Computational Social Sciences

Alvin Chin
Daqing Zhang
Editors

Mobile Social Networking

An Innovative Approach

Springer

Computational Social Sciences

A series of authored and edited monographs that utilize quantitative and computational methods to model, analyze, and interpret large-scale social phenomena. Titles within the series contain methods and practices that test and develop theories of complex social processes through bottom-up modeling of social interactions. Of particular interest is the study of the co-evolution of modern communication technology and social behavior and norms, in connection with emerging issues such as trust, risk, security, and privacy in novel socio-technical environments.

Computational Social Sciences is explicitly transdisciplinary: quantitative methods from fields such as dynamical systems, artificial intelligence, network theory, agent-based modeling, and statistical mechanics are invoked and combined with state-of-the-art mining and analysis of large data sets to help us understand social agents, their interactions on and offline, and the effect of these interactions at the macro level. Topics include, but are not limited to social networks and media, dynamics of opinions, cultures and conflicts, socio-technical co-evolution, and social psychology. Computational Social Sciences will also publish monographs and selected edited contributions from specialized conferences and workshops specifically aimed at communicating new findings to a large transdisciplinary audience. A fundamental goal of the series is to provide a single forum within which commonalities and differences in the workings of this field may be discerned, hence leading to deeper insight and understanding.

For further volumes:
http://www.springer.com/series/11784

Alvin Chin • Daqing Zhang
Editors

Mobile Social Networking

An Innovative Approach

Springer

Editors
Alvin Chin
Xpress Internet Services China
Nokia
Building 2, 5 Donghuan Zhonglu
Economic and Technological
Development Area
Beijing, 100176, China

Daqing Zhang
Institut Mines-Telecom/Telecom SudParis
9, rue Charles Fourier, 91011, Evry Cedex,
France

ISBN 978-1-4614-8578-0 ISBN 978-1-4614-8579-7 (eBook)
DOI 10.1007/978-1-4614-8579-7
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013951654

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The usage of mobile computing and social networking has skyrocketed in the past several years. Since the introduction of Apple's iPhone, the mobile phone has now truly become the device that Mark Weiser noted in his vision for ubiquitous computing (Weiser 1999). Since the introduction of Facebook, social networking has become one of the most popular activities on the Internet. We are now entering a new computing era where mobile computing and social networking have combined into mobile social networking – a means for people to socialize and connect directly through their mobile phones.

Mobile social networking is certainly not new terminology coined by us, as there have been many articles, whitepapers, and industry and academic books on or related to this subject. However, previous books have only discussed the business of social media and social networking applications using mainstream applications such as Facebook, Twitter, and LinkedIn (Roebuck 2011), user behavior within these networks (Zhao et al. 2011), and their implications in terms of analysis from the network, data, and sociological points of view (Aggarwal 2011; Fuhrt 2010). These books fail to investigate mobile social networks such as Foursquare that are being created through mobile devices and used during activities. Our book looks at mobile social networks from the micro point of view, that is, at a particular activity and how this can be recorded and shared easily via online social networks. We also investigate specific research issues that have been neglected by previous books. In addition, our book deals with the research and cutting-edge technologies of mobile social networking, and provides comprehensive coverage of applications, data analysis, design, human–computer interaction, and sociology.

The objectives of the book are as follows:

1. Identify current problems in mobile social networking and propose possible solutions
2. Provide examples of real-life applications that illustrate mobile social networking

3. Demonstrate real-life data extracted from deploying the applications in the field
4. Challenge the widely accepted preconceptions of what mobile social networking is within the industry and academic fields

Our book has contributions from leading experts in mobile social networking, covering the areas of data mining, machine learning, ubiquitous computing, mobile computing, trust, human–computer interaction, applications and services, and social computing. These experts come from academia or industry, and are well respected in their areas.

This book targets graduate students and researchers interested in mobile social networking. In addition, we believe that business professionals and CTOs can also benefit from understanding this new technology – how it affects their business, and what issues they should address in order to stay ahead of the competition. We identify the research issues in mobile social networking, and outline a research agenda as to what other research issues still need further study.

This book is by no means meant to be a complete book on mobile social networking, as it is impossible to cover all aspects in a book of this size. However, we do hope that this book becomes a useful tool for industry practitioners and researchers to help advance the field, and to increase awareness of mobile social networking.

We would have been unable to complete this book without the following individuals. First, we thank the contributing authors for taking time in their busy schedules to write their chapters. Second, we thank Springer for giving us the opportunity to help create this book. Third, we thank our institutions for providing us the environment, the inspiration, and the technical and financial support for our research. Fourth, we thank our families for supporting us. And finally, we thank you the reader, for finding this book and reading it. When you read the book, do not just blindly accept what is written. Think about the concept, the methods, the experiments and the results, and challenge them. Do they make sense, do you agree, what is missing, what are the opportunities? Then use that to formulate your own research, and discuss it with others on social media.

If you find any errors or would like to provide comments and feedback, please join our Google group at msnbook@googlegroups.com. And since this is a mobile social networking book, you can use social media to also follow us on Facebook (http://www.facebook.com/MobileSocialNetworkingBook) and Twitter (http://twitter.com/msn_book) to share your comments there.

Finally, we hope that you enjoy reading this book, and we look forward to your comments! Happy mobile social networking!

Beijing, China — Alvin Chin
Evry Cedex, France — Daqing Zhang

References

Aggarwal, C. C. (2011). *Social network data analytics.* Springer, New York.

Furht, B. (2010). *Handbook of social network technologies and applications.* Springer, New York.

Weiser, M. (1999). The computer for the 21st century. *SIGMOBILE Mobile Computing and Communications Review, 3*(3), 3–11, doi: 10.1145/329124.329126, http://doi.acm.org/10.1145/329124.329126.

Roebuck, K. (2011). Mobile *social networks: high-impact strategies – what you need to know: definitions, adoptions, impact, benefits, maturity, vendors*. Tebbo.

Zhao, H. V., Lin, W. S., & Liu, K. J. R. (2011). *Behavior dynamics in media-sharing social networks.* Cambridge University Press, Cambridge, UK.

Contents

1 **Introduction** 1
Alvin Chin and Daqing Zhang

2 **Socially Aware Computing: Concepts, Technologies, and Practices** 9
Zhiwen Yu and Xingshe Zhou

3 **Ephemeral Social Networks** 25
Alvin Chin

4 **Social Behavior in Mobile Social Networks: Characterizing Links, Roles, and Communities** 65
Martin Atzmueller

5 **Mobile Social Service Design for Special Context** 79
Huanglingzi Liu, Wei Wang, Dong Liu, Hao Wang, and Ying Liu

6 **Exploiting Personal and Community Context in Mobile Social Networks** 109
Daqing Zhang, Zhiyong Yu, Bin Guo, and Zhu Wang

7 **Enhancing Mobile Social Networks with Ambient Intelligence** 139
Kevin Doolin, Nick Taylor, Micheal Crotty, Mark Roddy, Edel Jennings, Ioanna Roussaki, and David McKitterick

8 **Data Analysis on Location-Based Social Networks** 165
Huiji Gao and Huan Liu

9 **Towards Trustworthy Mobile Social Networking** 195
Zheng Yan, Valtteri Niemi, Yu Chen, Peng Zhang, and Raimo Kantola

10 **Conclusions** 237
Alvin Chin and Daqing Zhang

Contributors

Martin Atzmueller Knowledge and Data Engineering Group, University of Kassel, Kassel, Germany

Yu Chen Human Computer Interaction Group, EPFL, Lausanne, Switzerland

Alvin Chin Xpress Internet Services China, Nokia, Beijing, China

Micheal Crotty Telecommunications Software and Systems Group (TSSG), Waterford Institute of Technology, Waterford, Ireland

Kevin Doolin Telecommunications Software and Systems Group (TSSG), Waterford Institute of Technology, Waterford, Ireland

Huiji Gao Computer Science and Engineering, Arizona State University, Phoenix, USA

Bin Guo School of Computer Science, Northwestern Polytechnical University, Xi'an, China

Edel Jennings Telecommunications Software and Systems Group (TSSG), Waterford Institute of Technology, Waterford, Ireland

Raimo Kantola Department of ComNet, Aalto University, Espoo, Finland

Ying Liu Intel, Beijing, China

Dong Liu University of Science and Technology of China, Hefei, China

Huanglingzi Liu China Mobile, Beijing, China

Huan Liu Computer Science and Engineering, Arizona State University, Phoenix, USA

David McKitterick INTEL, Collinstown Industrial Park, Leixlip, County Kildare, Ireland

Valtteri Niemi Department of Mathematics and Statistics, University of Turku, Turku, Finland

Mark Roddy Telecommunications Software and Systems Group (TSSG), Waterford Institute of Technology, Waterford, Ireland

Ioanna Roussaki National Technical University of Athens, Athens, Greece

Nick Taylor School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, UK

Wei Wang Adwo Mobile Media Technology, Beijing, China

Zhu Wang School of Computer Science, Northwestern Polytechnical University, Xi'an, China

Hao Wang Babytree Inc., Beijing, China

Zheng Yan The State Key Laboratory of ISN, Xidian University, Xi'an, China
Department of ComNet, Aalto University, Espoo, Finland

Zhiwen Yu School of Computer Science, Northwestern Polytechnical University, Xi'an, China

Zhiyong Yu Institut Mines-Telecom/Telecom SudParis, Evry Cedex, France

Daqing Zhang Institut Mines-Telecom/Telecom SudParis, Evry Cedex, France

Peng Zhang Xi'an University of Post and Telecommunications, Xi'an, China

Xingshe Zhou School of Computer Science, Northwestern Polytechnical University, Xi'an, China

Author Bios

Alvin Chin

Is Senior Researcher, Xpress Internet Services, Mobile Phones Services at Nokia, previously in the Mobile Social Experiences group at Nokia Research Center, Beijing. His current research involves studying user behavior on the mobile web and in social networks, mining the big data from browser logs, and creating recommendations of web content based on user profiling and context. His previous research explored how the mobile phone could be used for creating physical proximity social networks to capture and infer context for social activity and collaboration in offline physical environments and online social networking sites such as Facebook or Twitter, and also designing improved people recommendation systems that take into account physical context. Alvin has Bachelors and Masters degrees in Computer Engineering from the University of Waterloo and a Ph.D. in Computer Science from the University of Toronto, Canada. His research interests include social networking, computer-supported collaborative work, web and data mining, recommendations, context-aware computing, and pervasive and ubiquitous computing. Alvin is an active user of social networking and Web 2.0 technologies, and is active in program committees of many conferences such as CSCW, CHI, SocialCom, Hypertext, UbiComp, CPSCom, and UIC/ATC. Recently, he was the General Chair of the IEEE CPSCom 2013 conference. He is also an Associate Editor for the New Review of Hypermedia and Multimedia journal, and a guest editor in ACM Transactions on Internet Services and Technology. Alvin received a Best Paper Award at the IEEE CPSCom 2011 conference. Alvin has several book chapters, and has presented at both academic and industry events. He can be contacted at alvin.chin@nokia.com, and his web site is at http://www.alvinychin.com.

Daqing Zhang

Is a full professor on Ambient Intelligence and Pervasive System Design at Institute Mines-Telecom, TELECOM SudParis, and CNRS, France. His research interests include context-aware computing, urban computing, mobile social networking, pervasive elderly care, etc. Dr. Zhang has served as the general or program chair for more than 10 international conferences, giving keynote or invited speeches at more than 15 international events. He is the associate editor for 4 journals including ACM Transactions on Intelligent Systems and Technology, etc. In recent years, he has been exploring a new research direction called "social and community intelligence," making use of taxi GPS traces, social media data and mobile phone data to extract human and community intelligence and enable innovative services. He is the winner of the Ten-years CoMoRea impact paper award at IEEE PerCom 2013, the Best Paper awards at IEEE CPSCom 2013, IEEE UIC 2012 and Mobiquitous 2011. Daqing Zhang obtained his Ph.D. from University of Rome "La Sapienza" in 1996.

Chapter 1
Introduction

Alvin Chin and Daqing Zhang

Abstract We have come a long way from face-to-face communication to electronic communication. With mobile phones and devices, online social networks, and internet, we can connect our offline activities and experiences and share them online easily. Sensors in our phones and mobile devices collect context in order to record the activities that we do and the people that we meet. We can truly now do mobile social networking, that is, connect with people to create social networks directly through the phone, rather than connect to people indirectly by adding them on social networks on the phone, which we call social networking on mobile. This presents unique research challenges and opportunities which we introduce in this chapter, and outline the structure for the chapters in this book.

People are social creatures, and we want to communicate and socialize with others. In the physical real world, we form our own social networks of colleagues, friends, family, etc. In sociology, there has been much research into the social networking of individuals in public and private places. We form physical communities in order to provide support to each other and feel a sense of belonging, which McMillan and Chavis called a *sense of community* (McMillan and Chavis 1986). We can now connect with anyone else within six degrees of separation (Watts and Strogatz 1998), which was first demonstrated in Milgram's experiment (Milgram 1967).

As electronic communication started in the 1970s before the start of the Internet, people have communicated through e-mail, electronic bulletin boards, and newsgroups. This essentially has allowed people to communicate at a distance without

A. Chin (✉)
Xpress Internet Services China, Nokia, Beijing, 100176, China
e-mail: alvin.chin@nokia.com; ubiquitousdude@gmail.com

D. Zhang
Institut Mines-Telecom/Telecom SudParis, 9, rue Charles Fourier, 91011 Evry Cedex, France
e-mail: daqing.zhang@it-sudparis.eu

A. Chin and D. Zhang (eds.), *Mobile Social Networking: An Innovative Approach*,
Computational Social Sciences, DOI 10.1007/978-1-4614-8579-7_1,

having to be physically in each other's presence. One of the original electronic collaboration mediums was the WELL where people logged in and communicated with each other using text (Rheingold 2000). People started to talk with each other through this electronic medium and shared information, thus creating a sense of virtual community, i.e., a sense of community in a virtual setting (Blanchard 2008). Much research has been conducted in the area of virtual community from a psychological, social, and computer science perspective, which is comprehensively reviewed by Chin and Chignell (2008).

With the evolution of the Internet, and the Web becoming mainstream in the 1990s, communication and collaboration tools such as Internet Relay Chat (IRC), instant messaging (such as ICQ and Microsoft Windows Live Messenger), wikis, and blogs were used. This made it easier for people to share information and feel even closer than ever before, which made them communicate with not only text but also images and video. These collaborative tools allowed people to start creating persistent conversations (Erickson and Herring 2004). However, this did not create an immersive social space for recording information and connecting with people very easily. Friendster (http://www.friendster.com) was the first company that created a site and system for social networking with people online, which allowed people to find others by browsing other people's profiles and then becoming "friends" with them, enabling people to share content and communicate with each other (Rivlin 2006). The term "social networking" in the online virtual world was originally patented by Friendster (Abrams 2006).

Online social networking did not really truly become ubiquitous to the public and explode until Facebook (http://www.facebook.com) came along. In 2004, Mark Zuckerberg had an idea to allow graduating students to keep in contact with each other online using a virtual type of yearbook, and Facebook was born, as chronicled by David Kirkpatrick in his book 'The Facebook Effect' (Kirkpatrick 2010) which was made into the movie 'The Social Network', also in 2010. As students across the United States started to use this new social medium, and businesses realized the tremendous potential for sharing information and communicating more easily than ever before, the social networking and media revolution started. People could easily communicate their thoughts and share them with others using blogs and social networking sites such as Facebook, Friendster, and MySpace (http://www.myspace.com). Facebook started to create the Status page, where people could announce what they were doing and share that with others, and see the Feed of updates of all the friends, thus creating the 'stickiness' and community experience that to many serves as an addiction. In 2006, Twitter (http://www.twitter.com) was born, which even simplified the broadcasting of content by introducing the microblogging concept of posting an update (called a tweet) within 140 characters. This made it even easier to quickly post updates to friends and to the public, and this started to create real-time data streams and made possible much research on the user behavior, social influence, and spreading of information in Twitter (Java et al. 2007).

With the ubiquity of the mobile phone being used to take pictures and with GPS available in nearly all phones now, Foursquare (http://www.foursquare.com) was born in 2009. This enabled location-based social networking where people can check in at a location and share their status and photos with others (called social location sharing), which helps to bridge the gap between offline and online. Lindqvist et al. (2011) conducted a study to explain why people use Foursquare. Nonetheless, it is the phone that helps to connect people together in a mobile social network, whether it is by phone call, SMS, or an online social networking application. With Wi-Fi and 3G wireless technologies everywhere, the bandwidth is large enough to allow for an always-on connection, and since nearly everyone now has a mobile phone, online social networking and communication has never been so easy.

Social media and social networking technologies such as Facebook, Foursquare, Instagram (http://www.instagram.com), WhatsApp (http://www.whatsapp.com) and Path (http://www.path.com) connect the offline from the capturing of photos in the real world and the location from GPS and Wi-Fi, to the online social networking world of Facebook and Twitter. However, social networking technologies on the mobile phone are considered as social networking on mobile because people can still do social networking on other devices such as PCs and laptops. The mobile device, such as a mobile phone, captures context which includes location and accelerometer, so it is more than just accessing a social network through an application that connects to the Internet. It is about recording context and then connecting people through the common physical context, such as co-location, co-encounter, and co-activity. We are now beginning to see serendipity applications such as Glancee (http://www.glancee.com) which was acquired by Facebook, Highlight (http://www.highlig.ht.com), and Banjo (http://ban.jo), which are social discovery applications that help you discover friends and strangers near you offline then connect to them online, or notify you if you have any online friends near you.

