmed

to significantly impact personal data disclosure decisions in online social networks, where users were proven to be much more likely to disclose sensitive information to strangers if they have a friend in common [145].

The mixed methods study, presented in this paper, is based on the findings of the past work, discussed in this section. Especially, the results presented in this article target at complementing the outcomes of the empirical analysis in USN, described in [167]. However, differently from [167], in this investigation we focus on analyzing ad hoc data disclosure decisions, which were proven to be more accurate for preserving users' data privacy in USN [25, 26, 27, 103]. Moreover, we quantitatively analyze additional influential factors that were not taken into consideration in the previous empirical analysis [167]. We refer to research on the current mood and different aspects of the users' current location: type of current environment (e.g. work, social, holiday) and location familiarity, evaluated according to the amount of time that the users usually spend in a specific location (e.g. daily, monthly, first time in this location, etc). Finally, in this mixed methods study, we also supplemented empirical results with findings of qualitative interviews in order to further explore the participants' attitudes towards the impact of influential factors on their data disclosure decisions.

E.3 The influential factors

In this section we present the influential factors that might impact users' data disclosure decisions in USN. Importantly, some of the factors, introduced in Section E.2, were not taken into consideration. We refer to anonymity, granularity of disclosed

E. THE INFLUENTIAL FACTORS FOR THE VARIATION OF DATA SENSITIVITY IN UBIQUITOUS SOCIAL NETWORKING

information and identity of the inquirer. Even if acknowledging the importance of these influential factors for users' overall data disclosure, we did not find them to be relevant in USN for the reasons described in the following.

Firstly, the granularity of disclosed information is often applied in disclosure of social locations among acquaintances and refers to extent of details of current location, e.g. country, city or neighborhood to the exact address where the user is [8, 101, 178]. However, USN services commonly exploit opportunistic networks to promote social networking between strangers during physical meetings [65, 168, 183]. In opportunistic networks, the data exchange is restricted to the range of the adopted wireless technology and thus users are aware about each others' location only when they are in the proximity [168]. Consequently, the granularity of the disclosed information was not considered as a relevant factor for data disclosure in USN. Secondly, the anonymity and pseudoanonymity influential factors were not also included in this research because it would cause significant losses of potential networking opportunities. In USN users must be identifiable, as they must allow others to link their profiles to *real* people in order to have the possibility to gain USN benefits. Thirdly, we did not consider the identity of the inquirer to be a relevant influential factor in USN, due to its focus on initiation of relationships between strangers. Instead, we took into account other information that the users have in common with the inquirer, such as number of mutual friends and previous encounters.

The selected influential factors for data disclosure in USN were clustered into three different groups: contextual information, interrelated attributes and design properties. The first group regards influential factors related to the current con-

textual circumstances of the users' encounters in USN environments, e.g. where is the user, what he is doing, etc. The second group of influential factors consists of information that the user has in common with the inquirer, e.g. similar music preferences or number of mutual friends. Finally, the last group corresponds to design solutions that should be taken into consideration when implementing USN services. A definition of each of the selected influential factors is following provided.

Contextual data:

- **Environment:** is considered to be the current location of the users, grouped according to their ordinary activities in that location, e.g. work environments, social environments, work trips, etc.
- **Location familiarity:** is considered to be the users' familiarity with their current location evaluated according to the amount of time that users usually spend in a specific location, e.g. daily, monthly, first time in this location, etc.
- **Activity:** is considered to be the current action of the user, e.g. working, relaxing, etc.
- **Mood:** is considered to be the users' current status of emotion, e.g. depressed, happy, sad, angry, etc.

Interrelated attributes:

- **Familiar strangers:** is considered to be the number of times that the users have already encountered the inquirer, e.g. 120 times in the last 3 weeks, etc.

E. THE INFLUENTIAL FACTORS FOR THE VARIATION OF DATA SENSITIVITY IN UBIQUITOUS SOCIAL NETWORKING

Notably, encountering does not necessarily imply interaction - they may have just passed by each other without noticing.

- Mutual friends: is considered to be the number of mutual friends that the users have with the inquirer, e.g. 6 common friends, etc.
- Purpose of disclosure: is considered to be the reason why a specific personal information is disclosed, e.g. potential networking benefits are foreseen because users have related interests or career abilities and expectations.

Design properties:

- Access and control: is considered to be the right to add, remove or modify any information disclosed at any time. This design solution enables users to control other people's access to their personal data even after actual disclosure.

Importantly, even if acknowledging the importance of access & control as a relevant influential factor for data disclosure in USN, this design property will not be further discussed in this paper. We preferred focusing on contextual data and interrelated attributes influential factors in our mixed method study, because of our target to in-depth analyze the variation of human data sensitivity under different circumstances. The design properties are a complex problem worth investigating, however they were found to influence participants' data disclosure, due to perceived increase of comfort with data disclosure and better usability of USN services, rather than shaping data sensitivity under different circumstances [167, 168].

E.4 Investigation methodology and design

In this section we present the methodology and design of our investigation that aims at analyzing the influential factors, introduced in Section E.3. In order to ensure the validity of answers, we helped participants to get more familiar with the USN concept during an introductory meeting, at beginning of the study. We introduced to the participants the existing USN prototype Spiderweb [166] as well as its services (presented also in this video²³) and other USN applications, already available in the market, i.e. Sonar²⁴ and Aka-Aki²⁵. We also illustrated how potential networking benefits can be gained through USN, as shown in this video about Aka-Aki²⁶. Further, we presented different scenarios from everyday lives, where these services might be applied, such as professional areas, dating and big events, e.g. conferences and exhibitions, as described in [65, 168]. Finally, we discussed with the participants the potential networking benefits in the identified application areas as well as possible privacy threats that might arise as a result of the information disclosure in USN, e.g. potential undesired face-to-face interactions [168].

Afterwards, the participants engaged in a mixed methods study, composed of quantitative and qualitative investigations. We preferred to run a mixed methods study, in comparison to carrying out only one of the two selected investigations, because this approach allows to gain a broad understanding of the research problem as well as ensures greater overall validity of results [51]. As illustrated in Figure E.1, this study followed the sequential explanatory strategy characterized by col-

²³<http://www.youtube.com/watch?v=DgeVNv10CIM>

²⁴<http://www.sonar.me>

²⁵<http://www.aka-aki.com/>

²⁶<http://www.youtube.com/watch?v=mvRgtT4LawU>

E. THE INFLUENTIAL FACTORS FOR THE VARIATION OF DATA SENSITIVITY IN UBIQUITOUS SOCIAL NETWORKING

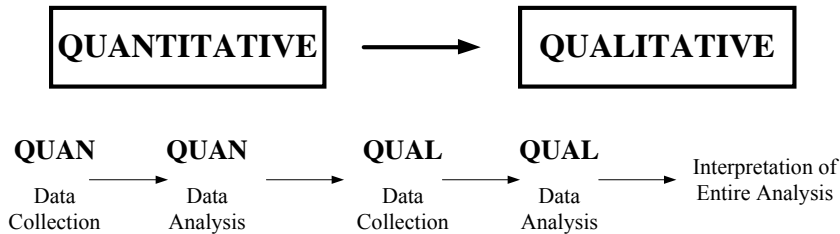


Figure E.1: Sequential explanatory design of the mixed methods study

lection and analysis of quantitative data in the first phase of research, which then provides input for the subsequent qualitative investigation. We selected a sequential explanatory strategy in comparison to others, e.g. concurrent triangulation or transformative designs, because of its straightforward nature that enables to gain in-depth understanding of obtained findings and especially pay particular attention to unexpected results, arising from the quantitative study [142]. In the following we explain in details the two phases of the study.