We believe that mobile social networking does not only encompass accessing online social networks from the mobile, but also connecting the offline to online and vice versa via the mobile. In other words, our definition of mobile social networking is as follows. Mobile social networking makes the mobile become an integral part of your social network and lifestyle. It combines distributed content sharing, social networks, sensor networks, and pervasive computing together on the phone in order to provide an integrated experience that fuses physical and digital social interactions through the mobile. An example can illustrate how mobile social networking can help to make your daily life easier, such as providing you an automatic reminder to complete a task, which we call a social reminder. Let's say that you see someone in an elevator that looks familiar. Your phone can display information to inform you who is that person (based on your previous meetings and encounters as recorded by the sensor and social networks), then you know how to greet that person. The system will automatically exchange business cards seeing that you do not have that person's contact in your phone, using pervasive computing technologies. When you get off the elevator, the phone will recognize that your manager is behind you and send you a social reminder that you need to send a sales report to him at the end of the day (after looking at your calendar and your e-mail from your manager).

So what differentiates a mobile social network from social network on mobile? Table 1.1 illustrates this difference, which we explain in detail.

Table 1.1 Social networking on mobile versus mobile social networking

Feature	Social networking on mobile	Mobile social networking
Duration	Continuous	Ephemeral
Offline interaction	None	Activity based
Proximity and location	Co-location	Nearby, encounter
Context sensing	None	Environment and phone sensing
Contact management	Indirect discovery	Direct discovery
Content sharing	Public	Public, private, group
Collaboration	Coordination	Ad hoc

First, when people meet each other, the duration that they meet and social network with each other has a definitive start and end time, which lasts for a temporary period of time before they leave and depart their separate ways. We call this duration of networking ephemeral, as opposed to continuous where the duration of networking has no specific start and end time. Social networking on mobile such as accessing Facebook on your mobile phone has a continuous duration of networking; there is no actual start and end time for the social interactions. Facebook has the Timeline which indicate events during which you and others were apparently doing things together; however, it does not record offline interactions directly. Second, people usually meet each other at events or activities, therefore the offline interactions are within the activity. Think about the last time you met someone and socially networked with them; it was probably at some type of event or activity. However, these interactions are not recorded automatically (at least, they can be recorded, but manually through check-in or posting a status indicating the activity that the two of you were doing together). Third, the granularity of proximity and location is different between social networking on mobile and mobile social networking. With social networking on mobile, the granularity of location is by co-location, i.e., you and someone else are at the same place since the location technology used is usually GPS. Co-location is used by social networking applications such as Foursquare to indicate if you are in the same place as another person, and allows you to check in. On the other hand, with mobile social networking, since the mobile can use Bluetooth, Wi-Fi, or NFC, these wireless technologies can be used to detect other mobiles nearby, or can detect whether two people encountered each other (Xu et al. 2011).

Fourth, with many sensors that are provided on a mobile phone such as the accelerometer, mobile phones can be used to capture various elements of context (Siewiorek et al. 2003). Activity can be inferred from the mobile context (Lee and Cho 2011), which means that in conjunction with proximity and location, it can be used for detecting social interactions for mobile social networking. Fifth, most social networking applications on mobile allow you to discover contacts indirectly, i.e., you have to manually find the contact you are looking for through searching on the social network (as for example in Facebook) for adding contacts. This is very cumbersome, especially if there is more than one person with the same name. However, this problem does not exist with mobile social networking, because through proximity technologies you can find who the person is close to you (within reasonable accuracy depending on the proximity technology), therefore you can easily add a contact directly through your encounter. For example, Find & Connect is a system by Chin et al. (2012) where you can directly add a contact around you from your mobile phone. Sixth, because of the proximity and groups that are formed from the mobile social network, privacy can be finely grained into categories of public, group, and private, with group privacy being specified from the people who were together at the activity. In social networking on mobile, the default for privacy is public, since most people find it cumbersome to specify who they wish to post to and share the message with. Finally, to collaborate with someone using social networking applications on your mobile, you need to coordinate with the person you want to collaborate by sending a message or notifying them. That means a setup time is required to add the person to your social network, before you can start collaborating. For example, in WhatsApp, if you want to set up a group to discuss organizing an event, first you have to find the people in the group in WhatsApp and then send a friend request (if they are not added as your friends already). Second, once they are friends, then you have to create a group and manually invite each friend to the group, and each friend has to accept the invitation. However, with mobile social networking, the collaboration is ad hoc because it records the people who you were with together in the group, which means that there is no need to manually add people to the online group.

We can see that mobile social networking presents the next generation of social networking that really does bridge the gap between offline and online. With that, it presents a new set of research issues and challenges which we outline and explain in the following chapters. Chapter 2 discusses socially aware computing, which is a combination of pervasive computing and social computing, and addresses the main research challenges pertaining to the acquisition, processing, and identification of the sensing data from smart phones. Chapter 3 introduces ephemeral social networks, a type of mobile social network captured during an event such as a conference, and discusses the theory, characteristics, and user behavior of people in ephemeral social networks from the deployment of Find & Connect, an application for finding and connecting with attendees in a conference. Chapter 4 addresses the social behavior in mobile social networks through the characterization of links, roles, and communities found in mobile social networks. Chapter 5 discusses the design of a mobile social service from a human–computer interaction point of view, and discusses the requirements analysis, service design framework, and two case studies applying the framework, one being about a large-scale exhibition service and the other, a local group-buying service. Chapter 6 addresses the personal and community context of mobile social networks, and provides a novel classification of mobile social networks; it further presents a context model, followed by data sources for obtaining context along with techniques for inferring this context. Chapter 7 discusses the community aspect of mobile social networks; the concept that the authors introduce here is that of pervasive communities which combine pervasive and social computing, along with Web and smartphones to become a self-organizing, self-improving and pro-active entity for enabling a personalized and optimum user experience. Chapter 8 addresses location-based social networks, such as Foursquare, and discusses the distinct properties, data analysis, and research issues of location-based social networks from the data mining perspective, as well as applications of data mining to real-world location-based social networks. Chapter 9 addresses the security, privacy, and trust in mobile social networks, where the authors present a trust management framework that supports context-aware trust/reputation generation, trustworthy content recommendations, secure communications, unwanted traffic control, user privacy recommendation and preservation, and other trust and privacy enhancement technologies. Finally, Chap. 10 concludes the book with discussion about outstanding research issues in mobile social networking, and a call for action for a research agenda from both academia and industry to realize this vision.

References

Abrams, J. (2006). System, method and apparatus for connecting users in an online computer system based on their relationships within social networks. United States Patent 7069308. http://www.google.com/patents/US7069308?dq=Friendster+social+networking+patent&ei=LRS7UN-nCOaX1AWDhID4Aw. Accessed 2 Dec 2012.

Blanchard, A. (2008). Sense of virtual community. In S. Kelsey & K. St. Amant (Eds.), *Handbook of research on computer mediated communication* (pp. 325–338). Hershey: Information Science Reference. doi:10.4018/978-1-59904-863-5.ch025.

Chin, A., & Chignell, M. (2008). Automatic detection of cohesive subgroups within social hypertext: A heuristic approach. *New Review of Hypermedia and Multimedia, 14*(1), 121–143.

Chin, A., Xu, B., Hong, D., Wang, Y., Yin, F., Wang, X., Wang, W., & Fan, X. (2012). Using proximity and homophily to connect conference attendees in a mobile social network. In *Proceedings of the IEEE ICDCS'12 international workshop on PhoneCom* (pp. 1–8). Macau, China: IEEE Press.

Erickson, T., & Herring, S. (2004). Persistent conversation: A dialog between research and design. In *Proceedings of the 37th annual Hawaii international conference on system sciences*, Hawaii, USA. Vol. 1. doi: 10.1109/HICSS.2004.1265280.

Java, A., Song, X., Finin, T., & Tseng, B. (2007). Why we twitter: Understanding microblogging usage and communities. In *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis (WebKDD/SNA-KDD '07)* (pp. 56–65). New York: ACM. doi:10.1145/1348549.1348556 http://doi.acm.org/10.1145/1348549.1348556

Kirkpatrick, D. (2010). *The Facebook effect: The inside story of the company that is connecting the world*. New York: Simon and Schuster.

Lee, Y. S., & Cho, S. B. (2011). Human activity inference using hierarchical Bayesian network in mobile contexts. In B-L. Lu, L. Zhang, & J. Kwok (Eds.), *Proceedings of the 18th international conference on neural information processing: Vol Part I. (ICONIP'11)* (pp. 38–45). Berlin/Heidelberg: Springer-Verlag. doi:10.1007/978-3-642-24955-6_5 http://dx.doi.org/10.1007/978-3-642-24955-6_5.

Lindqvist, J., Cranshaw, J., Wiese, J., Hong, J., & Zimmerman, J. (2011). I'm the mayor of my house: Examining why people use Foursquare – A social-driven location sharing application. In *Proceedings of the SIGCHI conference on human factors in computing systems (CHI '11)* (pp. 2409–2418). New York: ACM. doi:10.1145/1978942.1979295 http://doi.acm.org/10.1145/1978942.1979295.

McMillan, D. W., & Chavis, D. M. (1986). Sense of community: A definition and theory. *Journal of Community Psychology, 14*, 6–23.

Milgram, S. (1967). The small world problem. *Psychology Today, 2*, 60–67.

Rheingold, H. (2000). *The virtual community: Homesteading on the electronic frontier*. Cambridge: MIT Press.

Rivlin, G. (2006). Wallflower at the web party. New York Times. http://www.nytimes.com/2006/10/15/business/yourmoney/15friend.html?_r=1. Accessed 2 Dec 2012.

Siewiorek, D., Smailagic, A., Furukawa, J., Krause, A., Moraveji, N., Reiger, K., Shaffer, J., & Wong, F. L. (2003). SenSay: A context-aware mobile phone. In *Proceedings of the 7th IEEE international symposium on wearable computers (ISWC '03)* (pp. 248–256). Washington, DC: IEEE Computer Society.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature, 393*(6684), 440–442.

Xu, B., Chin, A., Wang, H., Chang, L., Zhang, K., Yin, F., Wang, H., & Zhang, L. (2011). Physical proximity and online user behavior in an indoor mobile social networking application. In *Proceedings of IEEE CPSCom 2011* (pp. 273–282). doi:10.1109/iThings/CPSCom.2011.74.

Chapter 2
Socially Aware Computing: Concepts, Technologies, and Practices

Zhiwen Yu and Xingshe Zhou

Abstract The advances of pervasive computing technologies significantly enhance the capabilities for data capture, processing, and usage. The combination of pervasive computing and social computing leads to a new emerging research topic called 'socially aware computing'. This new paradigm aims to leverage the large-scale diverse sensing devices that can be deployed in human daily lives to recognize individual behaviors, discover group interaction patterns, and support communication and collaboration. Smartphones, which are equipped with a variety of sensors and are popular around the world, bring new opportunities for socially aware computing. In this chapter, we introduce the definition of socially aware computing, discuss the main research challenges, and present our work of implementing socially aware computing by using smartphones.

2.1 Introduction

Ubiquitous computing has been proposed and evolved for more than 20 years. A lot of technologies have been investigated and various applications have been developed, such as smart home, smart classroom, and smart meeting. With the rapid advances of embedded devices, wireless sensor networks, and mobile computing, more and more ubiquitous intelligent systems are being deployed in human daily lives. Such systems are integrated with the capabilities of sensing, computation, and communication, which significantly enhance data capture, processing, and usage in ubiquitous computing. Furthermore, it is now possible to sense social context and support social activity.

Z. Yu (✉) • X. Zhou
School of Computer Science, Northwestern Polytechnical University, Xi'an, China
e-mail: zhiwenyu@nwpu.edu.cn; zhouxs@nwpu.edu.cn

A. Chin and D. Zhang (eds.), *Mobile Social Networking: An Innovative Approach*, Computational Social Sciences, DOI 10.1007/978-1-4614-8579-7_2,

In 2005, Alex Pentland proposed the notion of socially aware computing in his paper entitled "Socially Aware Computation and Communication" (Pentland 2005). It aims to capture, quantify, and visualize social context, such as tone, gesture, and posture for enhancing human social interaction. In 2009, David Lazer et al. (2009) proposed using massive data streams collected in the real world for understanding individuals, organizations, and even our society. Their motivation and target are very similar to those of socially aware computing.

Social awareness is a concept that comes from sociology. It is used to describe the capability or phenomena of social communication, such as knowing what behavior is accepted in the society and following the specification. In the area of computer science, social awareness refers to sensing and reacting to social context by computer systems. A system with social awareness can help people understand the current situation, improve their social communication skills, and facilitate efficient social interaction.

According to International Data Corporation (IDC), the number of mobile phones in existence in 2010 was three times the number of personal computers. The number of mobile phones used around the world then reached 5.9 billion in 2011. On the other hand, smartphones are becoming cheap and are now becoming more popular.

Smartphones are equipped with a variety of sensors, such as accelerometer, GPS, digital compass, microphone, camera, etc. More importantly, smartphones are programmable, which enables the development of context-aware applications based on the built-in sensors. Smartphones bring new opportunities for socially aware computing, such as activity recognition, large-scale sensing, mobile social networking, etc.

In this chapter, we first introduce the definition of socially aware computing, then describe the main research issues associated with it. We then present our work of implementing socially aware computing by using smartphones, specifically activity recognition based on smartphones, enhancing social interaction with smartphones, and understanding social relationship with mobile phone data.

2.2 What Is Socially Aware Computing?

2.2.1 The Origin of Socially Aware Computing

Understanding the behavior and interaction of human beings has been a fundamental research topic for many years. Most of the existing studies investigate social relationships based on user survey data (Carley and Krackhardt 1996; Vaquera and Kao 2008). They are constrained in accuracy, breadth, and depth because of their reliance on the data derived from using a self-reported questionnaire. This kind of data has several limitations. First, it is subjective, as the input might be influenced by the subject himself/herself, e.g., concealing the facts for certain reasons. Second, it is a snapshot that is not dynamic or real time. Furthermore, the self-reported survey cannot be conducted with large-scale populations because that results in controlled and limited data.

With the emergence of the Internet and the Web, it is possible to obtain large-scale data for analyzing online behavior and social networks. Tang et al. (2012) propose a method to detect community based on social media data acquired from the Web, such as BlogCatalog and Flickr. Lin et al. (2009) use the Blog and DBLP data to discover communities and analyze community evolution. Chen and Saad (2012) adopt the citation network and trust network on the Web to extract community. However, online behavior is virtual, which makes it essentially different from the temporal–spatial human behavior and interaction in the real physical world.

By using mobile and pervasive computing technologies, it is possible to obtain real-world sensing data for sociological studies. Compared with the self-reported data, the automatically captured data has many advantages: it is objective (or honest) without user bias, and the continuous field data is particularly appropriate for longitudinal studies of human behaviors in their daily lives. Moreover, the data capture can be automatically performed in a large-scale population, and different data processing algorithms can be compared based on a common data set. Leveraging pervasive sensing to collect and analyze the "digital footprints" at community scale, social and community intelligence can be realized (Zhang et al. 2011).

Socially aware computing is essentially the analysis of human beings and their societies, as well as the development of pervasive computing technologies. Therefore, we can say that it is the convergence of social computing and pervasive computing.

2.3 The Definition of Socially Aware Computing

Here we give our general definition of socially aware computing. *Socially aware computing aims to leverage large-scale, dynamic, continuous, and real-time sensory data to recognize individual behaviors, discover group interaction patterns, and support human communication and collaboration.*

The large number of various sensing devices, such as ubiquitous sensors (e.g., RFID, motion sensors, microphone, camera, etc.) and mobile phones (recording the logs of GPS, calls, and messages), combined with e-mail and Web (e.g., DBLP, BBS, social network sites, blogs, Wiki), offer the in situ data for analyzing human behavior and interaction. In addition to analysis, socially aware computing also emphasizes intelligent assistance and support of human behavior and social interaction from the individual, group, and society perspectives respectively.

Figure 2.1 shows the relationship between socially aware computing, physical world, and macroscopic society. By using various large-scale sensors, socially aware computing captures sensory data from the physical world. After processing the sensory data, socially aware computing provides decision support for macroscopic society. On the other hand, socially aware computing receives theoretical guidance from macroscopic society, and gives direct effect to the physical world through intelligent devices and actuators.

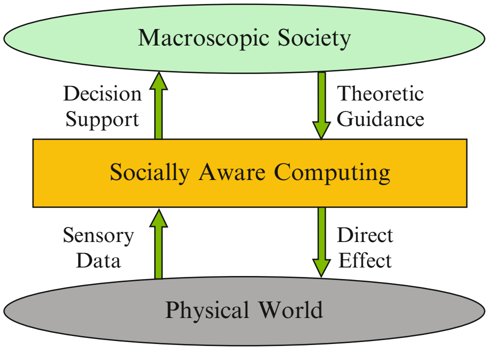


Fig. 2.1 The relationship between socially aware computing, physical world, and macroscopic society

2.4 Main Research Issues of Socially Aware Computing

There are basically five research topics related to socially aware computing: large-scale pervasive sensing, activity and interaction analysis, social interaction support, software framework and methodology, and applications.

2.4.1 Large-Scale Pervasive Sensing

Large-scale pervasive sensing is required for capturing human beings and their society. Three main sensing data sources are mobile sensors, social web, and static sensing infrastructure. Mobile sensors are attached to moving objects, e.g., vehicles and persons. For instance, camera and GPS loggers equipped with a shuttle bus are mobile sensors. A smartphone with various kinds of built-in sensors is a typical mobile sensing platform. Social web refers to the web sites through which people maintain their social networks and interact with each other in the cyber space. The online interaction and user-generated content can be used to analyze human behavior and relationship. Sheth (2009) label Web 2.0 service users as "citizen sensors", and have worked on social event detection from user-contributed contents. Static sensing infrastructure is deployed in our daily life statically, such as surveillance cameras, environmental sensors, and positioning sensors. For instance, it is possible to detect abnormal events using the surveillance cameras widely installed in a city. Temperature, light, and humidity sensors are also widely used for environmental monitoring, for example, to detect a forest fire.

As the sensing data comes from different sources, before it can be used for analyzing human behavior and interaction, several data management operations should be conducted.

2.4.1.1 Multimodal Data Processing

The sensing data can be in different modalities, such as video, image, audio, and structured text. Different sensor types have different attributes and capabilities, such as varying accuracy in sensing the physical and virtual world. Extracting features from the multimodal data is the basis for high-level processing. Another interesting but challenging piece of work is to discover the relevancy from different data sources and modalities.

2.4.1.2 Semantic Representation

To make the data understandable for the machine and usable for external applications, raw data from different sensor sources must be transformed to the same metrics and represented by a shared ontology.

2.4.1.3 Large-Scale Sensing Data Fusion

Sensing data usually has noise, uncertainty, and varying accuracy. Isolated sensing data provides limited information in a complex scenario. Data fusion can be used to address this issue. On the other hand, it is better to fuse the data in order to decrease the data size in transfer and storage.

2.4.1.4 Large-Scale Sensing Data Storage

The three main pervasive sensing technologies mentioned above lead to a very large amount of data generation. Furthermore, the sensing data is continuously generated, which poses hard challenges for data storage, backup, and addition. Data access, i.e., searching or querying particular information from the large sensing data efficiently, is a challenging research topic.

2.4.2 Activity and Interaction Analysis

With the sensing data, it is possible to recognize individual activity and analyze group interaction as mentioned below.

2.4.2.1 Individual Activity Recognition

Activity recognition has been drawing increasing interest from the researchers in the fields of artificial intelligence, personalization, and ubiquitous computing. Chen et al. (2012) present a comprehensive survey to examine the development and current status of various aspects of sensor-based activity recognition. From the viewpoint of sensor usage, activity recognition can be divided into two categories. One is monitoring the movement of the human body by using sensors that are placed on the body. The recognized activities include walking, running, scrubbing, and exercising.

The other approach is monitoring how people interact with objects (e.g., how people move things, usage of objects). Usually this approach is effective in recognizing activities such as grooming, cooking, phoning, toileting, and washing hands. It requires that objects are instrumented with tags, and that users wear an RFID reader affixed to a glove or a bracelet. Patterson et al. (2005) perform fine-grained activity recognition (i.e., not just recognizing that a person is cooking but determining what they are cooking) by aggregating abstract object usage.

The inference model is the key issue of activity recognition. In terms of a learning model, the approaches can be grouped as supervised learning-based activity recognition and unsupervised learning-based activity recognition. The former one is basically a classification problem. It first uses a number of records composed of features and activity labels for training, then it uses the learned model to predict an unlabeled record. The classifiers can be static or temporal. Static classifiers include support vector machine (SVM), naïve Bayes, Bayesian network, and decision tree, while hidden Markov model (HMM), conditional random field (CRF), and dynamic Bayesian networks (DBN) are temporal classifiers. The unsupervised learning-based methods use clustering or mining for activity detection. Phung et al. (2009) recognize user motion state, significant places which the user visits, and user rhythms by using a density-based clustering technique based on Wi-Fi observations. Gu et al. (2009) build activity models by mining a set of emerging patterns from the sequential activity trace only, and apply these models in recognizing sequential, interleaved, and concurrent activities.

Current work on activity recognition has mainly focused on simplified use scenarios involving single-user single-activity recognition (Chen et al. 2012). In real-world situations, human activities are often performed in complex ways. For example, a single user performs concurrent, interleaving, and multi-goal activities. Multiple users perform a cooperative activity. A major issue when observing multiple people is the data association problem: what observations belong to which person? Another complex scenario is recognizing abnormal activities, which is a particularly important task in security monitoring, where suspicious activities need to be dealt with, and healthcare applications, where assistance needs to be provided for incapable users. The most challenging issue of abnormal activity recognition is the unbalanced data problem. A much larger proportion of sensing data is about normal activity, while the data for abnormal activities are extremely scarce, which makes training the classification model quite difficult.

2.4.2.2 Group Interaction Analysis

Compared with individual activity, group interaction is a higher level social semantic. Based on the extracted individual behaviors, social network analysis, machine learning, and data mining techniques can be used to analyze group interaction. Main research topics include group relationship reasoning, interaction pattern discovery, community structure detection, and evolution analysis.

Group relationship can be inferred from sensory data. Eagle et al. (2009) propose to infer friendship based on proximity (via Bluetooth scan). It is based on the common sense and experience that friends usually spend time together in the same physical sites. By using the extra-role factor, i.e., off-campus proximity, they can predict most reciprocal friends and non-friends.

Human interaction is one of the most important characteristics of group social dynamics. Yu et al. (2010a) present a multimodal approach for detecting human interaction based on a variety of contexts, such as head gestures, attention from others, speech tone, speaking time, etc. Discovering interaction patterns is useful for understanding how people interact within a group or with the people in another community. Yu et al. (2013) propose tree-based mining algorithms for discovering patterns of group interaction flow and interaction network in meeting discussion.

Community structure is useful for understanding how people are organized and how information is propagated in the community. Social network analysis methods are often adopted in this study. Onnela et al. (2007) analyze the structure and tie strengths of social and communication networks by using the call records of millions of mobile phones. They find a coupling between interaction strengths and the network's local structure, e.g., social networks are robust to the removal of the strong ties but fall apart after a phase transition if the weak ties are removed.

Investigating how a social group evolves is important to understand community dynamics and predict its future structure. Investigating community evolution is challenging due to the difficulty in obtaining dynamic and continuous human interaction data reflecting the evolution process in the real world. Palla et al. (2007) investigated the stability, group lifetime, and member abandonment in social group evolution by using both co-authorship network and mobile phone call records. Kossinets and Watts (2006) analyzed social network evolution by using e-mail contact data, and found that network evolution is dominated by the network topology and the organizational structure in which the network is embedded.

2.4.3 Social Interaction Support

Differently from social computing and social network analysis, which mainly focus on data analysis, socially aware computing aims to support human communication and collaboration based on the sensed activity and interaction. Thus, social interaction support is the core function in realizing social awareness. It serves as the interface between human beings and the system. The underlying techniques are personalized recommendation, social status visualization, group collaboration, and smart decision-making.

VENETA (Arb et al. 2008) is an application that recommends new friends based on a mobile social networking platform. It uses a decentralized method to explore the social neighborhood of a user by detecting friends of friends that are in the user's current physical proximity. Whenever two mobile phones come into Bluetooth connection range, they compare their contact book entries. If neither of the users appears in the other's contact book (i.e., the users are not friends already) and they share at least one common contact, then the two users are identified as friends of a friend. The objectively mined friendship could nudge a user to be aware of a (statistical) relationship with others.

Yu et al. (2010b) present a graphical user interface for visualizing group social dynamics in a meeting. It helps with meeting people in the organization, and improves people's meeting participation skills. For instance, knowing the current status of the meeting (e.g., did all members agree on a conclusion, who was quiet, who was extroverted), the organizer can make some adjustments to make the meeting more efficient. On the other hand, through interaction visualization, the members become aware of their own and others' behavior in a discussion (e.g., one person speaks for a long time, two people always discuss in a subgroup), and can then make changes to increase the group's satisfaction with the discussion process.

DeaiExplorer (Konomi et al. 2006) extracts social networks from DBLP, a web-based publication database. It builds personal connections from historical records of research activities by taking into consideration co-authoring, publishing in the same proceedings, citation, co-citation, and bibliographic coupling. The extracted social networks are revealed on a big display installed at a conference venue. The co-located conference participants can discover interpersonal connections, and find each other in the physical space through RFID technology. Combining the Web data and physical sensing data, the system successfully supports academic collaboration.

2.4.4 Software Framework and Methodology

To facilitate socially aware computing application development, testing, and deployment, software framework and infrastructure are needed. It offers systematic support for fulfilling common functions, such as heterogeneous data management, social context inference, personalized recommendation, information visualization, etc. Software methodology, i.e., principles guiding the design of models and algorithms, is also required. Furthermore, evaluation standards and methods are required for evaluating socially aware systems. There have been several attempts to develop the software framework and methodology of socially aware computing.

WearCom (Kortuem and Segall 2003) is the wearable community design methodology which facilitates application creation, and provides a framework for investigating the social and technical issues involved. Wearable communities denote the social networks that might emerge when enough people use wearable computing technology throughout their daily lives. WearCom supports an exploratory design approach based on rapid prototyping of wearable community systems. It integrates social and technical concerns, and guides designers from scenario development to implementation. WearCom provides a design language, a design process, and a software platform. The design language permits the specification of important design decisions. The design process outlines an iterative sequence of individual design activities, each of which generates a specific design artifact. The software platform supports the implementation and execution of proactive, presence-aware wearable community applications. Furthermore, six design principles are proposed that contribute to successful wearable communities. Developers can apply these guidelines to evaluate existing designs, guide the design process, and educate designers about the characteristics of successful wearable community systems.

Raento and Oulasvirta (2008) indicate that social awareness applications are based on the idea of a group sharing real-time context information via personal and ubiquitous terminals. Based on the social psychological findings derived from 3 years of research with the mobile social awareness system, nine design principles are proposed specifically for a mobile, ubiquitous social awareness application. The principles are: (1) support lightweight permissions, (2) assume reciprocity, (3) make it possible to appear differently to different people, (4) allow for commenting, modifying, and framing automatic disclosure, (5) provide for feedback, (6) allow the user to lie, (7) do not take control away from the user, (8) allow opportunistic use, and (9) do not try to do everything within the system.

2.4.5 Applications

Socially aware computing can be applied in various areas, such as public health, public safety, and urban planning. For example, through determining friendship based on sensory data, the public health organization can choose to inoculate friends of those randomly inoculated individuals, which has been proved to be efficient in controlling an epidemic outbreak (Cho 2009). Public safety involves the prevention of and protection from events that could endanger the public, such as crimes or disasters. Public video surveillance systems have greatly enhanced citywide event sensing and safety monitoring. With the gathered data, particularly data focused on the time, distribution, and geography of past events, the Los Angeles Police Department generates daily probability reports about when and where crimes are most likely to occur (Greengard 2012). MIT's Real Time Rome project (Calabrese and Ratti 2006) uses aggregated data from cell phones, buses, and taxis in Rome to better understand urban dynamics in real time. It offers support for intelligent transportation management. The Biketastic project (Reddy et al. 2010) improves bike commuting by collecting and mining data that bikers have contributed through their mobile phones. Bikers can then plan routes with the lowest probability of traffic accidents and the best air quality.

2.5 Socially Aware Computing Practice with Smartphones

2.5.1 Activity Recognition Based on Smartphones

We propose recognizing user activities based on a single tri-axial accelerometer in the smartphone. The smartphones are randomly put in the users' pant pockets without any limitations about the phone orientation.

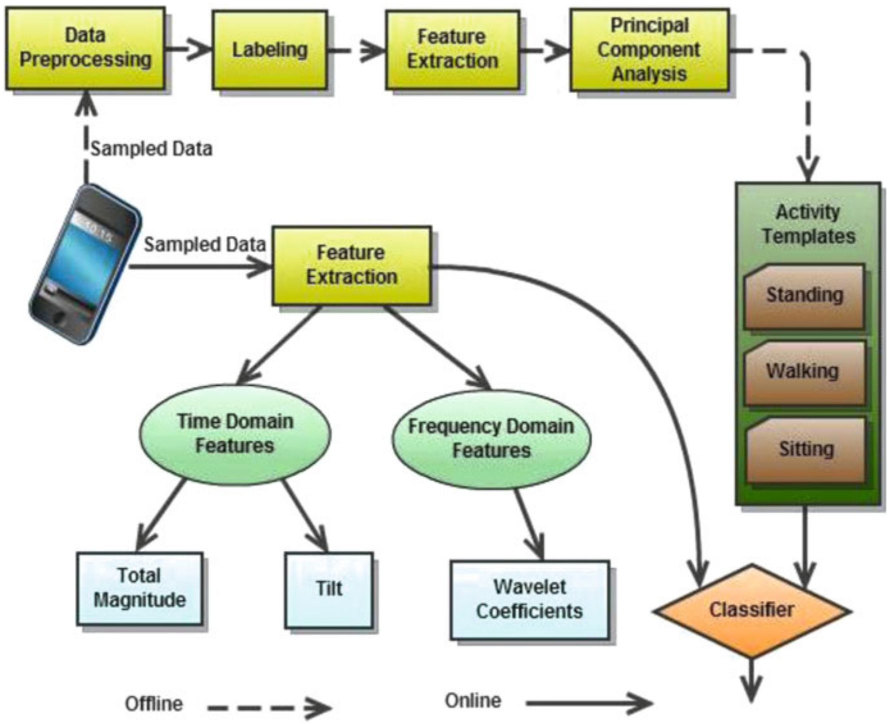


Fig. 2.2 Flow chart of activity recognition including offline learning and online classification

The activities targeted include two types: static activities (e.g., standing, sitting, lying, and driving) and repetitive activities (e.g., walking, running, ascending stairs, descending stairs, cycling, and jumping). Both time-domain features and frequency-domain features are investigated, such as mean of each axis, deviation of each axis, mean of total magnitude, deviation of total magnitude, tilt, linear regressive coefficients, and wavelet coefficients.

Figure 2.2 shows the framework of activity recognition consisting of two parts: the offline data training and the online classification. The offline data training part extracts features from the sampled data, and constructs templates for each activity. The online classification part extracts features of the sliding window, calculates the similarities between the target activity and templates, and selects a suitable class as the label of the sampled data in the sliding window.

The offline data training process consists of four steps. The data preprocessing step takes charge of data cleaning and data representation. Labeling defines the class of each sampled data using all the target activities. Feature extraction captures characteristics of each activity. To minimize the size of features involved, principal component analysis (PCA) is introduced to select the most discriminative features. Finally, a template will be generated for each activity, which describes the crucial feature parameters. To reduce time consumption, offline data training is performed

Table 2.1 Recognition rate

Activity	Percentage of records correctly recognized Time-domain features (%)	Frequency-domain features (%)	Recognition rate (%)
Standing	98.98	1.02	98
Sitting	100	0	100
Lying (prone)	100	0	100
Lying (supine)	99.28	0.72	100
Driving	37.69	62.31	80
Walking	0	100	80
Running	56.76	43.24	86
Ascending	0	100	88
Descending	0	100	82
Cycling	97.50	2.50	84
Jumping	0	100	82
Average	53.66	46.34	89.1

on the PC or workstation. Only those results are transplanted onto the smart phone to serve as templates of user activities.

The online classification process extracts features and calculates similarities using activity templates. According to those similarities, the current inputs are classified into the corresponding type. To reduce computational complexity significantly, no low-pass filter is used. We design a lightweight, hierarchical recognition algorithm with adjusting step length. First, time-domain features are utilized to classify user activities based on template-based classification. However, it is difficult to discern some activities when only the time-domain features are taken into consideration. To discriminate the details of user activities, frequency-domain features are introduced. This algorithm adjusts the size of sliding window according to similarities to enhance recognition accuracy.

The recognition rate for each activity is presented in Table 2.1. The contributions of the time-domain and frequency-domain features are calculated. First, the average recognition rate reaches up to 89.1 %. The recognition rate demonstrates that activity recognition based on the low-resolution accelerometer with low sampling frequencies is feasible. Second, the majority of static activities are recognized in the first phase based on time-domain features. By contrast, most of the repetitive activities are discerned in the second phase based on the combined features. This demonstrates that those selected time-domain features are very useful in discriminating user activities, which favors a decrease in computational load. On one hand, opportunities resulting from the time-consuming frequency-domain feature extraction and the heavyweight decision tree algorithm are minimized. On the other hand, the introduction of classification based on combined features favors improvements in recognition rates, during which complex activities such as ascending and descending are discriminated based on combined features.

2.5.2 Enhancing Social Interaction with Smartphones

We build a service-oriented system architecture to support social interactions in campus-wide environments. The basic functions of the system consist of semantic extraction, pattern mining, ubiquitous search, and location management. The client side is running on smartphones, which collects contexts such as location, proximity, cell phone log, etc., and provides social services to the user for enhancing social interactions.