E.4.1 Phase 1: Quantitative investigation

The first phase of the study comprises a quantitative investigation that analyzes the statistical relationships between the selected influential factors and participants' in situ data disclosure decisions. In the following, firstly, we introduce the techniques utilized for collection of data about users' personal data disclosure decisions and afterwards we describe the statistical methods chosen for analyzing the acquired information.

Data collection

In order to collect data about ad hoc information disclosure decisions, participants were asked to utilize a mobile application that simulates the USN behavior. We preferred to provide a new mobile application, designed specifically for this investigation, rather than utilizing the Spiderweb mobile social network or other existing USN applications, due to two reasons that are explained in the following. First, these applications are not widely spread yet and participants would probably encounter difficulties in finding opportunities to disclose their personal information to other *real* users. Second, the provided USN prototype was explicitly designed to collect data about participants' information disclosure decisions for further analysis, based on the selected influential factors.

Several times a day, the USN prototype was randomly asking participants to specify their current circumstances and their related ad hoc data disclosure decisions. Three screenshots of the USN prototype are shown in Figure E.2. Firstly, participants were inquired to specify their current circumstances, as illustrated by an example in Figure E.2-A:

- Which is your current mood? The participants could choose between the following range of answers: happy, angry, excited, stressed, sad, worried or other, which implied unrestricted description of their current mood;
- Where are you? The participants could choose between the following range of answers: work environment, social environment, holiday, work trip, on the move or other, which implied unrestricted description of their current location;
- How often are you usually here? The participants could choose between the

E. THE INFLUENTIAL FACTORS FOR THE VARIATION OF DATA SENSITIVITY IN UBIQUITOUS SOCIAL NETWORKING

following range of answers: daily, weekly, monthly, few times per year, first time here;

- What are you doing? The participants could choose between the following range of answers: working, partying, relaxing, shopping, socializing, walking or other, which implied unrestricted description of their current activity.

After the participants provided information about the current circumstances, the USN prototype was asking them to express their ad hoc data disclosure preferences, as shown by an example in Figure E.2-B. Participants were aware that potential networking benefits would be directly proportional to the amount of shared information, thus their ad hoc data disclosure decisions were representing a compromise between privacy risks and potential networking benefits. The selection of data types to be disclosed was provided in accordance to data categorization in popular online

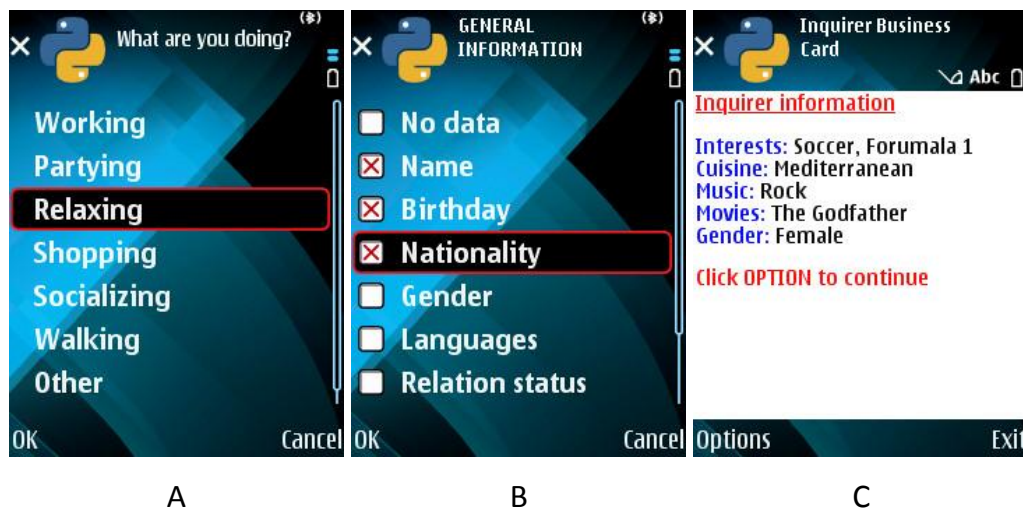


Figure E.2: Three screenshots of mobile prototype simulating the ubiquitous social networking behavior

social networks sites (e.g. gender, age and favorite music). This categorization was already used in previous investigations about disclosure of personal information in USN and the detailed description of the provided data types can be found in [170].

When participants had expressed their ad hoc data disclosure preferences according to the current circumstances, the USN prototype presented them a business card of a hypothetical inquirer, composed of some interrelated attributes between the participants and the inquirer, as illustrated by an example in Figure E.2-C. Three different kinds of attributes were randomly selected by the application:

1. Participants were informed about the number of times they had encountered the hypothetical inquirer during the last 3 months under similar circumstances. This number was randomly ranging from 2 to 900.
2. Participants were informed about a number of mutual friends with the hypothetical inquirer. This number was randomly ranging from 2 to 80.
3. Participants were informed about personal information of the hypothetical inquirer, randomly related either to work or social activities. The presented inquirer's personal information was purposely matching the one of the participants, e.g. shared tastes in music, movies, food or career skills, abilities and expectations. Notably, we were capable of finding these interrelated attributes between the participants and hypothetical inquirers, because we had previously collected participants' personal information about their work and social activities and preferences, during the introductory meeting at the beginning of the study.

Finally, after highlighting the interrelated attributes with the hypothetical in-

E. THE INFLUENTIAL FACTORS FOR THE VARIATION OF DATA SENSITIVITY IN UBIQUITOUS SOCIAL NETWORKING

quirer, the USN prototype asked the participants whether they would like to extend their ad hoc data disclosure decisions with any other personal information, which was previously preferred to be kept private. Both initial and extended ad hoc data disclosure decisions were stored in the local memory of the provided mobile phones together with the respective circumstances in order to be applied for further statistical analysis.

Data analysis

The collected data disclosure decisions were analyzed by applying the logistic regression method. We selected this approach because it does not require strict assumptions as other statistical methods like ordinary least squares regression or linear discriminant function analysis [151]. In contrast to the other two mentioned methods, the logistic regression does not assume linearity between independent and dependent variables nor normality or equal variance within each group of the independent variables [28, 71, 158, 182]. Moreover, we decided to run a binary logistic regression, instead of other kinds of logistic regression methods, such as multiple or ordinary, because our dependent variable was dichotomous, i.e. either disclose the information or not, and the categorical typology of our independent variables, e.g. environment, activity, mood, data type. The research hypothesis posed to the data was that the likelihood of a USN user to disclose specific personal information is dependent on the investigated influential factors. Thus, the variables were defined as follows:

- Dependent variable: whether specific users' personal information, e.g. music taste, is disclosed (1 = yes, 0 = no);

E.4 Investigation methodology and design

- Independent variables (or predictors): data type (e.g. name, music taste or career skills) and selected influential factors introduced in section E.3, e.g. environment, current activities, etc.

To test the research hypothesis, a six-predictor logistic model was applied for data collected from each participant. We preferred to consider each of the participants as a separate test case, rather than evaluating all the participants' data disclosure decisions collectively. This choice was motivated by our focus on ensuring user's personal privacy, i.e. the process where an individual selectively shares his/her own personal information, such as email address, career skills and abilities, to others [122]. In order to separately analyze participants' data disclosure decisions, we aimed at collecting enough data sharing preferences to run the logistic regression statistical method per each of the participants. Consequently, we asked participants to utilize the USN prototype for 11 days. At least three times per day, participants provided their initial and extended data disclosure decisions that were composed of 25 data types each, as introduced in Section E.4.1. The total number of collected data disclosure decisions per each of the participants during the time of the test was the following:

$$11 \text{ (days)} * 3 \text{ (times per day)} * 2 \text{ (initial and extended data disclosure decisions)} * 25 \text{ (data types)} = 1650,$$

which considerably exceeded the minimum recommended sample size, e.g. at least 50 cases per predictor, as suggested in [28, 151].

E. THE INFLUENTIAL FACTORS FOR THE VARIATION OF DATA SENSITIVITY IN UBIQUITOUS SOCIAL NETWORKING

The logistic regression analysis was carried out in SPSS version 19 in the Windows 7 environment and addressed:

- Overall evaluation of the model: we present results of the chi-square statistic in order to evaluate the overall significance of the model. When the significance level of the chi-square statistic is lower than *.050*, we can determine that the overall model is statistically significant. In such cases, there is not a high probability to obtain the presented chi-square statistic value under the condition that the data types and influential factors, taken together, do not have impact on users' data disclosure decisions;
- Goodness-of-fit-statistics: we present results of the Hosmer-Lemeshow (H-L) test, which assesses how accurately the model's estimates fit to the actual data. The accuracy is considered to be acceptable when the H-L significance is greater than *.050*. Additionally to the H-L statistic, we also present results of the Nagelkerke index (R^2) that investigates the strength of relationship between the dependent and independent variables. This index ranges from 0 to 1, with the value 1 representing the strongest relationship between the variables;
- Assessment of predicted probabilities: we present information about the overall proportion of cases that the model classified correctly. In the ideal model, this proportion would amount to 100%;
- Statistical tests of individual predictors: we show results of the Wald chi square statistic, which provides an index of the significance of each analyzed independent variable. The predictors are considered to be relevant when the

corresponding significance value is less than $.050$. In this case, it is possible to determine that the analyzed independent variable has a significant impact on the users' data disclosure decisions.

E.4.2 Phase 2: Qualitative investigation

The second phase of the study comprises a qualitative investigation that was conducted to better understand the impact of the influential factors on participants' personal data disclosure decisions as well as research on subjective motivations causing the quantitative results. In the following, we introduce the techniques utilized for collection of qualitative data, followed by the strategy for data analysis.

Data collection

Qualitative interviews were preferred alternatively to other investigation methods, such as handing out questionnaires or establishing a focus group interview. This method was chosen because of the following two reasons: (i) lack of participants' extensive experience in utilizing USN services and (ii) potential misinterpretation of the research questions due to their complexity and ambiguity. Moreover, we decided to run semi-structured interviews to better understand the motivation behind the participants responses and ensure that general areas of information are collected from each participant, however still allowing adaptability of the interview process [50, 118, 136].

Questions were related to the selected influential factors and respective statistical results, obtained during the first phase of the study. Specifically, per each influential factor, we asked the participants to reflect on how important each of the

E. THE INFLUENTIAL FACTORS FOR THE VARIATION OF DATA SENSITIVITY IN UBIQUITOUS SOCIAL NETWORKING

factors was for their personal data disclosure decisions and to elaborate on the reasons. Moreover, after showing the statistical results of the quantitative investigation to the participants, we asked them whether they could confirm these results and comment on any surprising outcomes, obtained in the quantitative investigation.

Data analysis

The strategy utilized for analysis of the information, collected during the qualitative interviews, follows a hierarchical approach, illustrated in Figure E.3. At beginning we transcribed the qualitative interviews (step 1) and reviewed them in order to gain a generic understanding of the participants' attitudes towards influential factors that

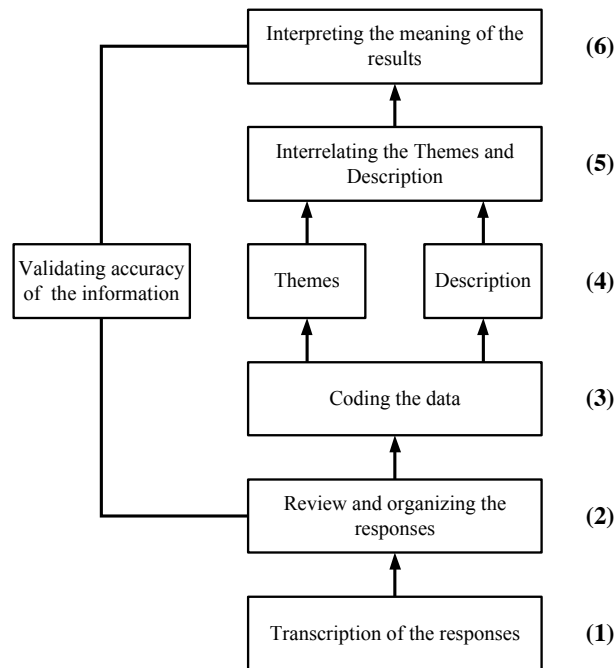


Figure E.3: Qualitative data analysis of the mixed methods study

E.4 Investigation methodology and design

might impact their data disclosure decisions in USN (step 2). In step 3, we organized the transcribed answers in different segments based on the influential factors and, in the next step, we generated the description of participants and themes. The former regards information about the participants (e.g. gender, privacy clusters), while the latter refers to the categories of the major research findings. In step 5, we interrelated the themes and description data categories and, in the last step, we interpreted the qualitative data while taking into account, when relevant, the interconnection between themes and description.