The server backend consists of several modules ranging from context aggregation, social network analysis, context storage, and knowledge mining to peer communication and admission control.

Generally speaking, campus life consists of study, communication, and entertainment. It is useful for people to learn about the usage of the facilities before planning. For example, *is the study lounge available? which classroom does my classmate sit in? is the tennis-court crowded?* Three applications were implemented and deployed based on the proposed architecture, which are closely related to daily campus life and aims to enhance the social interactions among college students.

The three applications are named *Where2Study, I-Sensing,* and *BlueShare* respectively. *Where2Study* aims to help users find a suitable place to study and locate his/her friends based on Wi-Fi positioning technology. *I-Sensing* is a campus information-sharing system based on participatory sensing, through which every user can publish his/her sensing requests and accomplish others' sensing tasks by using the sensors in their smartphones. *BlueShare* is a media-sharing application among Bluetooth devices based on the opportunistic network. The interesting media is sent to all users close to the Bluetooth devices. The advantage is the ability to transfer large files without any payments.

For space limitation, we here only show the details of the *Where2Study* application. Readers can refer to Yu et al. (2011) for more details about the applications. *Where2Study* not only presents the navigation map of a building to help students find classrooms (Fig. 2.3a), but also shows the status of all classrooms as shown in Fig. 2.3b, such as which classrooms are full and which ones have free seats. To check the detailed information of a particular room, a user can click a button in Fig. 2.3b and then the status of the seats in the room is displayed, as shown in Fig. 2.3c. Furthermore, the application supports querying the location and activity of user friends, as shown in Fig. 2.3d.

A key feature of this application is the capability to browse the status (e.g., name, location, and activity) of close friends. This allows users to reach out and be aware of their social network established by friendship, which will help each other to study (e.g., all my friends are studying at the moment). In addition, when a user encounters a problem during study, he or she could turn to their friends for discussion according to their location shown on the mobile phone.

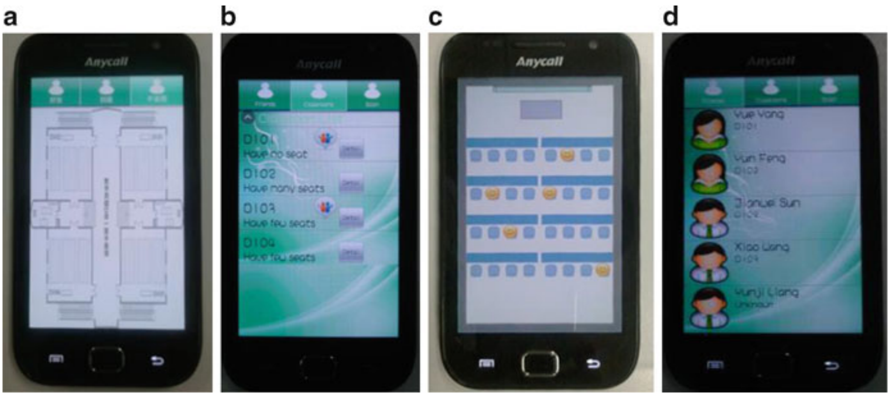


Fig. 2.3 Where2Study user interface: (**a**) map navigation, (**b**) rough status of all (**c**) detailed status of a particular classroom, and (**d**) friend location

2.5.3 Understanding Social Relationship with Mobile Phone Data

Investigating how a social group evolves is important for understanding its community dynamics and predicting its future structure.

We present a study of understanding social relationship evolution by using real-life anonymized mobile phone data. The data was captured by the MIT Reality Mining project (Eagle et al. 2009). An application running on the mobile phones continuously records user behaviors and communications such as location, proximate users, voice calls, and text messages. The collected data is anonymized in further analysis to protect user privacy.

We define a friendship as a directed relationship, i.e., person A regards another person B as his or her friend but not necessarily vice versa. The support vector machine (SVM) approach is adopted as the inference model to predict friendship based on a variety of features extracted from the mobile phone data, including proximity, outgoing calls, outgoing text messages, incoming calls, and incoming text messages. Second, we demonstrate the social relation evolution process by using the social balance theory. For the friendship prediction, we achieved an overall recognition rate of 97.0 % by number and a class average accuracy of 89.8 %. This shows that social relationships (not only reciprocal friends and non-friends, but non-reciprocal friends) can probably be predicted by using real-world sensing data. With respect to the evolution of friendship, we verified that the principles of reciprocity and transitivity play an important role in social relationship evolution.

2.6 Conclusion

This chapter presents the concepts and technologies of socially aware computing. The research paradigm has three features: sensing-based, data-driven, and field-study-based. Our current work in the field is described. Privacy is an important issue we need to consider. Sensing data captured in human daily lives, such as phone call and short message information, is highly sensitive. We need to balance between the benefit derived from the information and user privacy. Robust models and mechanisms are needed to safeguard user privacy during the sharing and usage of the sensing data. In the future, we also plan to apply the technologies in other applications, such as public health, urban transportation management, and environment monitoring.

Acknowledgments This work was partially supported by the National Basic Research Program of China (No. 2012CB316400), the National Natural Science Foundation of China (No. 60903125, 61222209), and the Program for New Century Excellent Talents in University (No. NCET-09-0079).

References

Arb, M., Bader, M., Kuhn, M., & Wattenhofer, R. (2008). VENETA: Serverless friend-of-friend detection in mobile social networking. In *Proceedings of IEEE 2008 international conference on wireless and mobile computing, networking and communications (WiMob 2008)* (pp. 184–189). New York.

Calabrese, F., & Ratti, C. (2006). Real time Rome. *Networks and Communication Studies, NETCOM, 20*(3–4), 1–12.

Carley, K. M., & Krackhardt, D. (1996). Cognitive inconsistencies and non-symmetric friendship. *Social Networks, 18*(1), 1–27.

Chen, J., & Saad, Y. (2012). Dense subgraph extraction with application to community detection. *IEEE Transactions on Knowledge and Data Engineering, 24*(7), 1216–1230.

Chen, L., Hoey, J., Nugent, C., Cook, D., & Yu, Z. (2012). Sensor-based activity recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part C, 46*(6), 790–808.

Cho, A. (2009). Ourselves and our interactions: The ultimate physics problem? *Science, 325*(5939), 406–408.

Eagle, N., Pentland, A., & Lazer, D. (2009). Inferring social network structure using mobile phone data. *Proceedings of the National Academy of Sciences (PNAS), 106*(36), 15274–15278.

Greengard, S. (2012). Policing the future. *Communications of the ACM, 55*(3), 19–21.

Gu, T., Wu, Z., Tao, X., Pung, H. K., & Lu, J. (2009). epSICAR: An emerging patterns based approach to sequential, interleaved and concurrent activity recognition. In *Proceedings of the 7th IEEE international conference pervasive computing and communications*, (pp. 1–9). New York.

Konomi, S., Inoue, S., Kobayashi, T., Tsuchida, M., & Kitsuregawa, M. (2006). Supporting colocated interactions using RFID and social network displays. *IEEE Pervasive Computing, 5*(3), 48–56.

Kortuem, G., & Segall, Z. (2003). Wearable communities: Augmenting social networks with wearable computers. *IEEE Pervasive Computing, 2*(1), 71–78.

Kossinets, G., & Watts, D. J. (2006). Empirical analysis of an evolving social network. *Science, 311*(5757), 88–90.

Lazer, D., Pentland, A., et al. (2009). Computational social science. *Science, 323*(5915), 721–723.

Lin, Y. R., Chi, Y., Zhu, S., Sundaram, H., & Tseng, B. L. (2009). Analyzing communities and their evolutions in dynamic social networks. ACM Transactions on Knowledge Discovery from Data 3, 2: Article 8.

Onnela, J. P., Saramäki, J., Hyvönen, J., Szabó, G., Lazer, D., Kaski, K., Kertész, J., & Barabási, A. L. (2007). Structure and tie strengths in mobile communication networks. *Proceedings of the National Academy of Sciences (PNAS), 104*(18), 7332–7336.

Palla, G., Barabasi, A. L., & Vicsek, T. (2007). Quantifying social group evolution. *Nature, 446*(7136), 664–667.

Patterson, D. J., Fox, D., Kautz, H., & Philipose, M. (2005). Fine-grained activity recognition by aggregating abstract object usage. In *Proceedings 9th IEEE international symposium on wearable computers*, (pp. 44–51). New York.

Pentland, A. (2005). Socially aware computation and communication. *IEEE Computer, 38*(3), 33–40.

Phung, D., Adams, B., Tran, K., Venkatesh, S., & Kumar, M. (2009). High accuracy context recovery using clustering mechanisms. In *Proceedings of IEEE international conference on pervasive computing and communications (PerCom'09)*, (pp. 1–9). New York.

Raento, M., & Oulasvirta, A. (2008). Designing for privacy and self-presentation in social awareness. *Personal and Ubiquitous Computing, 12*(7), 527–542.

Reddy, S., Shilton, K., Denisov, G., Cenizal, C., Estrin, D., & Srivastava, M. (2010). Biketastic: Sensing and mapping for better biking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)* (pp. 1817–1820). New York: ACM.

Sheth, A. (2009). Citizen sensing, social signals, and enriching human experience. *IEEE Internet Computing, 13*(4), 87–92.

Tang, L., Wang, X., & Liu, H. (2012). Scalable learning of collective behavior. *IEEE Transactions on Knowledge and Data Engineering, 24*(6), 1080–1091.

Vaquera, E., & Kao, G. (2008). Do you like me as much as I like you? Friendship reciprocity and its effects on school outcomes among adolescents. *Social Science Research, 37*(1), 55–72.

Yu, Z. W., Yu, Z. Y., Aoyama, H., Ozeki, M., & Nakamura, Y. (2010a). Capture, recognition, and visualization of human semantic interactions in meetings. *The 8th IEEE International Conference on Pervasive Computing and Communications (PerCom 2010)* (pp. 107–115). New York.

Yu, Z. W., Yu, Z. Y., Zhou, X., & Nakamura, Y. (2010b). Multimodal sensing, recognizing and browsing group social dynamics. *Personal and Ubiquitous Computing, 14*(8), 695–702. New York.

Yu, Z., Liang, Y., Xu, B., Yang, Y., & Guo, B. (2011). Towards a smart campus with mobile social networking. In *The 2011 IEEE international conference on internet of things (IEEE iThings 2011)*. New York.

Yu, Z., Zhou, X., & Nakamura, Y. (2013). Extracting social semantics from multimodal meeting content. *IEEE Pervasive Computing, 12*(2), 68–75.

Zhang, D., Guo, B., & Yu, Z. (2011). The emergence of social and community intelligence. *IEEE Computer, 44*(7), 21–28.

Chapter 3
Ephemeral Social Networks

Alvin Chin

Abstract Location-based mobile applications such as Foursquare help bridge the gap between offline and online. People that we encounter and connect with, around physical resources such as conferences, provide opportunities for extending our social networks from offline to online. We call these proximity-based networks that revolve around encounters and activities *ephemeral social networks* because they are created at a specific point in time for a specific duration at a specific event. Ephemeral social networking is the next evolution of mobile social networking, which aims to help us to connect with and recall the people we meet in our daily lives. However, there are many questions that need to be answered. What are the characteristics of ephemeral social networks? How to record and identify an ephemeral social network? In this chapter, we explain the theory behind ephemeral social networks, and create a platform called Find & Connect to show its properties. We then describe our application and system for connecting offline to online, then finally study the influence on user behavior of offline on online and vice versa by deploying Find & Connect at three conference events.

3.1 Introduction

Online social networking sites (OSNs) such as Facebook have blossomed over the past few years; however, they have not been well integrated with real-life interactions. The growing usage of GPS-enabled mobile phones, and location-based applications such as Foursquare, connect users' offline activities with their online social networks, in what is called location-based social networks. At events, people use

---

This work was done while the author was at Nokia Research Center

A. Chin (✉)
Xpress Internet Services China, Nokia, Beijing, 100176, China
e-mail: alvin.chin@nokia.com; ubiquitousdude@gmail.com

A. Chin and D. Zhang (eds.), *Mobile Social Networking: An Innovative Approach*, Computational Social Sciences, DOI 10.1007/978-1-4614-8579-7_3,

social media such as Facebook and Twitter to converse and connect with people. For making contacts, people are beginning to use proximity-based social networking applications such as Bump, Banjo, Glancee, Highlight, and WeChat which use Bluetooth, Wi-Fi, GPS, or even shaking the phone.

However, recording the contacts of people you meet offline onto your online social networks is not trivial. First, this involves remembering the name of the person that you met. Second, it requires finding the person's name in the online social network, which may consist of multiple people under the same name. Third, you must send a friend or follow request to add that person to your network. One of the main objectives when we attend an event, such as a conference, is to meet more people and to establish connections with them. To achieve this, we need to find the right persons to know in the conference who can help us with our line of work and extend our social network. Considering that a conference has a series of social events and an organized program, for new conference attendees it is usually not easy to find similar people in their research area at the conference and make a persistent record of contacts exchanged, because people are busy moving around talking with others. The physical interactions that happen among attendees and their communications are rarely recorded, leaving it hard to recall the person whom you had an enlightening conversation with, or why you should add someone as a contact.

According to previous research (McPherson et al. 2001), when people communicate and interact with each other, they tend to add those who share similar interests and have historical physical proximities. In social science, homophily or social selection is defined as the similarities that people have, such as similar interests, similar education background, and similar consuming habits (McPherson et al. 2001). By discovering the homophily between conference attendees, we can provide them with a way to find the right people that they may want to know. For physical proximity, applications like Highlight help you to discover the people around you, and what you have in common. However, these applications are not designed for a conference or an event, and thus will not help much for connecting attendees from offline encounters to online interactions to extend an ephemeral social relationship to a sustainable social relationship. Another important factor in making friends is social influence, in which an individual's interactions with others who are already friends, his/her social behaviors and activities may change and converge to be in accordance with the behaviors of their friends. Social influence is present in many social settings (Easley and Kleinberg 2010; Vannoy and Palvia 2010), especially in online social networks, for example, the influence from friends such as joining a community in LiveJournal and attending a DBLP conference (Backstrom et al. 2006). In our work, we try to understand how social selection and social influence have an effect on physical proximity in friendship or contact formation. We hypothesize that: (1) for social selection, more physical interactions will result in an increased probability that a person will add another as a friend or contact, and (2) for social influence, being friends or contacts will result in an increased frequency of physical proximity between each other.

Our research problem is to find the relevant people and connect with them easily in a dynamic, ephemeral environment such as a conference. We want to be able to capture dynamic social networks at a particular point in time for a specific duration at a place, so that we are able to remember the event and the people that were there. We call these ephemeral social networks. To achieve this, we develop an application and system called Find & Connect which integrates both proximity and homophily within a conference. For proximity, we use the concept of an 'encounter', which is defined as a physical interaction between two people that move from far away and then come close together, stay together to perform some activity, and then move away. Encounters capture the opportunistic interactions between two people who did not physically meet face-to-face, but should get to know each other due to the homophily. We use Wi-Fi and RFID technology for recording the position of each user, and then calculate the relative distance between the users. We create the system that integrates the encounters and positioning with the conference program, and add social features such as messaging, adding friends, following others, and exchanging contacts. Then, we provide a recommendation system based on common contacts, similar interests, common sessions attended, and common physical encounters. We develop a web client so that users can run the Find & Connect client using a mobile phone, and then conduct a field deployment of our system at three different events: an academic conference with parallel sessions (UIC 2010), a business meeting with a single session (GCJK 2011), and an academic conference with a single session (UbiComp 2011).

We then provide an analysis of user behavior of offline and online (O2O) from Find & Connect. The reason why we study O2O is to clearly understand the influence and interactions between offline and online, which can help us design applications to better connect people from offline to online automatically. Also, the O2O analysis can help us find the causes that trigger the transformation between offline and online in the future. Our contributions are the following. First, we build a system and application called Find & Connect based on social proximity and homophily in a conference context to better help attendees find and connect with each other. Second, we quantify the offline to online (and online to offline) networks created from the encounters, the follow, the friends, and exchanged contacts to study the O2O user behavior. To the best of our knowledge, this is the first work to study the influence of offline to online, and vice versa, using a social proximity-based system. Finally, we discuss the similarities and differences in the results for each of the trials, and the implications for conducting future research in this area.

The chapter is organized as follows. In Sect. 3.2, we first provide theoretical concepts for the definition of what we mean by an ephemeral social network and how it helps to bridge between the offline physical environment and online, and then explain how we create an ephemeral social network in Sect. 3.3. In Sect. 3.4, we describe our platform for ephemeral social networking called Find & Connect, and explain the system architecture and the implementation of how we integrate online and offline together to augment the conference program. We also illustrate the user interface for our Find & Connect application, and go through the use cases that attendees would normally use in the conference for the three trials. In Sect. 3.5, we analyze the user behavior of the attendees from the three trials for the offline and online networks, and study how offline affects online. In Sect. 3.6, we then discuss the lessons that we learned from the trials and from our research, and explain the implications and suggestions for improving our research. Section 3.7 concludes this chapter and discusses avenues for future research.