In order to ensure accuracy of the findings, two different techniques were applied during the analysis of the qualitative data, as shown in Figure E.3. The first one was the triangulation of different data sources, carried out during the second step of our analysis. Particularly, we provided to participants a questionnaire for classifying the influential factors according to the impact that they had on their personal data disclosure decisions. Consequently, we were capable of understanding whether our first review of the transcripts resulted in correct assumptions. Afterwards, during the last step of our analysis, we applied the second technique, i.e. member checking, which determined the interpretations' accuracy of the collected qualitative responses. Specifically, we sent out the final findings of the qualitative investigation back to the participants in order to get feedback on the accuracy of interpretation. When needed, follow-up interviews with the participants were conducted to give them the opportunity to additionally comment on the findings.

E.5 Participants

The participants were randomly selected by sending out email invitations to take part in this study. The selection was limited to online social networks users. We determined this category to be the most relevant because of their advanced experience in social networks, even if the perception towards the services might vary between virtual and physical worlds.

Respondents were asked to provide information about their demographic characteristics. Particularly, we focused on three demographic features, namely gender, age and occupation, which were further applied for clustering purpose. Participants were also asked to indicate their privacy preferences on visibility of their own personal data (e.g. user profile, pictures, posts) in their main OSN site. Based on these answers, we were able to observe patterns among data disclosure attitudes. Consequently, we also classified the participants into three privacy clusters, following the Westin/Harris privacy segmentation model [198]:

- **Fundamentalists:** these respondents were extremely concerned about sharing their personal data with any other online social networks users (friends or strangers);
- **Pragmatists:** these participants also cared about loss of privacy due to the disclosure of their personal information. However, they often had specific concerns and particular strategies for addressing them. For example, this category of respondents generally preferred sharing personal information only among their friends;

- Unconcerned: these respondents were trusting online social networks sites and believing that the privacy of their data was not jeopardized. Thus, they were willing to share their personal data not only with people who were their friends, but as well with users who were complete strangers to them.

When recruiting the participants, we aimed to achieve stratification between participants' privacy clusters to ensure that specific characteristics of individuals are represented in the sample in accordance to the proportion in the entire population [75]. Consequently, in this study, we target at obtaining similar proportions of participants' privacy clusters in reference to our latest empirical investigation where a random sample was selected [167]. In total we recruited 13 participants with the following privacy and demographic characteristics:

- Gender: 8 of the participants were male, while 5 of them were females;
- Age: 6 of the respondents were between 26 and 35 years old, 5 of them were younger than 26 years and 2 participants were older than 35 years;
- Occupation: 7 of the participants were studying and 6 of them were working at the time of the investigation;
- Privacy: 7 of the respondents were pragmatists, 3 of them were fundamentalists and 3 of the participants were unconcerned.

The detailed demographic and privacy characteristics of each participant are illustrated in Table E.1.

E. THE INFLUENTIAL FACTORS FOR THE VARIATION OF DATA SENSITIVITY IN UBIQUITOUS SOCIAL NETWORKING

Table E.1: Information about the participants

| User | Gender | Age | Occupation | Privacy |
|------|--------|-------|------------|----------------|
| 1 | Male | > 35 | Employed | Pragmatic |
| 2 | Male | 26-35 | Employed | Unconcerned |
| 3 | Male | 26-35 | Employed | Fundamentalist |
| 4 | Male | 26-35 | Employed | Pragmatic |
| 5 | Female | 26-35 | Employed | Fundamentalist |
| 6 | Male | 26-35 | Student | Fundamentalist |
| 7 | Female | < 26 | Student | Pragmatic |
| 8 | Male | > 35 | Employed | Unconcerned |
| 9 | Female | 26-35 | Student | Pragmatic |
| 10 | Female | < 26 | Student | Unconcerned |
| 11 | Male | < 26 | Student | Pragmatic |
| 12 | Male | < 26 | Student | Pragmatic |
| 13 | Female | < 26 | Student | Pragmatic |

E.6 Investigation results

In this section we present the results of our mixed methods study that investigates whether users' personal data disclosure decisions in USN are impacted by the influential factors, defined in Section E.3. In the following, we present the quantitative results followed by the outcomes, obtained during the qualitative investigation.

E.6.1 Quantitative results

Table E.2 presents the results of the binary logistic analysis that was conducted to predict participants' data disclosure decisions using the data type of the disclosed information and the selected influential factors as predictors. Importantly, we did not consider all the 13 participants sharing preferences as a collective sample, instead we run the model per each of the participants as a separate test case, due to

Table E.2: Results of the binary logistic regression analysis

| User ID | Prediction correct | Model's overall evaluation | | | Goodness-of-fit | | Individual predictors | | | | | |
|---------|--------------------|----------------------------|-----|------|-----------------|----------------|-----------------------|------|-------------|-------------|-------------|------|
| | | Chi-square | Dif | Sig. | H-L Sig. | R ² | Type | Env | Mood | Fam | Act | Int |
| 1 | 93.4% | 895.310 | 44 | .000 | .977 | .886 | .000 | .000 | .000 | .002 | .060 | .000 |
| 2 | 88.4% | 1268.333 | 43 | .000 | .000 | .663 | .000 | .000 | .800 | .004 | .000 | .000 |
| 3 | 92.1% | 589.700 | 47 | .000 | .978 | .563 | .000 | .002 | .269 | .145 | .000 | .000 |
| 4 | 92.6% | 1380.669 | 43 | .000 | .115 | .827 | .000 | .000 | .000 | .538 | .000 | .000 |
| 5 | 90.4% | 1169.623 | 47 | .000 | .075 | .764 | .000 | .004 | .262 | .832 | .004 | .000 |
| 6 | 87.5% | 1202.384 | 48 | .000 | .200 | .702 | .000 | .001 | .000 | .267 | .000 | .000 |
| 7 | 84.6% | 1093.781 | 48 | .000 | .079 | .652 | .000 | .013 | .000 | .238 | .000 | .000 |
| 8 | 91.8% | 1123.657 | 39 | .000 | .673 | .808 | .000 | .031 | .297 | .765 | .010 | .000 |
| 9 | 93.9% | 991.477 | 44 | .000 | .377 | .785 | .000 | .049 | .000 | .236 | .122 | .000 |
| 10 | 85.3% | 1339.721 | 47 | .000 | .113 | .744 | .000 | .000 | .000 | .017 | .000 | .000 |
| 11 | 88.1% | 1413.007 | 47 | .000 | .695 | .767 | .000 | .003 | .108 | .033 | .074 | .000 |
| 12 | 92.1% | 1533.357 | 41 | .000 | .840 | .837 | .000 | .008 | .163 | .007 | .000 | .000 |
| 13 | 88.1% | 1185.852 | 44 | .000 | .386 | .772 | .000 | .000 | .795 | .000 | .201 | .000 |

Type: Data type; **Env:** Environment; **Fam:** Location familiarity; **Act:** Activity; **Int:** Interrelated attributes.

E. THE INFLUENTIAL FACTORS FOR THE VARIATION OF DATA SENSITIVITY IN UBIQUITOUS SOCIAL NETWORKING

focus on personal privacy (refer to Section E.4.1 for more details).

The overall evaluation of the model was found to be statistically significant for each of the participants, as the significance value of the Chi-square statistics was always observed to be lower than $.000$. Further, in the majority of the cases the model was also found to accurately fit to the actual data, because the H-L significance value was observed to be more than $.050$ in 12 out of 13 cases. As well, the Nagerlkerke R^2 index was observed to be higher than $.750$ for the majority of the participants, which indicated moderately strong (75% or more) relationships between predicted outcomes and predictors. The overall success of predictions ranged from 84.6% to 93.9% with a mean value of 90%.

In order to analyze whether the selected influential factors could be considered as relevant predictors for users' data disclosure in USN, we present results of the Wald chi square statistic applied to 6 predictors: data type, environment, location familiarity, mood, activity and interrelated attributes. The data type was found to be a statistically significant predictor, because the *p-values* for all participants were observed to be less than $.000$, as shown in Table E.2. In fact, the information disclosure decisions were strongly influenced by the kind of shared data, e.g. participants might have decided to disclose music tastes, but not home address. Moreover, the other investigated predictors, related to the contextual data and interrelated attributes, are discussed in the following subsections.

Contextual data influential factors

In this section we describe the Wald statistic results in regard to the contextual data influential factors, i.e. environment, location familiarity, activity and mood.

As shown in Table E.2, the Wald criterion demonstrated that the environment influential factor was found to be a statistically significant predictor for participants' data disclosure. The *p-values* for all the participants were observed to be less than .050. Notably, contradicting results were discovered in regard to the other three contextual data influential factors, i.e. activity, location familiarity and mood. The current activity was found to be statistically significant for 9 out of 13 participants, while the mood and location familiarity were both found to be statistically significant for 6 out of 13 participants. In Table E.3, we present the impact of activity, location familiarity and mood influential factors on different participants' clusters.

In regard to the privacy clusters, the most relevant results can be observed among the fundamentalists where none of the participants was found to be influenced by the location familiarity, while all of them together with the unconcerned participants were strongly affected by the activity influential factor. No relevant differences were noted among the gender clusters. In relation to the age groups,

Table E.3: Impact of mood, location familiarity and activity influential factors on different clusters

| | | N | Mood | LF | Act |
|------------|-------|---|------|----------|----------|
| Privacy | Fund. | 3 | 1 | 0 | 3 |
| | Prag. | 7 | 4 | 4 | 3 |
| | Unco. | 3 | 1 | 2 | 3 |
| Gender | Male | 8 | 3 | 4 | 6 |
| | Fema. | 5 | 3 | 2 | 3 |
| Age | < 26 | 5 | 2 | 4 | 3 |
| | 26-35 | 6 | 3 | 1 | 5 |
| | > 35 | 2 | 1 | 1 | 1 |
| Occupation | Stud. | 7 | 4 | 4 | 4 |
| | Empl. | 6 | 2 | 2 | 5 |

N: Total number of participants; **LF:** Location Familiarity; **Act:** Activity.

E. THE INFLUENTIAL FACTORS FOR THE VARIATION OF DATA SENSITIVITY IN UBIQUITOUS SOCIAL NETWORKING

4 out of 5 participants younger than 26 years were found to be impacted by the location familiarity in contrast to other clusters, while participants from 26 to 35 years old presented significant results in regard to the activity influential factor. Finally, among the occupation clusters, the employed participants were observed to be highly impacted only by the activity influential factor.

Interrelated attributes influential factors

As shown in Table E.2, the Wald criterion demonstrated that the interrelated attributes predictor was found to be statistically significant for personal data disclosure in USN, as the *p-values* for all the participants were observed to be less than *.000*. Consequently, in this section, we get insight into each of the interrelated attributes influential factors by separately analyzing whether informing the USN users about varying numbers of mutual friends and previous encounters as well as diverse profiles similarities might differently influence participants' data disclosure preferences.

In order to achieve this goal, we selected the initial ad hoc data disclosure decisions, when no information about the inquirer was provided, as baseline reference category for the binary logistic regression model. Afterwards, the baseline reference category was compared to the respective extended ad hoc data disclosure decisions, categorized according to various scenarios of the three interrelated attributes influential factors, described in the following.

For familiar strangers, firstly, we took into consideration answers about data disclosure, when few previous meetings, ranging from 2 to 9, were indicated. Afterwards, as second scenario, we analyzed only answers based on encounters, ranging

E.6 Investigation results

from 50 to 100. As a third scenario we only considered data disclosure decisions, where previous meetings with the inquirer were indicated to range from 700 to 900. Three different scenarios were also applied for the mutual friends influential factor, with the first scenario ranging from 2 to 4 mutual friends, second scenario ranging from 10 to 23 and the last scenario ranging from 50 to 80 common friends with the inquirer. In regard to the purpose of disclosure, firstly, we considered answers that were only related to social profile similarities and afterwards we evaluated only disclosure preferences, based on work profile similarities. The individual results of the Wald statistics for familiar strangers, mutual friends and purpose of disclosure factors are presented in Table E.4 and Figure E.4. In Table E.4, we show the significance values for each scenario, while Figure E.4 presents the $Exp(B)$ mean values, which indicate the average change in probability of disclosing personal data, caused by providing information about the inquirer.

Table E.4: Wald statistic significance values for different scenarios of the interrelated attributes influential factors

| User ID | Familiar Strangers | | | Mutual Friends | | | Purpose of Disclosure | |
|---------|--------------------|-------------|-------------|----------------|-------------|-------------|-----------------------|-------------|
| | (2,9) | (50,100) | (700,900) | (2,4) | (10,23) | (50,80) | Social | Work |
| 1 | .029 | .009 | .000 | .063 | .028 | .201 | .005 | .646 |
| 2 | .439 | .006 | .023 | .000 | .052 | .001 | .000 | .000 |
| 3 | .044 | .000 | .315 | .322 | .001 | .001 | .748 | .001 |
| 4 | .675 | .000 | .000 | .001 | .000 | .000 | .000 | .000 |
| 5 | .315 | .024 | .000 | .022 | .000 | .000 | .016 | .000 |
| 6 | .006 | .000 | .000 | .144 | .000 | .008 | .008 | .000 |
| 7 | .228 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| 8 | .312 | .013 | .000 | .023 | .637 | .099 | .009 | .000 |
| 9 | .188 | .000 | .000 | .134 | .000 | .000 | .000 | .000 |
| 10 | .016 | .065 | .000 | .103 | .165 | .003 | .028 | .236 |
| 11 | .000 | .000 | .000 | .001 | .000 | .000 | .000 | .000 |
| 12 | .090 | .130 | .000 | .004 | .954 | .001 | .004 | .174 |
| 13 | .097 | .000 | .000 | .017 | .033 | .079 | .162 | .000 |