3.2 Definition of an Ephemeral Social Network

We encounter many people every day, and the people whom we encounter and meet could present opportunities to make new social connections. This is based on the concept of the 'familiar stranger' (Paulos and Goodman 2004), where we repeatedly observe and are co-located with, but do not directly interact with a stranger. In online social networks (OSNs) such as Facebook, our own social network consists of micro-social networks where we physically interact with and are surrounded by people during an activity at a specific event or location. We define these micro-social networks as *ephemeral social networks* (ESN) because the network connections between people are spontaneous (usually not in advance and not scheduled) and temporary (ephemeral), which occur at a specific place or event at a specific time and last for a specific duration. This is similar to the concept of ephemeral groups (Wang et al. 2004), and similar to Kermarrec's (Kermarrec and Le Merrer 2012) paper on ephemeral groups of people in an activity. We can consider the following scenario at a conference to be an example of an ephemeral social network, where people meet during the demo session in the meeting room. It is possible for some people to know each other; some may be friends, and some may not. However, the ephemeral social network is created by interaction, potentially socially, among the participants during the session. Thus, they might find some interesting information, such as who just passed by, or who just gave a talk.

We can then view the ephemeral social network (ESN) as a time slice of your overall social network, which captures the offline interactions during an activity, and can be synchronized with the online social network (OSN), as shown in Fig. 3.1. In addition, ephemeral social networks can bridge the gap between offline and online, since the ephemeral social network provides a permanent trace of your activities, who you met, and any social interactions, from which you can then interact with those people offline after the event.

Current social networking applications focus on online activity, and rely on an internet connection. However, most social networking activity occurs offline at places that we go to, such as entertainment venues, school, family activities, enterprise, emergency situations, and any type of events. The problem is that the social interactions are often not recorded during these times and places, and therefore many opportunities for meeting people are missed. This is where the ephemeral social network helps to record those missed opportunities, and from this, ephemeral social networks can help carry out many daily use cases such as a social reminder. How many times have we forgotten to do a particular task because we forget to record it manually on a piece of paper or a to-do list? With ephemeral social networks, we can remember how we have met another person, which then serves as the context and trigger to perform a particular task. For example, if we see someone

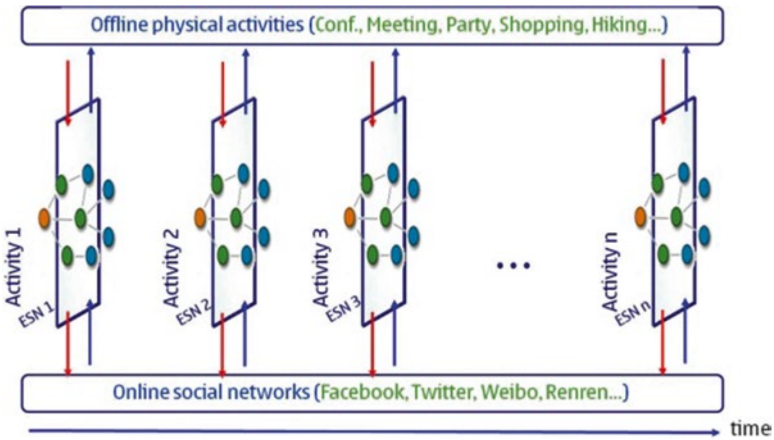


Fig. 3.1 Ephemeral social network bridges the gap between online and offline

who is familiar but do not know who that person is, we could use our phone or glasses to discover who that person is and how, where, and when we last met from the ephemeral social network. Then, by examining the to-do list and calendar entry in our phone, the system could determine that we might have to reply to that person regarding a sales report that she sent. In fact, ephemeral social networks can be used to predict if you are likely to meet another person again in the future (Zhuang et al. 2012), which can help you organize your activity in a conference.

3.3 Creating an Ephemeral Social Network

Now that we understand what an ephemeral social network is, we can begin to explain how the ephemeral social network is created, and how it can be found from the recording of offline and online interactions. During an activity, we can observe that there are three elements that are captured and help to form the ephemeral social network. They are *contact*, *content*, and *context*, as shown in Fig. 3.2. To demonstrate how contact, content, and context revolve around an activity and form the ephemeral social network, we use the example of a salon event as an activity.

3.3.1 Contact

We are surrounded by people during an activity, and we usually meet and interact with people during the activity. For example, if the activity is a salon event, people attend this event to meet new people and engage in interesting dialogue and discussion. People will also exchange business cards so they can remember the person that they meet and record their contact information, so that they can contact that person later.

Fig. 3.2 Capturing an ephemeral social network from the activity

These are the offline contacts that are formed from the meeting and exchanging of business cards. Some examples of applications that record offline contacts include Zaizher, WhatsApp, and WeChat (Weixin), which use physical proximity and shaking for finding contacts nearby. During the activity, people also use their phones and computer devices to connect through social media, such as Facebook and Twitter. This is especially true during a salon event or during a conference. For example, a person may tweet or update a status about the activity, which others can read and comment on, and then decide to add that person to her social network. Or, a person may search for another person's contact in Facebook or Twitter after hearing about them in the activity, and then add that person to her online social network. Therefore, these people form the online contacts. However, adding contacts to your social network is still a tedious process, necessitating a search for a particular person and the possibility that there may be multiple persons of that same name. In addition, it is impossible to meet with everyone during the activity, as everyone is always talking or engaging a discussion with someone else, so the problem is how to recommend particular people that you should meet and add to your social network. The people that you meet, engage with, and encounter in an offline activity then form the basis for the contacts in the ephemeral social network.

3.3.2 Content

When an activity happens, it is usually centered around content, and that content is usually captured by the individuals during the activity, in order to remember what happened during the activity. For example, a salon event usually centers around a topic where a speaker is chosen who is an expert on this topic, which forms the content. The content that is captured during the activity includes the recordings of text, photos, and video that are used for an individual to remember the event. This content can then be shared with others through social networking sites such as Facebook, in order to create conversations and create a memory system such as the memex as advocated by Vannevar Bush (Bush 1949). In this age, with smartphones being ubiquitous, people always capture moments using their smartphone cameras and share them on social networking sites such as Facebook, Instagram, and Path. In this sense, the content captured is the offline event which is synchronized to online. However, the content that is uploaded to social networking sites from mobile devices is still associated with the person's profile and/or associated with a place if it is uploaded using check-in sites such as Foursquare. Therefore, it is difficult to search for and view that content from that particular activity, since due to the nature of the social network stream, that content becomes buried as new recent content appears. The contacts that form the ephemeral social network will upload photos, video, and text, and therefore that forms the content in the ephemeral social network.

3.3.3 Context

An activity occurs at a specific time at a location for a fixed duration, and the data that captures the situation and the environment is known as context, as defined by Dey (2001). In our example with the salon event, the GPS location (geographic location), the name of the place and room that the salon event is held (semantic location), and the particular actions that the contacts in the ephemeral social network are performing, form this context. There may be people in the salon event who are in a group together talking with a speaker of the event, thus forming the context. The context can be captured using the phone's sensors as well as from the offline environment. Context can be divided into physical context, social context, and situational context. Much research has been done for capturing context (Henricksen et al. 2002), inferring context (Dey et al. 2001) and for activity recognition (Choudhury et al. 2008). Many types of context can now be captured by the phone (Lane 2012), and datasets have been created such as the one created by Nokia (2012), that record the context from mobile phones for research purposes.

However, there still does not exist an application or system that can easily record encounters and interactions between people, and memorable moments in offline social activities, in order to relive the experience. This is because the contact, content, and context within an activity or event are not integrated with each other, and multiple applications are currently used in order to capture this. The opportunity is then how to connect online and offline social experiences through cyber–physical social convergence.

3.3.4 Recording Encounters

Before creating an ephemeral social network, we need to first record proximity interactions between people. We can choose to use Wi-Fi (Moraes and Nunes 2006), Bluetooth (Kostakos and O'Neill 2008), RFID (Cattuto et al. 2010), or GPS (Eagle

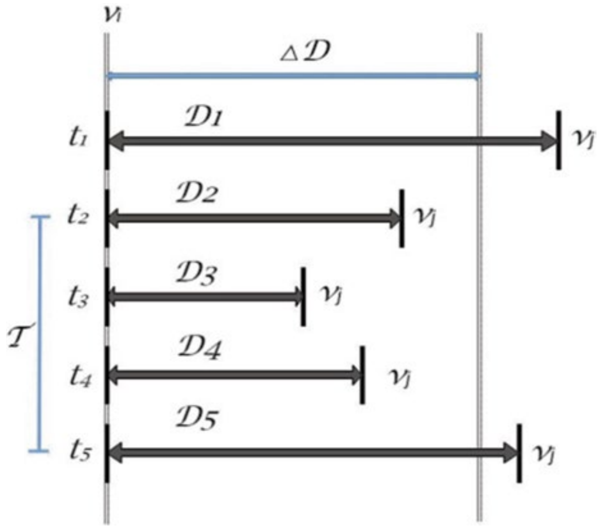


Fig. 3.3 Defining an encounter between two users v_i and v_j

and Pentland 2009) wireless technologies to record the positioning of users, and then compute the distance between these users in order to determine if they are proximate or not. We use encounter as the concept for defining a proximity interaction between two people. We define an encounter if the distance between two people is within the *encounter distance threshold* and their distance lasts for at least the *encounter duration threshold* before they move away and are beyond the encounter distance threshold. We then create an encounter graph $G_{en}(V, E)$ that represents the proximity graph of all proximity interactions between people where V is the set of nodes ($v_i \mid 1 \leq i \leq N$), N is the number of nodes and E is the set of edges ($e_{ij} \mid 1 \leq i \leq N, 1 \leq j \leq N, i \neq j$) and

- Node v_i designates user i, node v_j is user j, and the edge e_{ij} is a link when two users (v_i and v_j) encounter each other.
- Edge (e_{ij}) has a timestamp attribute to define when the encounter happens called $T_{en\ start}(e_{ij})$ and when the encounter ends called $T_{en\ stop}(e_{ij})$.
- Edge e_{ij} is built only if the encounter distance $D_{en}\ (e_{ij})$ is less than the encounter distance threshold ΔD and the encounter duration $\Delta T_{en}\ (e_{ij}) = T_{en\ stop}(e_{ij}) - T_{en\ start}(e_{ij})$ is larger than ΔT (the time duration threshold that is defined to be an encounter).

Figure 3.3 illustrates an example that explains the concept of an encounter. Suppose we have two users v_i and v_j who encounter each other. At t_1, v_i, and v_j are at distance D_1, which is greater than the encounter distance threshold ΔD. These two users may be at the coffee break at a conference where they will eventually encounter each other and meet. At t_2, v_j may discover that she recognizes v_i and starts to move closer to meet her. She is now within the encounter distance threshold as $D_2 < \Delta D$. At t_3, v_j moves closer and at this point, v_i and v_j start talking to each other and their distance is now D_3, which is also within the encounter distance threshold ΔD.

After they finish talking, at t_4, v_j then moves apart from v_i (perhaps to go and meet someone else) and their distance is now D_4, which is still within the encounter distance threshold ΔD. At t_5, v_j has moved away from v_i and is beyond the encounter distance threshold ΔD. According to our definition, the encounter duration is T which is $t_5 - t_2$, and we record an edge in the encounter graph between v_i and v_j.

We then record all the encounters between any two people throughout the activity duration. The collection of all the encounters forms the encounter network which is recorded as the encounter graph $G_{en}(V, E)$. Note that the encounter is defined as a pairwise proximity interaction between two people, and that by mining the encounter graph, our goal is to identify encounter patterns between groups of people in co-located areas that can possibly indicate the existence of an ephemeral social network. Therefore, we study the user behavior through offline interactions in the encounter network.

3.4 Platform for Ephemeral Social Networking: Find & Connect

Ephemeral social networking happens in many environments such as the workplace, conference, social event, sports event, and entertainment event. Find & Connect is a platform that we create for providing ephemeral social networking among users at a conference (Chang et al. 2011; Chin et al. 2012), meeting (Xu et al. 2011a), or workplace environment (Zhu et al. 2010), by combining the features of offline and online. Users can find where the room, session, and people are on the map, and then connect with the people by adding them as a connection, sending them a message, or sharing an item with them. We provide the system description and architecture and then explain the features that we create for social networking at three conferences: the UIC 2010 conference (research conference with parallel track program), GCJK meeting (business conference with single-track program), and the UbiComp 2011 conference (research conference with single-track program). We now explain the system and architecture for Find & Connect below.

3.4.1 System and Architecture

Figure 3.4 shows the system architecture of Find & Connect, which is a client–server system. The user's position is calculated by having a *Positioning Client* run on a mobile phone or a separate positioning device such as a badge that sends the signal strengths from nearby access points on a floor and sends them to the *Positioning Server*. The wireless technology to record the position of the user can be Wi-Fi, RFID, Bluetooth, GPS, or cell ID. For recording the position of users in indoor environments such as the workplace or conference, RFID and Wi-Fi

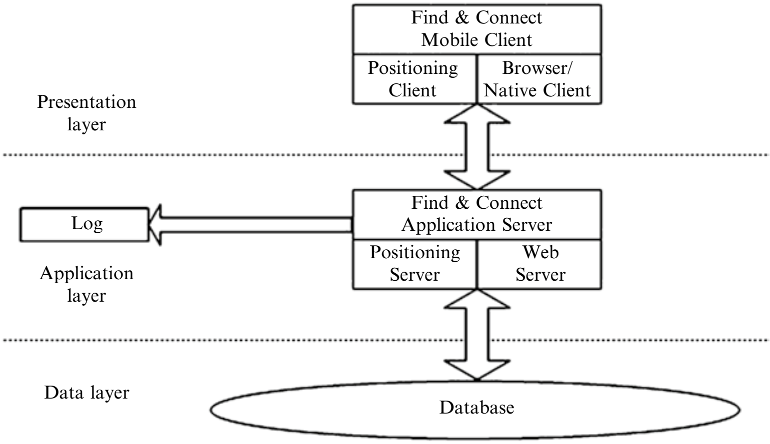


Fig. 3.4 System architecture of Find & Connect

technologies are most often used. The Positioning Server uses the *Positioning Model* to approximate the positioning of the user as an (x, y) coordinate according to the dimensions of the floor in an indoor environment. For Wi-Fi, the Positioning Model is created by performing a site survey that records the Wi-Fi signal strengths from Wi-Fi access points on the floor. For our implementation, we use an off-the-shelf commercial Wi-Fi positioning system, called the Ekahau Real Time Location System (Ekahau 2011). The *Browser/Native Client* presents the location-based social network services in a web browser or a native application on the mobile phone which is accessed from the Find & Connect *Application Server*. The *Log* records client requests including time, username and page accessed, and location, while the *Database* records all the data pertaining to the user and the conference/meeting.

In terms of functionality, Find & Connect contains five major components as illustrated in Fig. 3.5, which explains how the system revolves around the conference/meeting and how we integrate the offline and online together. The novel part is that we add a personalized recommendation of which contacts to connect to, based on offline and online, whereas previous work did the latter but not the former.

1. Offline System
   The *Offline System* records all offline connections and interactions, and forms the context in the ephemeral social network. It consists of a *Positioning System* to record the updated locations of all the attendees. The positioning wireless technology could be RFID, Wi-Fi, or even cell ID. Regardless of the technology, the *Encounters* module calculates the encounters through the absolute positioning coordinates recorded from the positioning system to measure people's offline interaction and frequency, as explained in Sect. 3.3.4.

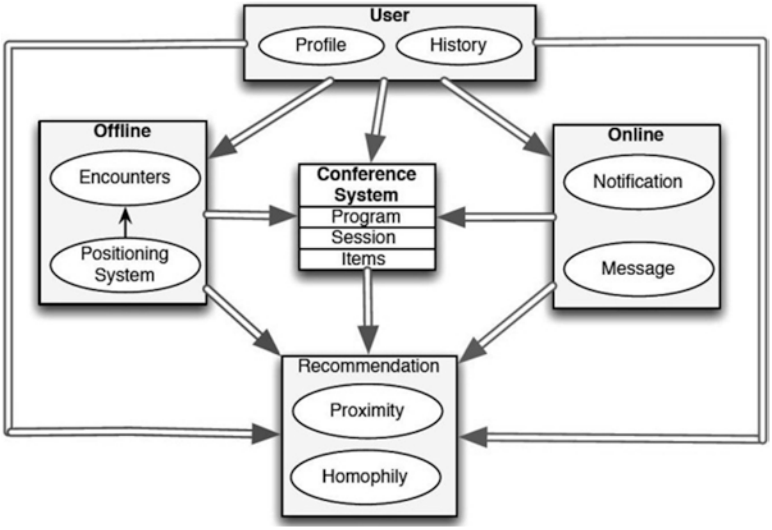


Fig. 3.5 Functional architecture of Find & Connect

2. Online System
   The *Online System* supplies users with a platform to prolong their ephemeral offline relationships to permanent online connections. People can send messages to the target people through the *Messages* module, such as contact requests, as well as get notifications of contact, friend, or follow requests from the *Notifications* module. The conversion of permanent online connections into friends, followers, or exchanged contacts forms the contacts in the ephemeral social network.
3. Conference System
   The *Conference System* is the central part of Find & Connect, and works as a bridge to help people connect their offline social networks to online social networks in the ephemeral social network through the conference program. It manages all the conference contents, including *Program*, *Sessions*, and other *Items* in the session such as papers, demos, posters, and videos. Therefore, this forms the content component of the ephemeral social network, from which the offline and online components are integrated.
4. Recommendation
   Even though Find & Connect allows attendees to find others and connect to them easily by sending friend or contact requests or even following others, this still requires manual effort, and attendees may still miss other important people they should connect with. Therefore, we create an intelligent *Recommendation System* to suggest to the user who she should add to her social network based on *Proximity* and *Homophily*. For proximity, we include the history traces, offline encounters, and common sessions attended. For homophily, we include the

common research interests from her user profile, messages, and common friends with others. This enables users to find the relevant people effectively, and then connect to them in online networks.