E. THE INFLUENTIAL FACTORS FOR THE VARIATION OF DATA SENSITIVITY IN UBIQUITOUS SOCIAL NETWORKING

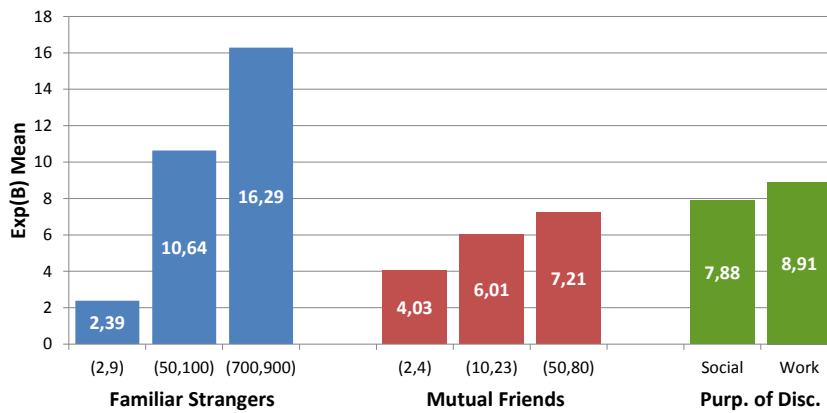


Figure E.4: Change of probability of data disclosure under different scenarios

As shown in Table E.4, for familiar strangers, the majority of participants was not influenced by this predictor when few meetings with the inquirer were known before any actual data disclosure, i.e 2-9 previous encounters. When this number was increased, many more participants were impacted by being familiar strangers with the inquirer. In fact, in the second scenario, 11 out 13 participants were influenced and, in the last scenario, only one participant was not impacted by this factor. These outcomes were also confirmed in the results presented in Figure E.4, where the probability of disclosing personal information significantly increased proportionally to the number of previous meetings with the inquirers. Participants were approximately 16 times more likely to disclose their personal information when being aware about a large number of previous encounters, i.e. between 700 and 900. In regard to the mutual friends influential factor, we did not observe a significant difference between the scenarios, despite the varying numbers of mutual friends. As shown in Table E.4, 8 out 13 participants were influenced by having 2-4 common friends with the inquirers, while in the last scenario (i.e 50-80 common friends) only

2 additional participants were found to be impacted by this influential factor. As well, in Figure E.4, we can still observe a relevant, but not significant, increase of probability to disclose personal information, proportional to the number of mutual friends. Finally, the purpose of disclosure was found to be a very strong predictor, as we observed that all the participants were influenced by this determinant in at least one of the two scenarios, i.e. either work or social. As shown in Table E.4, 10 out of 13 participants were found to be influenced by this factor in the work scenario and 11 out of 13 of them were observed to be impacted by knowing beforehand to have social similarities with the inquirer. Both scenarios also presented relevant changes of probability to disclose personal information with slightly higher results when relevant professional networking benefits could be foreseen, as illustrated in Figure E.4.

E.6.2 Qualitative results

In this section we present the outcomes of the second phase of the study. Firstly, we describe the results of the qualitative investigation about the contextual data influential factors, i.e. mood, environment, location familiarity and activity, followed by the ones regarding the interrelated attributes influential factors, i.e. familiar strangers, mutual friends and purpose of disclosure.

Contextual data influential factors

During the qualitative interviews, initially, we inquired the participants about the environment influential factor, as it was found to be statistically significant for all them during the quantitative investigation. As shown in Figure E.5, all the partici-

E. THE INFLUENTIAL FACTORS FOR THE VARIATION OF DATA SENSITIVITY IN UBIQUITOUS SOCIAL NETWORKING

participants confirmed the importance of the current environment, because it significantly shaped the relevancy of certain data types to be disclosed. Participants emphasized that they utilized different strategies for differentiation of data disclosure, based on different environments, as one of them noted:

"I prefer to split my work and social lives, because I don't want that my lifestyle would be known at work. Thus, even if I did not consider some of the personal data to be sensitive, such as interests or hobbies, I wished to keep it private in my ordinary work environment, as I do not consider them relevant for those situations"

Similarly to the quantitative outcomes, during the qualitative investigations, we also discovered contradicting results in regard to the other three contextual data influential factors, i.e. current mood, activity and location familiarity. When comparing the qualitative answers with quantitative results, the most significant difference was observed in regard to the mood influential factor, because 8 out of 13 participants claimed that their data disclosure is affected by their current humor, as shown in Figure E.5. In fact, two participants, who did not present statistically significant results for the mood factor in the quantitative investigation, stated that their data disclosure is affected by their current humor. The participants commonly agreed that this influential factor would determine their acceptance to exploit USN services, rather than shaping the extent of their data disclosure, as one of them said:

"When I was stressed, tired or irritated, I did not disclose any of my personal information, because I did not want to engage in any new social"

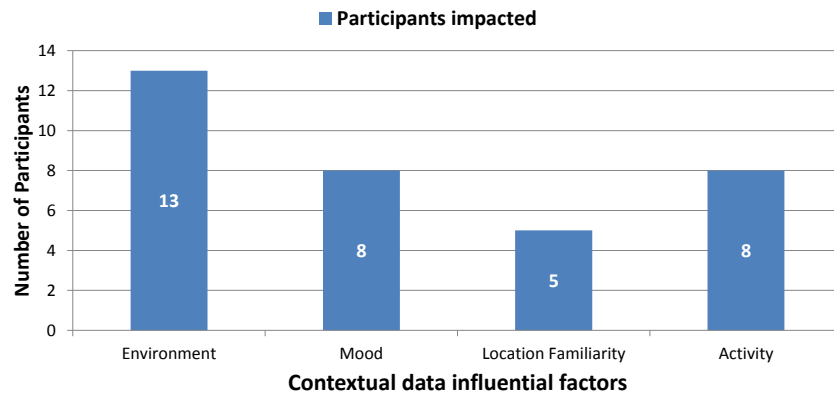


Figure E.5: Impact of the context data influential factors on participants

interaction, even if I probably lost relevant networking benefits. Actually, when I am in those moods, I would prefer to switch off these services”

In regard to the current activity, during the qualitative interviews, 8 out of 13 participants claimed that their data disclosure decisions were influenced by this factor. The majority of the participants noted a relationship between the current environment and current activity for their data disclosure in USN. They discussed that when deciding their sharing preferences, the current environment had higher influence than the current activity. In fact, the participants’ data disclosure decisions were essentially based on the impact of the current environment, but refined by taking into consideration the current activity. However, one of the respondents emphasized that the impact of the current activity influential factor might increase proportionally to the duration of the activity:

”When I had quick coffee breaks with my work colleagues, I did not relevantly change my data disclosure preferences, but during a barbecue event at my work, I additionally disclosed some of my personal infor-

E. THE INFLUENTIAL FACTORS FOR THE VARIATION OF DATA SENSITIVITY IN UBIQUITOUS SOCIAL NETWORKING

mation related to social activities”

Finally, 5 out of 13 participants confirmed the importance of the location familiarity influential factor for their data disclosure decisions. Such inclination can be explained by the fact that some people develop an unconscious trust in more familiar places. This led them to also share more personal information, which would have been detained otherwise. For instance, one of the respondents claimed:

”When I went to a very familiar cafe in the city center, I disclosed personal information that I usually share in all leisure places, but I additionally shared other data, e.g. my political views, which I usually kept private in other social environments. My political views is sensitive information, but I knew that cafe very well and the kind of people that go there, so I believed that most of them were very open-minded. I did not feel that my political views were so sensitive anymore and I decided to share it, as it was relevant in that case”

However, all the other participants, i.e. 7 out of 13, did not provide similar comments. They believed that a more familiar location does not necessarily lead to a more trustable environment, as in such places there is still no control over other people surrounding the user.

Interrelated attributes influential factors

The first interrelated attribute that we discussed with the participants was familiar strangers. As shown in Figure E.6, 11 out of 13 participants confirmed the relevance of being familiar strangers with the inquirer for their data disclosure decisions. Par-

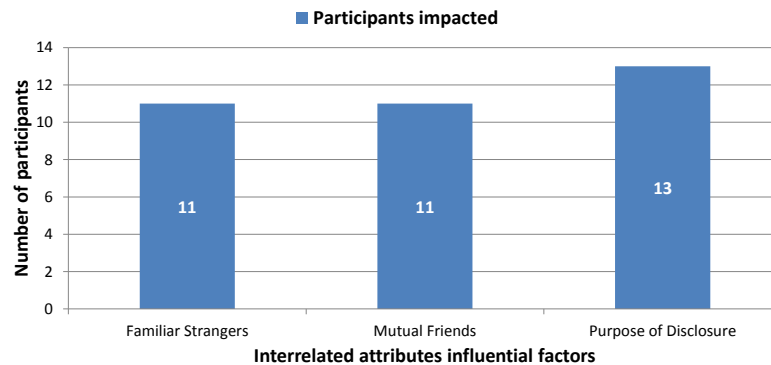


Figure E.6: Impact of the interrelated attributes influential factors on participants

Participants discussed that being familiar stranger with the encountered person might mean that they have been many times at the same locations. Thus, participants expected to have common interests related to that particular location, as for example one of them noted:

”I was aware that we had something in common: we lived in the same neighborhood, we often went to the same poker club, etc. In such cases, I was additionally sharing personal information, previously preferred to be kept private, which was specifically relevant for those circumstances”

Further, some of the participants as well highlighted that knowing the number of previous encounters also provided them a feeling of increased comfort with data disclosure. They were aware that these users were not malicious, i.e. people only interested in retrieving other users’ personal information. A few respondents also felt that it was worth disclosing personal data to these particular people, because they were active users, often exploiting these services. As a result, participants were expecting to have higher probability of receiving potential networking benefits in

E. THE INFLUENTIAL FACTORS FOR THE VARIATION OF DATA SENSITIVITY IN UBIQUITOUS SOCIAL NETWORKING

exchange to their information disclosure.

In regard to the mutual friends, all the participants confirmed the statistical results, as 11 out of 13 of them claimed to be impacted by this factor. They emphasized that this information unconsciously increased a feeling of curiosity about the encountered users and, consequently, motivated them to share more personal data, in order to easily initiate a face-to-face interaction. The respondents also confirmed the statistical results, which did not indicate a significant proportional relation between the increasing number of mutual friends and the probability of disclosing personal information. Instead, all of them emphasized that their data disclosure decisions would probably be impacted by knowing about the identity of the mutual friends, even if this feature was not tested during the quantitative phase of this study. They provided comments similar to the following:

”If the friend that I have in common with the inquirer was a close friend of mine, then I would definitely like to share more of my personal information. On the contrary, if the mutual friend was a person that I do not like or someone who had a strong influence on me (e.g. my boss), then I would probably not disclose some of my personal information”

As shown in Figure E.6, the last interrelated attribute, purpose of disclosure, was found to be relevant for all the participants, because after knowing about similarities with the inquirers, the participants had a reason for sharing their personal information, which would be kept hidden otherwise. Moreover, participants discussed that they were highly motivated to share the same data types as the ones, disclosed by the inquirers. Participants thought that sharing this data might be

relevant for initiating potential face-to-face interactions. For example, one of the respondents claimed:

”When I was utilizing the USN prototype, I received a business card from another user who was from Senegal as me and I also learned that we were both studying at the same university. This information strongly motivated me to share my personal data because I really wanted to know her. Naturally, after receiving her business card, I decided to share the matching data types (i.e. nationality and university) as well as other data types that she had disclosed to me, even if her preferences were not matching mine (e.g. favorite books, movies, etc). I did so, because I assumed that she wanted to know this information about other users, as she was sharing it herself”

Moreover, after receiving the inquirers’ personal information, in many cases, participants significantly changed their data disclosure decisions, which were previously based only on the contextual data influential factors (e.g. location, activity). Participants explained that they were motivated to change their sharing preferences, because they could foresee the relevance for disclosing other personal data. For instance, when being at a social environment, they usually did not include data related to work activities. However, after knowing that encountered users were working in their same professional area, respondents felt motivated to share also data related to work activities, because they expected to receive relevant professional networking benefits in exchange. Finally, the purpose of disclosure influential factor as well encouraged participants to disclose personal information that was usually considered

E. THE INFLUENTIAL FACTORS FOR THE VARIATION OF DATA SENSITIVITY IN UBIQUITOUS SOCIAL NETWORKING

to be too sensitive to be shared, as one of the participants noted:

"I am usually very cautious about disclosing my political views or religion information, because I don't know how other people might react to it. However, when utilizing the provided USN application, after receiving information that the inquirer had matching political views or religion, I did not have anymore concerns about disclosing this information and I felt that it was relevant to do it"

E.7 Discussion

In this paper we describe a mixed methods study, which investigated the influential factors for variation of human data sensitivity upon different circumstances in order to contribute to the design of privacy management systems of ubiquitous social networking. The results of this investigation showed that users prefer to share different subsets of their profiles under different situations. The disclosed personal information was selected by compromising between perceptions of data sensitivity for the current circumstances and evaluations of data relevance for gaining potential networking benefits. We found that participants' data sensitivity was decreasing inversely proportionally to the relevance of information disclosure for initiation of networking.