5. User

   For each user, the *User* module in our system records *Profile* and *History* traces, including where the user has been in the conference and who the user encountered, to support offline and online interactions as well as supplying the proximity and homophily information to the recommendation system.

3.4.2 Connecting Offline to Online

We describe the information that we collect and use for offline and online, and how offline is used to connect to online in our Find & Connect application, for the UIC 2010 research conference, GCJK 2011 business meeting, and UbiComp 2011 research conference.

A. Offline Information

1. *Location*: At a conference/meeting, we can record the position of users using RFID or Wi-Fi. At the UIC 2010 conference, we used Wi-Fi to record the position of phones (for which we deployed Nokia XpressMusic 5800 and Nokia X6 phones to attendees) with Wi-Fi access points deployed in each room on the conference floor. To calculate the actual position, the Positioning Client on the phone collects the Wi-Fi signal strengths for the nearest Wi-Fi access points, then sends it to the Positioning Server where it uses positioning algorithms to estimate the actual position compared to a positioning model (which records the signal strength map of the entire conference floor). For the GCJK 2011 business meeting, we also used Wi-Fi, similarly to UIC 2010 where the attendees used their own Nokia phones. For these two events, we used the Ekahau technology for the Wi-Fi positioning (Ekahau 2011). For the UbiComp 2011 conference, we used active RFID technology instead of Wi-Fi to decouple the positioning from the access and allow for any device to be used. To calculate the actual position of the user (who carries an RFID badge) in the conference rooms, we adapt the LANDMARC algorithm (Ni et al. 2004; Huang et al. 2007), where we use RFID readers and RFID reference tags (which are RFID badges) deployed in the room. We do not explain the details here due to space limitations; however, for example, in the main conference room, we use four RFID readers and 14 RFID reference tags (Lyu et al. 2013).

   Another piece of offline information is the session, where we record which session a user is currently in by using the user's position and comparing it to the room's coordinates, in order to determine if the user is in the room. From the current time of the user's position and the room, we can determine from the conference schedule which session the user is in, and therefore maintain a history of sessions that the user attended.

2. *Encounters*: We calculated the encounters according to Sect. 3.3.4 for each of the UIC 2010, GCJK 2011, and UbiComp 2011 conferences, for each user. By recording a history of all encounters, we determine whether a user encountered another person before, when, and the number of times, which is displayed in the In Common page when a user views that person's profile, as shown in Fig. 3.12b.

B. Offline Information

In Find & Connect, the online information consists of browsing and interacting with people and the conference schedule within our application. Browsing includes seeing the people that are in the session, browsing through the conference schedule, and viewing the notifications sent to the user such as contact requests, recommendations of people a user should connect with, and public notices. For example, if a person is close to you, people with similar interests are close to you, or if you know of a particular person that you can search for online, then you can easily add that person as a contact by sending a request message. The recipient will receive the contact request, along with the reason, from which she can also add you back as a contact.

3.4.3 Find & Connect at a Research Conference with Parallel Tracks (UIC 2010)

We deployed Find & Connect first at the UIC 2010 conference, which was an international conference for researchers on ubiquitous, intelligent, and autonomic computing held in Xi'an, China, in October 2010 and had parallel track paper sessions. There are five functional modules in our web client, which are *My Agenda*, *Program*, *Map*, *Social Network*, and *Buzz*; we describe the features of each below.

1. Program and My Agenda
   We create a *Program* page (Fig. 3.6a) for showing the conference program items. Users add program items to *My Agenda* by selecting "*Add to agenda*" (Fig. 3.6d). Users can view the session details in Fig. 3.6b by selecting the *Detail* button from Fig. 3.6a, see all papers in that session in Fig. 3.6b, then see a specific paper in that session in Fig. 3.6c. If users like this paper, they select "*Favorite this*" and can share to others in the conference by selecting the "*Share*" button, from which others will receive an SMS message indicating the name of the paper that they shared.
2. Map
   In the *Map* screen shown in Fig. 3.7, users can see the locations of people who are using this system in the conference. Selecting "*Search Online People*" can help them find the locations of the people they want to meet. Users can filter the list of people on the map by showing just friends, people in sessions, or only themselves. By selecting the link of the person in the "*User list*" or on the person icon on the map, users can see that person's profile. By selecting a specific user in the user list, users will see only that person on the map.

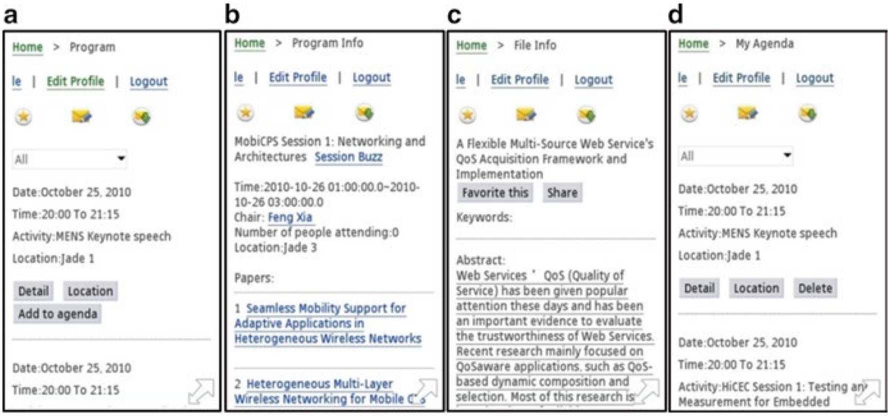


Fig. 3.6 Program and My Agenda pages from the UIC 2010 conference. The program schedule is shown in (**a**), details of the program item in (**b**), papers in the session in (**c**), and the agenda of sessions a user has added in (**d**)

Fig. 3.7 Map page of the conference floor and locations of people at the UIC 2010 conference

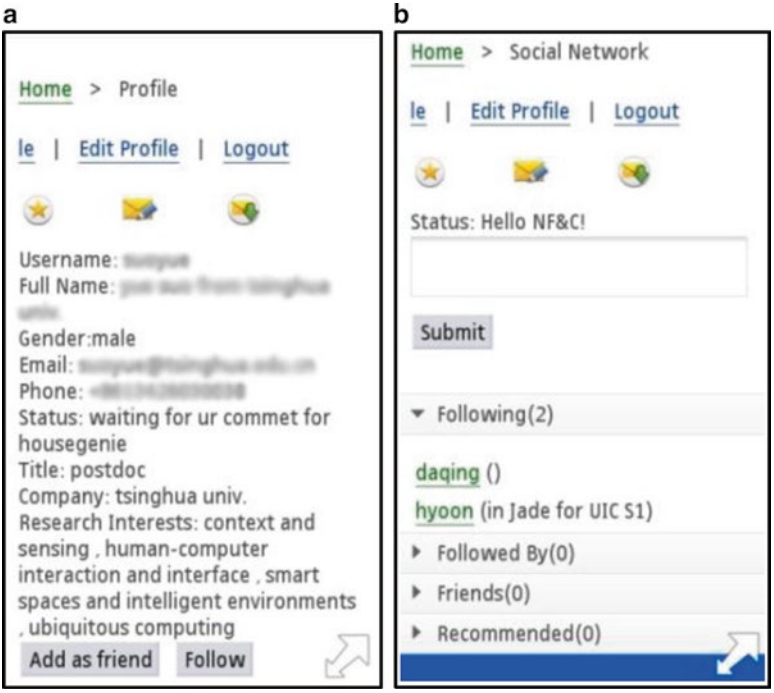


Fig. 3.8 Profile (**a**) and Social Network (**b**) pages from the UIC 2010 conference

3. Profile and Social Network
   Figure 3.8a shows the *Profile* page of a person. When the user selects a friend or a contact nearby, the user can look at the contact details, download the contact's business card to the phone, and find out the last encounter time and location with that person. When someone wants to add a user as a friend, they can select the *Add as friend* button; the other individual will then receive a SMS and then decide whether to accept. Users can also choose to follow the person whom they are interested in by selecting the *Follow* button. In the *Social Network* page on Fig. 3.8b, users can update their status and monitor their social network activity by receiving latest updates from all the people they are following, those that follow them, their friends, and those that Find & Connect recommends to them as friends (Xu et al. 2011b).
4. Messaging and Buzz
   In Find & Connect, users can communicate in three ways. First, users can send a standard message to one or many users, as shown in Fig. 3.9a. Second, users can send a location message where the message is sent to only those people in the specified location selected from the "*Location*" dropdown box. If users choose "Location", they can select the "*Expiration Time*" as to when that message will continue to be sent to people in that room until the expiration time. Third, users

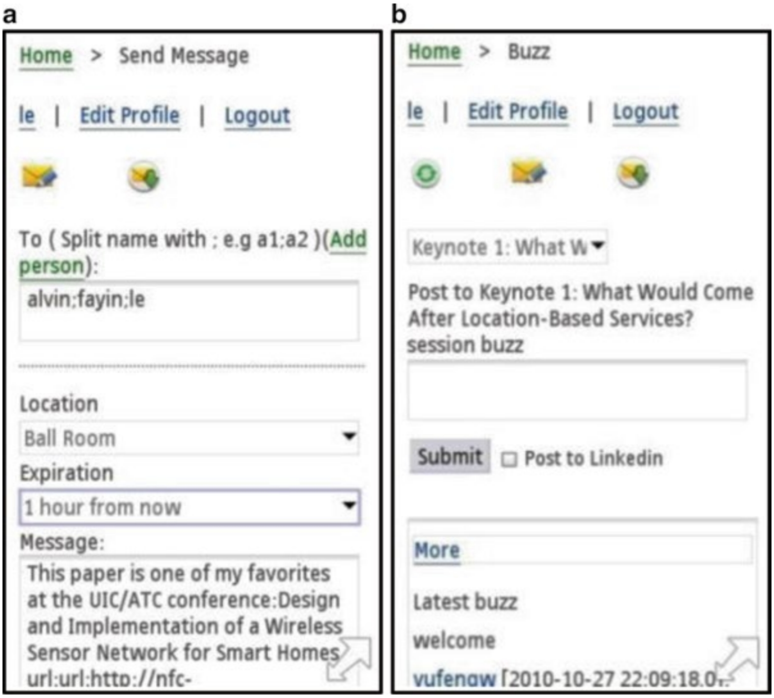


Fig. 3.9 Messaging (**a**) and Buzz (**b**) pages from the UIC 2010 conference

can also discuss the various sessions in the conference by posting a message to that session and viewing other people's posts, as shown in Fig. 3.9b. We call this *Buzz* and users who post the buzz can also post to LinkedIn for others to view, especially those that did not attend. In this way, this is like an activity-based tweet where it is grouped by session, unlike tweets in Twitter which are unstructured.

3.4.4 Find & Connect at a Business Conference with a Single Track (GCJK 2011)

Compared to the UIC 2010 conference, the GCJK business conference was single track and lasted for 1 day on April 13, 2011 in Beijing, was held in a large main meeting room of a conference hotel with eight Wi-Fi access points placed in various locations in the room, and did not have any papers. Therefore, we concentrated more on the social networking features, and developed a simplified version of Find & Connect (Xu et al. 2011a). The event was divided into 19 activities. Users were encouraged to download and use the Find & Connect client throughout the event; a

Fig. 3.10 User interface of Find & Connect in the GCJK business conference, where (**a**) shows the Status and Map, (**b**) shows the Profile page, (**c**) shows the Network page, and (**d**) shows the All Updates page

total of 76 users downloaded and used the client. The features are similar to the UIC 2010 conference, except there was no Program or My Agenda, and we relaxed the social linking to be one-way follow instead of two-way friendship, to allow more users to feel comfortable in using the social features of our software. Here, users connected only with people who were registered, and the follow recommendations were only among the participants. We also added an integrated social network feed to view the latest network activity, in similar fashion to Facebook and Twitter. The user interface is shown in Fig. 3.10, and instead of using horizontal tabs, we use a vertical list format so this can be easily expanded and contracted and can be displayed on all Nokia phones. Figure 3.10a shows the home page consisting of *Status*, *Map*, *Profile*, *Network*, and *All Updates*.

1. Status and Map
   Similar to the UIC 2010 conference, attendees could also post and edit their latest status and share their status with other users of Find & Connect in the *Status* section shown in Fig. 3.10a. The *Map* section, expanded in Fig. 3.10a, shows the location of the current user and other online attendees who had allowed location sharing (set by selecting the *Privacy* button).
2. Profile
   Selecting the name of a person on the map will show the person's profile, as shown in Fig. 3.10b. Users can visit each other's profile, share their location with others (who to share location with is governed by the *Privacy*), and view where others are and their status. Then, users can connect to that person by following that person and/or exchanging contacts with each other through sending business cards via SMS, where the profile information is converted into a virtual card format (vCard), exactly as in the UIC 2010 conference. Find & Connect also

shows whether users have encountered that person, which then provides the incentive to follow or exchange contacts with that person.

3. Network
   The *Network* section shown in Fig. 3.10c is similar to the Social Network page in the UIC 2010 conference. We presented an aggregated list of all the users that they have a connection with, grouped into "*Contacts Exchanged"* (those you have exchanged business cards with), "*Following*" (those that a user follows), "*Following me*" (those that follow the user) and "*Encounters*" (those who have been in close proximity with the user within 4 m). Finally, "*People You Should Follow*" shows all the people that a user should follow (similar to Twitter) and is obtained from our recommendation algorithm (Xu et al 2011b).
4. All Updates
   The *All Updates* section, as shown in Fig. 3.10d, is a new section that we added to GCJK, whose function is similar to the Wall feature in Facebook, where it provides a feed of all the updated interactions and content from the people that the user follows, from which she can follow their activities and make connections with others. These updates include their status history, people they followed, and people they exchanged contacts with.

3.4.5 Find & Connect at a Research Conference with a Single Track (UbiComp 2011)

We deployed Find & Connect at the UbiComp 2011 conference, which was a top international conference for researchers in ubiquitous and pervasive computing held in Beijing, China in September 2011 (Chin et al. 2012; Zuo et al. 2012). In contrast to UIC 2010, UbiComp 2011 had single-track paper sessions. Since the positioning system was decoupled from the phone, to make Find & Connect easily useable by all attendees in the conference, we decided to create a web application that could be run in any web browser on any device such as a smartphone, laptop, or tablet. Our user interface is divided into three parts: *People*, *Program*, and *Me*.

1. People
   We provide a view for grouping people based on location so a user can view who these people are, and then decide if she wants to establish a connection with them. Searching is also offered by using names of target users. The user interface is shown in Fig. 3.11.

   Nearby, Farther, and *All*. The People page in Fig. 3.11a is broken down into view tabs which are *Nearby* (shows a list of all people that are nearby within 10 m of a user's location), *Farther* (shows a list of all people that are farther away greater than 10 m but still in the same room where the user is), or *All* (shows a list of all attendees in the conference). A user can then view each attendee's profile by selecting the icon of that attendee. The user can also select the *Interests* button to group the list of users according to their research interests (which they

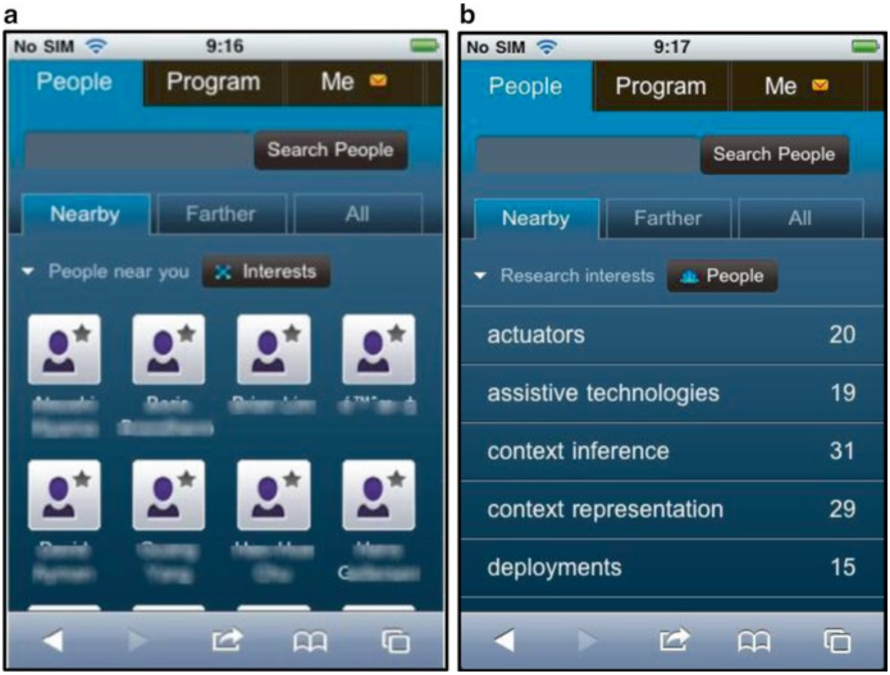


Fig. 3.11 People page in Find & Connect, where (**a**) shows the people nearby and (**b**) shows the interests of the people nearby

enter in their Profile) as shown in Fig. 3.11b. We design it like this so users can find others based on similar research interests easily.