The current environment contextual data influential factor was considered as a crucial determinant for data disclosure, because it primarily guided the participants in evaluation of their data sensitivity and relevance for exploiting ubiquitous social networking services. Similarly to the current environment, the purpose of disclosure

interrelated attribute as well significantly guided the participants in taking their data disclosure decisions and in some cases, when potential significant networking benefits could be clearly foreseen, this factor was found to motivate participants to alter their data disclosure decisions, based on the contextual data influential factors (e.g. environment, activity).

Following the results of this mixed methods study, we suggest designers of privacy management systems of ubiquitous social networking to take into consideration the other two contextual data influential factors, i.e. current activities and location familiarity, however as indexes of secondary importance if compared to the current environment. In fact, these two influential factors motivated participants to refine grained selection of disclosed personal information, rather than being significant primary predictors for personal information disclosure. The current activity refined data disclosure decisions for the majority of the participants, while the location familiarity presented contradictory results where only a few of the participants were influenced. The last contextual data influential factor, i.e. mood, was found to have impact on overall acceptance to exploit ubiquitous social networking services, rather than shaping the participants' data sensitivity. Thus, we suggest privacy designers to utilize information about user's current humor as a trigger to interrupt their participation in ubiquitous social networking environments.

Among the interrelated attributes, familiar strangers and mutual friends can be considered as relevant predictors for data disclosure in ubiquitous social networking, as they were found to be statistically significant during the quantitative investigation. However, we suggest privacy designers to consider them as indexes of secondary importance, when compared to the purpose of disclosure. In fact, these

E. THE INFLUENTIAL FACTORS FOR THE VARIATION OF DATA SENSITIVITY IN UBIQUITOUS SOCIAL NETWORKING

factors provided a feeling of increased comfort with data disclosure as well as motivated curiosity to start an interaction with other users, rather than guiding the participants in evaluating data relevance for better exploiting these services.

During the qualitative interviews, participants highlighted many aspects of the investigated influential factors that still need further attention. Firstly, participants emphasized that the duration of the current activity might have a different influence on their data disclosure decisions. Especially in case of activities with very long duration, it is suggested to analyze whether the current activity might impact the evaluation of data disclosure relevance more than the current environment influential factor. Further, it is also important to statistically investigate whether knowing the identity of the mutual friends would influence users' data disclosure decisions. Two relevant aspects are suggested to be taken into consideration: self-reported closeness and clustering of users' friends into manageable categories (e.g. co-workers). Lastly, additional analysis with a large scale of participants is required to confirm the results of the pilot test, presented in this paper. As well, due to contradictory results, gained when investigating the location familiarity influential factor, further research is needed to in-depth analyze its influence for the variation of data sensitivity in ubiquitous social networking.

Acknowledgments

This work is supported by Nokia and developed as a part of the Converged Advanced Mobile Media Platforms (CAMMP) project²⁷, funded by the Danish Advanced

²⁷<http://www.cammp.dk>

Technology Foundation. The author is extremely grateful to the participants of the investigation who took the time to take part in this study. Without their participation and feedback, this work would not have been possible. Finally, the author thanks Lene Sørensen and the anonymous reviewers of the journal for their valuable comments on the paper.

**E. THE INFLUENTIAL FACTORS FOR THE VARIATION OF
DATA SENSITIVITY IN UBIQUITOUS SOCIAL NETWORKING**

Appendix F

Legal Documentation

Thesis title Disclosure of Personal Data in Ubiquitous Social Networking

PhD student Antonio Sapuppo

Supervisors Associate Professor Lene Tolstrup Sørensen and Associate Professor Reza Tadayoni

List of papers

A Privacy and Technology Challenges in Ubiquitous Social Networking. Accepted for publication in International Journal of Ad Hoc and Ubiquitous Computing.

B Ubiquitous Social Networking: Concept and Evaluation. Sensor Letters, Vol. 10, No. 8, Pages 1632-1644, 2012.

C Designing for Privacy in Ubiquitous Social Networking. Accepted for publication in International Journal of Ad Hoc and Ubiquitous Computing.

- D** Privacy Analysis in Mobile Social Networks: the Influential Factors for Disclosure of Personal Data. *International Journal of Wireless and Mobile Computing*, Vol.5, No. 4, pp. 315-326, 2012.
- E** The Influential Factors for the Variation of Human Data Sensitivity in Ubiquitous Social Networking. *International Journal of Wireless and Mobile Computing*, Vol. 6, No. 2, Pages 115-130, 2013.

This thesis has been submitted for assessment in partial fulfillment of the PhD degree. The thesis is based on the submitted or published scientific papers which are listed above. Parts of the papers are used directly or indirectly in the extended summary of the thesis. As part of the assessment, co-author statements have been made available to the assessment committee and are also available at the Faculty. The thesis is not in its present form acceptable for open publication but only in limited and closed circulation as copyright may not be ensured.

Co-author statement in connection with submission of PhD thesis

With reference to Ministerial Order no. 18 of 14 January 2008 regarding the PhD Degree § 12, article 4, statements from each author about the PhD student's part in the shared work must be included in case the thesis is based on already published or submitted papers.

Paper title:

Privacy and Technology Challenges for Ubiquitous Social Networking

Publication outlet:

International Journal of Ad Hoc and Ubiquitous Computing

List of authors:

Antonio Sapuppo and Boon-Chong Seet

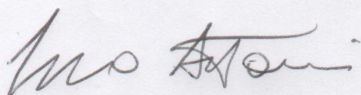
PhD student:

Antonio Sapuppo

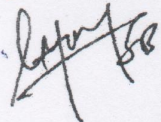
Scientific contribution of the PhD student (all participating PhD students) to the paper:

The first author, i.e. Antonio Sapuppo, carried out the major part. The co-author, i.e. Boon-Chong Seet, contributed to the introduction (Section 1), was consulted on Section 3 in relation to the context acquisition, and provided feedback on the overall structure and content of the paper.

Signature, PhD student



Signatures, co-authors



Co-author statement in connection with submission of PhD thesis

With reference to Ministerial Order no. 18 of 14 January 2008 regarding the PhD Degree § 12, article 4, statements from each author about the PhD student's part in the shared work must be included in case the thesis is based on already published or submitted papers.

Paper title:

Designing for Privacy in Ubiquitous Social Networking

Publication outlet:

International Journal of Ad Hoc and Ubiquitous Computing

List of authors:

Antonio Sapuppo and Joao Figueiras

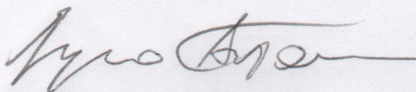
PhD student:

Antonio Sapuppo

Scientific contribution of the PhD student (all participating PhD students) to the paper:

The first author, i.e. Antonio Sapuppo, carried out the major part. The co-author, i.e. Joao Figueiras, contributed to the technical insights in Section C.5.3, as well provided feedback on the overall layout and text of the paper.

Signature, PhD student



Signatures, co-authors