Profile and In Common. Figure 3.12a shows a user's *Profile* page, and by selecting the *In Common* tab in Fig. 3.12b, she can find the similar interests, common contacts, sessions attended, and historical encounters between her and the user, and then can add that user as a contact. We provide detailed reasons for explaining what you have in common with another user; in particular, our new features are the common sessions attended and encounters, which are the physical features.

Adding a contact. After viewing another user's profile, the user can select the *Add as contact* button, and optionally send a message to introduce herself as shown in Fig. 3.13a. The user can then record how she knows this person (Fig. 3.13b). Existing systems do not incorporate specific reasons for why you want to add this person based on previous acquaintances, such as that you met this person from an event, as well as additional homophily information such as attending the same event. For example, LinkedIn only provides categories of colleagues, classmates, we've done business together, friend, and other where you provide an e-mail address as acquaintance reasons, and is general (includes free

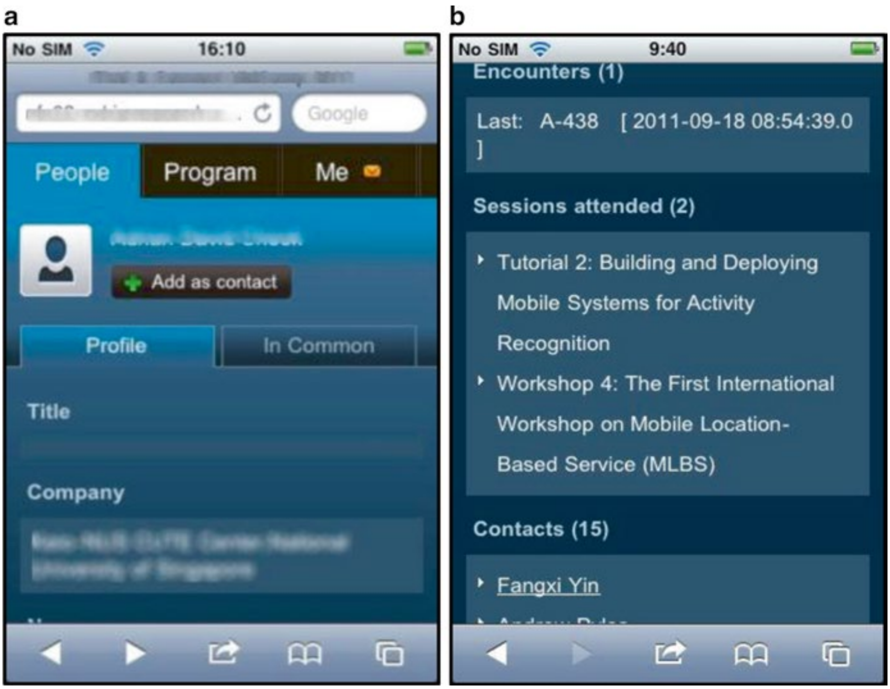


Fig. 3.12 Profile page of a user is shown in (**a**) and (**b**), showing the common research interests, common encounters, sessions attended, and common contacts between you and the user

form text). The specific reasons are important, as they provide more information to the user to help them decide whether to add this person or not (Wang et al. 2011). The acquaintance reasons shown in Fig. 3.13b are based on the survey which we administered to users before the conference, as described in our previous work (Chin et al. 2012).

2. Program
   In the *Program* shown in Fig. 3.14a, the user can see the conference schedule and session details just as in other conference systems (Barrat et al. 2010; Atzmueller et al. 2011), and then decide which sessions she wants to attend, similarly to the UIC conference (Chang et al. 2011). Here, we add the feature of recording the users who attended a particular session (as shown by selecting the *Attendees* button in the session page of Fig. 3.14b) because we know the position of each attendee. We do this because often it may be important to know whether a particular user was also in the same session as you, and for adding speakers to your contact list during their presentations so you do not forget to do it later. From this list of attendees at the session, the user can easily select any attendee and browse and connect from her profile.

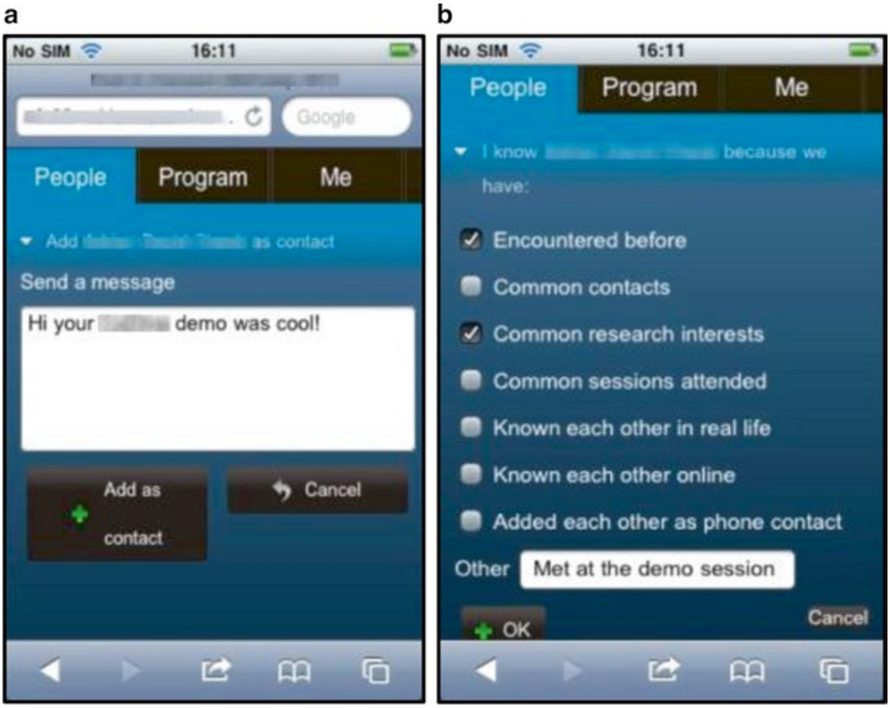


Fig. 3.13 Adding a contact (**a**) by having the user introduce herself and sending a message, and then selecting how she knows this person (**b**)

3. Me
   If the user wants to see all features personalized to her including notifications, her profile, and contacts that she has added, she goes to the *Me* page shown in Fig. 3.15a. By selecting the *Notices* tab as shown in Fig. 3.15b, the user can see the notifications of users who have added her recently (shown in *Contacts Added*), her recommended contacts from our contact recommendation system (*Recommendations*) and also *Public Notices*. She can then see their profile in Fig. 3.15c. The user can also see her contacts list (*Contacts*) and edit her own profile (*Profile*).

3.5 Analysis of Find & Connect: Offline and Online

In this section, we analyze the offline and online networks formed from the three conferences that we deployed Find & Connect to understand the user behavior for demonstrating ephemeral social networking. For user behavior, we use the following network properties to characterize each of these networks, as they are the most

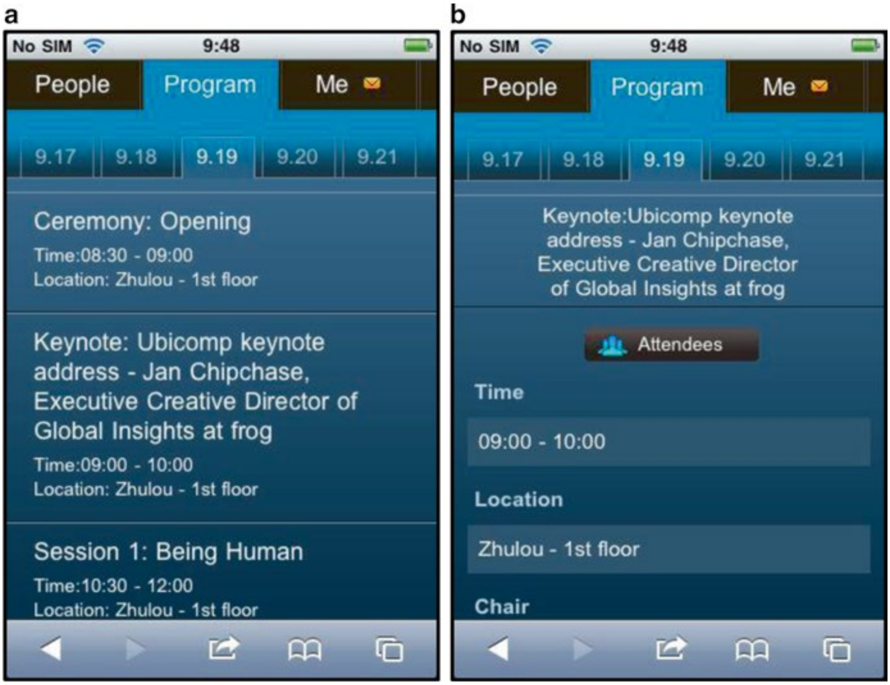


Fig. 3.14 Viewing the conference program and sessions in (**a**), then finding out who attended a session such as the keynote in (**b**)

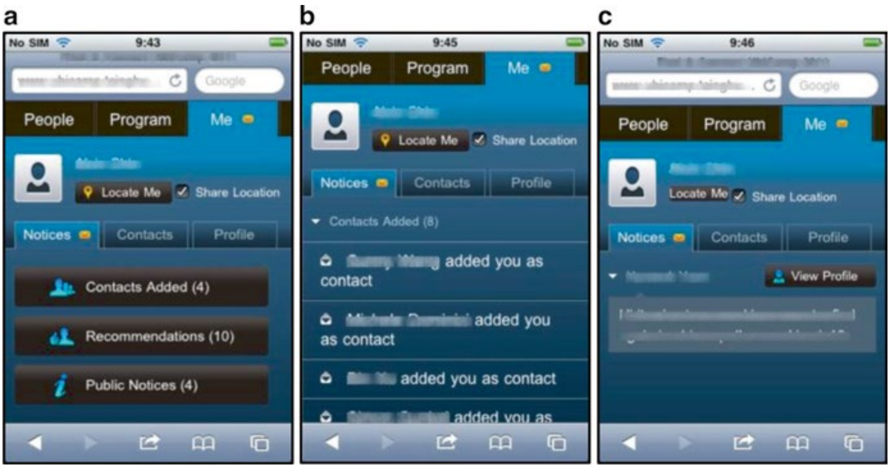


Fig. 3.15 Notifications section is shown in (**a**), where the user can see who added her as contacts from the Me page (**b**), and see the message that someone sent to her who added her as a contact (**c**), from which she can view that person's profile and decide to add that person back

frequently used (Wasserman and Faust 1994). Density is the proportion of ties in a network relative to the total number possible. Average shortest path is defined as the average number of steps along the shortest paths for all possible pairs of network nodes. Diameter is the maximum length of all shortest paths between any two connected nodes. Average clustering coefficient is a measure of the degree to which nodes in a graph tend to cluster together. We then analyze the effect that offline encounters have on forming online connections such as friends and follow.

3.5.1 Analysis at a Research Conference with Parallel Tracks (UIC 2010)

A total of 112 people registered for the trial, of which 62 users were paper authors and 50 users were non-authors. We recorded the user's position from the phone (which we lent to them to use) every 10 s, and calculated the encounters with our encounter algorithm every 5 min. In our database and log, we recorded the profile, friends, followers, recommended friends, agenda items, favorite papers, encounter information, and messages for each user. The friend recommendation algorithm is based on common interests, common sessions, encounters, messages sent and received, and common friends (Xu et al. 2011b), and was computed every 10 min. From the 86 of the 112 people surveyed, 73 of them knew at least one person at the conference; therefore, a high percentage (85 %) already knew people at the conference.

We studied the social connection networks formed from the acceptance of friend requests, follow requests, and exchanged contacts. We also studied the encounters recorded between people. The social network properties for each network are listed in Table 3.1. Users connected only with people who were registered, and the friend recommendations were only within the participants. The friend network includes also friendship that is established from the friend recommendations.

3.5.1.1 Network Properties

In the UIC conference, users followed less people (on average 3) than adding friends (average of 7.5) and exchanging contacts (7.9). The friend and exchanged contacts networks are much denser than the follow network (0.146 and 0.129 vs. 0.049 respectively), have a smaller diameter (4 for friends and follow vs. 6 for exchanged contacts), are more highly clustered with a larger clustering coefficient (0.505 and 0.462 for friends and follow vs. 0.387 for exchanged contacts), and have similar average shortest path length (2.12 for friends and follow vs. 2.68 for exchanged contacts). This means that both the friend and exchanged contacts networks are tighter and more intimate than the follow network, due to the reciprocity of friend requests and exchanged contacts, and people want to establish strong social connections with others. Users have a sense of closeness with being friends and

Table 3.1 Social network properties of the social connection and encounter networks formed in the UIC 2010, GCJK, and UbiComp 2011 conferences from Find & Connect

	Contact (UIC)	Friend (UIC)	Follow (UIC)	Encounter (UIC)	Contact (GCJK)	Follow (GCJK)	Encounter (GCJK)	Contact (UbiComp)	Encounter (UbiComp)
# of users	55	59	62	83	41	72	70	244	234
# of links	217	221	184	1,000	51	123	592	595	15,960
Average degree	7.89	7.49	2.97	24.1	2.5	1.71	8.46	2.44	68.2
Density	0.146	0.129	0.049	0.294	0.062	0.024	0.246	0.01	0.586
Diameter	4	4	6	3	6	6	4	8	3
Average clustering coefficient	0.505	0.462	0.387	0.711	0.195	0.221	0.683	0.174	0.876
Average shortest path length	2.12	2.12	2.68	1.69	2.62	2.78	2.02	3.30	1.414

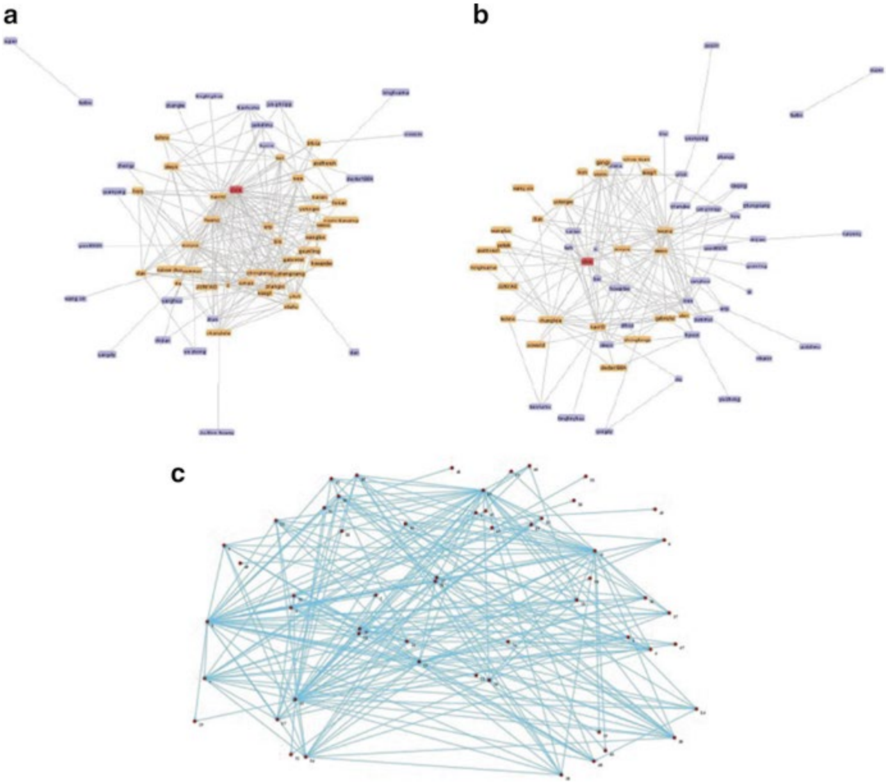


Fig. 3.16 Friend (**a**), follow (**b**), and exchanged contacts (**c**) networks for Find & Connect at the UIC 2010 conference

exchanging contacts, whereas follow is more for just casual interaction and is a loose one-way connection, unlike the strong influence of friendship. Offline encounters have the highest metrics in terms of number of users, number of links, average degree, density, diameter, and average clustering coefficient, and lowest metric in shortest path length compared to the online connections (friend, follow, and exchanged contacts). Figure 3.16 illustrates the friend, follow, and exchanged contacts networks as graphs.

The position distribution for all users follows a power-law similar to those for social networks and human behavior (Mislove et al. 2007), where only a small number of users are highly active offline while a large number of users are less active offline. Aggregated positioning of the users mapped to the conference floor show that the majority of user activity happened in the large meeting room and corridors outside the meeting rooms during the breaks.

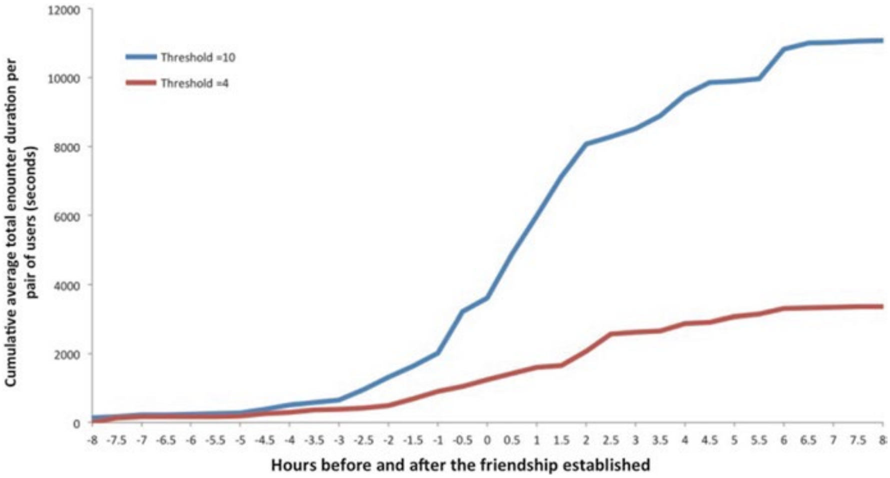


Fig. 3.17 Cumulative total encounter duration averaged per pair of users over time for different encounter distance thresholds in the UIC 2010 conference

3.5.1.2 Relationship of Offline Encounters on Online Connections

We investigate whether having greater encounter duration with a user will increase the probability of becoming a friend with that user. For the encounter duration, we select 4 m and 10 m as the threshold, with 4 m as the minimum threshold as from our earlier work (Wang et al. 2011). We choose 10 m as the maximum threshold because from the distribution of the total number of encounters defined by different encounter distance thresholds in the conference, we discover that the maximum total number of encounters occurs at 10 m, and then the encounter numbers decrease with increasing threshold. Figure 3.17 shows the cumulative total encounter duration averaged per pair of encountered users at each discrete time unit (30 min) before and after the time 0 point when the friend request was sent, with the encounter distance threshold of 4 m and 10 m.

The distributions for encounter distance thresholds of 4 m and 10 m are similar, and can be divided into four phases described below.

- *Phase I.* More than 2 h before the friendship connection is established, the cumulative average total encounter duration is very small and rises very slowly. Attendees at the beginning do not know many people, and therefore there is no online social selection that occurs here.
- *Phase II.* Around 2 h before the friend request is sent until time 0 when the friend request is sent, the cumulative average total encounter duration rises to be considerably large. As attendees begin to meet people, their encounters with that person increase as well as the encounter duration, thus causing attendees to know more about this person and add this person as a friend. Social selection on physical proximity in becoming friends becomes strong here.

- *Phase III*. Around 2 h after the friend request has been sent, the cumulative average total encounter duration continues to rise, but at a higher rate than before. After becoming friends, attendees want to know more about each other's work, so it is natural to continue to have more offline interactions, thus short-term social influence becomes strong.
- *Phase IV*. More than 2 h after time 0 when the friend request is sent, the average cumulative total encounter duration continues to rise but the curve starts to stabilize and flatten out. The frequency of encounters is less and encounter durations are smaller, and since the social connection has been established, users spend less time being physically proximate to each other because they go and meet others, causing long-term social influence to be weak.

The above behavior reflects the actual behavior of users at a conference where the intention is to meet more people.

3.5.2 Analysis at a Business Conference with a Single Track (GCJK 2011)

Our system was also deployed at a business conference for 1 day on April 13, 2011 in Beijing at the main meeting room of a conference hotel with eight Wi-Fi access points placed in various locations in the room (Xu et al. 2011a). The event was single track and divided into 19 activities. Users were encouraged to download and use the Find & Connect client throughout the event, where from a total of 779 people that were at the conference, 76 users downloaded and used the client. For the event, we relaxed the social linking to be one-way follow instead of two-way friendship, to allow more users to feel comfortable in using the social features of our software. Here, users connected only with people that were registered, and the follow recommendations were only within the participants. We recorded the exchanged contacts, followers, and encounter information for each user, similar to the UIC 2010 conference, and the networks are visualized in Fig. 3.18.

From Table 3.1, a total of 72 users are in the follow network with 123 follow links generated, 27 unique users follow others, and 66 unique users are followed. We discover 41 users in total who have at least exchanged contacts with one person in the exchanged contacts network, and 51 exchanged contact links are generated. There are 16 contacts that are exchanged in the day before the event, two are exchanged in the morning before the first activity, and 33 are exchanged during the activities. As expected, we discover that people mostly exchange contacts not during the actual main talks, but after, during lunch and leisure time. For the encounters, there are a total of 70 users who have at least one encounter, with 592 encounter links generated.

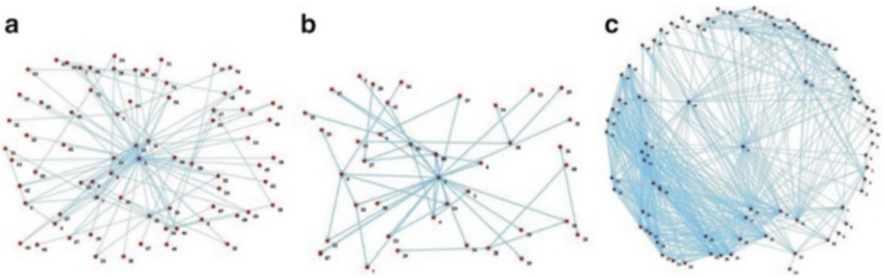


Fig. 3.18 Follow (**a**), exchanged contacts (**b**), and encounter (**c**) networks for the GCJK conference

3.5.2.1 Network Properties

As shown in Table 3.1, the encounter network has similar number of users as the follow network (70 vs. 72), but has nearly 5 times the number of links as the follow network (592 vs. 123). Also, the encounter network has almost 2 times the number of users as the exchanged contacts network (70 vs. 41), yet more than 10 times the number of links as the exchanged contacts network (592 vs. 51). In addition, the encounter network has the highest clustering coefficient (0.683) and highest density (0.245), which indicates that the encounter network is the most tightly connected, as can be seen in Fig. 3.18c. All these results show that the encounter network is better connected and more cohesive. Comparing the two online social networks, we observe that although the follow network is at a larger scale than the exchanged contacts network, the follow network is less dense than the exchanged contacts network (0.024 vs. 0.062). While the average clustering coefficient and network diameter are quite similar for follow and exchanged contacts (0.221 vs. 0.195 for average clustering coefficient of follow and exchanged contacts, and 6 for network diameter), this means that the exchanged contacts network and the follow network have a similar level of connection ties, while the follow network is larger than the exchanged contacts network.

Overall, in terms of highest density and smallest average shortest path length, the order is encounter, exchanged contacts, and follow, which intuitively makes sense because there are many encounters between people (recorded automatically), compared with exchanged contacts and follow which are manual. We can also see that the size, connectivity, and density as reflected in the visualization in Fig. 3.18 also confirm the order of the network.

3.5.2.2 Relationship of Offline Encounters to Online Connections

Similarly to UIC 2010, depending on the social relationship of a user pair, we divided all user pairs into two types; those who exchanged contacts, denoted as Exchange-Pairs, and those who established followship between each other, denoted

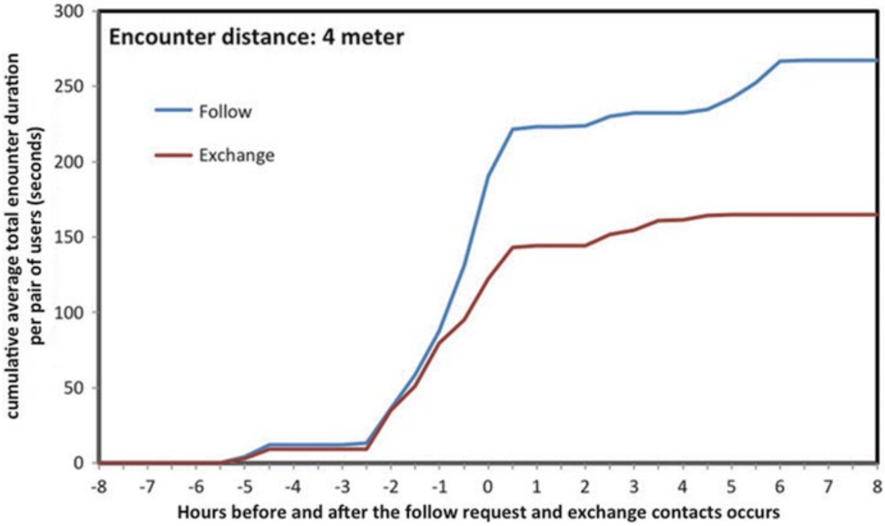


Fig. 3.19 Cumulative total encounter duration averaged per encounter over time for follow and exchanged contacts between any two attendees in the GCJK conference

as Follow-Pairs. Similarly, encounters of all user pairs can be divided into two categories: encounters of Follow-Pairs, denoted as Follow-Encounters, and encounters of Exchange-Pairs, denoted as Exchange-Encounters. For the encounter distance threshold, again we use 4 m as in UIC 2010. For each of the user encounter pairs (Follow-Encounters and Exchange-Encounters) we sum out the cumulative duration value of their encounters in discrete time intervals before and after their respective behaviors are committed, and then average them by the number of this type of user pair, which results in the Cumulative Average Total Encounter Duration for any pair of users. Figure 3.19 shows the cumulative total encounter duration averaged per pair of encountered users at each discrete time unit (30 min) before and after the time 0 point, for follow and exchanged contacts.

From Fig. 3.19, we can see three phases that emerge.

- *Phase I.* More than 2 h before any online social interaction request (exchanged contacts or follow), the cumulative average total encounter duration is very small and rises very slowly. At the beginning, people do not meet many of each other in the room because they are probably outside the room, or if they do encounter, it is short due to them passing by each other to attend to an activity.
- *Phase II.* Around 2 h before the online social interaction request until time 0 when the online social interaction request is sent, the cumulative average total encounter duration rises sharply to be considerably large. Before the exchanged contacts request or follow request, the cumulative average total encounter duration increases rapidly with time. This indicates that just before two users exchange their contacts or follow each other, they are spending more time physically proximate to each other.

- *Phase III*. After time 0 when the online social interaction request is sent (exchanged contact or follow), the average total encounter duration decreases, which causes the cumulative average total encounter duration curve to stabilize and flatten out. This indicates that after the user finishes pressing the Exchange Contacts or Follow button, the social connection has been established, and the users spend less time being physically proximate to each other. Most likely this is because after the user adds this person to her social network, then she can easily see the person's social activity updates from the All Updates page of the Find & Connect application, so there is no need to be physically proximate to that person.

The above behavior is similar to that of friend requests from the UIC 2010 conference presented earlier in Fig. 3.17, except that there is no Phase IV. This suggests that the social selection on physical proximity in exchanged contacts and follow is strong (Phase II), but that social influence is weak (Phase III). Greater physical proximity encounter duration results in an increased probability for a person to follow another one, or exchange their contacts with. However, after users have established an online social relationship, the probability of users being physically proximate to each other decreases because due to limited time, users want to meet others. At the beginning (Phase I), you do not know many people and there is no online social selection. Then, in Phase II, as you begin to meet people, you start talking with them face-to-face, and therefore your encounters with that person increase as well as the encounter duration. As you begin to know more about that person, you want to add her to your social network and therefore you choose to either follow or exchange contacts to make a record of this connection. Social selection is at play here. After that, in Phase III, you want to meet other people; therefore the frequency of encounters is less and encounter durations are smaller, because there is no need to be physically proximate to that person, since you can follow their updates in the application. Therefore, social influence is weak here.

3.5.3 Analysis at a Research Conference with a Single Track (UbiComp 2011)

We set up the Find & Connect trial at the UbiComp 2011 conference in Tsinghua University, Beijing from September 17 to 21, 2011 for all registered attendees in the conference to use. From a total of 421 registered attendees, 241 (57 %) used our system and each registered attendee wore an RFID badge to identify herself, which also acted as a conference badge. All participants used our Find & Connect web client from their own mobile device with the distribution as follows: 31.34 % of all web visits came from the Safari web browser (Apple iPhone, MacBook, or iPad), 23.85 % came from the Google Chrome browser, 22.12 % came from the Android browser (Android phone or tablet), 9.08 % came from Firefox and 8.29 % came from Internet Explorer (laptop computer). Therefore, we see that most attendees use an Apple or Android mobile device.

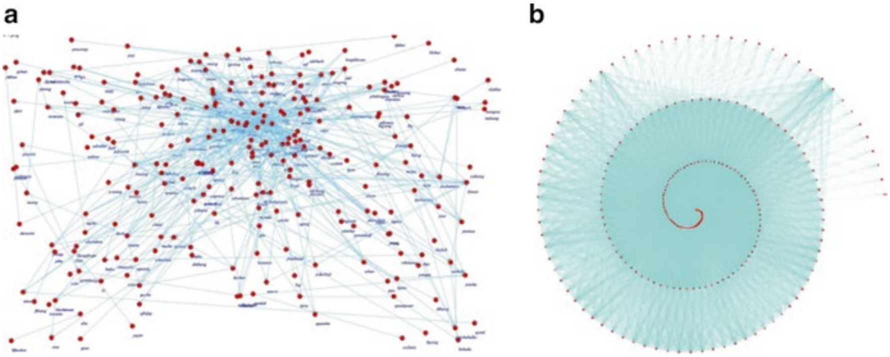


Fig. 3.20 Network graphs of the contacts (**a**) and encounters (**b**) from Find & Connect at the UbiComp 2011 conference

3.5.3.1 Usage Analysis

We used Google Analytics to track the usage of the features of Find & Connect. Participants spent an average of 11 min and 44 s in Find & Connect per visit, with an average of 16.5 pages browsed per visit, which shows that users spent a significant amount of time using the system, demonstrating that it is useful. Usage rose from the beginning of UbiComp, which was the tutorials (Sept. 17), until the first day of the conference (Sept. 19) when most people arrived, and then decreased as expected since people start to leave. From the page views, users mostly found people nearby (11.66 %), and made new connections (where the contact requests are shown in the notices page) (10.30 %), followed by login (6.27 %), viewing the conference program (4.97 %), and finding people farther away (3.29 %). This is expected because when the user logs in, the first page she sees is the finding people nearby page, which motivates her to explore who are the people close by in the room where she is.

3.5.3.2 Network Properties

We discuss the contact and encounter networks formed at UbiComp 2011, with the properties listed in Table 3.1 and their networks visualized in Fig. 3.20. We compare the results between all registered users who use Find & Connect (112) and add at least one contact or are added at least once by someone else, and the authors (62), who are 55 % of all registered users. There are a total of 571 contact requests, of which 40 % are reciprocated by the recipients, and 309 of the requests come from our contact recommendation algorithm out of a total of 15,252 contact recommendations. This results in a total of 2 % of the contact recommendations being converted into contact requests. The contact network is strongly driven by the authors, with almost all authors (55) out of all registered users (59) having added at least one contact (93 %), and the large majority of contact links come from authors. In addition, the network density, network diameter, average clustering coefficient, and average shortest path length are almost the same for authors and all registered users. The network diameter is 4, meaning that any user can be reached and added as a contact at a maximum of four hops, similar to the social influence theory of 3 degrees of separation in social networks (Cacioppo et al. 2009). With an average shortest path length of about 2, any user on average can be indirectly contacted via another direct contact, indicating that the UbiComp conference is a tight, well-connected community.

We discover that having common contacts is an important reason for adding a person as a friend/contact (similarly to online social networks) as well as having common research interests, thus validating the homophily principle and social influence theory of 3 degrees of separation. However, the majority of participants only have one to two contacts, and very few have more than ten contacts, showing that users do not add many contacts, contrary to our expectation. Perhaps users are too busy to look up the other user in Find & Connect and add as contact, or users already exchange business cards so there is no need to use Find & Connect because there is no incentive for doing so. We discovered that for those who do add others as contacts, the reasons for doing are primarily because they have encountered them before and that they know them in real life (Chin et al. 2012).

For the encounters at UbiComp 2011, there are 12,716,349 encounters and from Table 3.1, 234 users had an encounter with another attendee at the conference, with 15,960 unique encounter links established between any two people. Each user had an average of 68 encounters with other users. The encounter network is very dense (network density = 0.5861) as expected, due to the many people being co-located together during the session. The encounter network density is significantly higher than the contact network density. The network diameter in the encounter network is smaller than that in the contacts network (3 vs. 4), showing that any user can be indirectly encountered at a maximum of three hops (which is the same number as in social influence theory for establishing friends (Cacioppo et al. 2009), with an average shortest path of 1.414 hops (direct encounter). This means on average, users can directly encounter others at the UbiComp conference, without having to encounter another person first. Also as expected due to the conference environment, encounters between users are highly clustered (average clustering coefficient = 0.876) compared to the clustering in the contact network (0.462).

3.5.3.3 Relationship of Offline Encounters to Online Connections

Instead of computing the cumulative average encounter duration for a pair of contacts, which would likely show the same type of behavior as in UIC 2010 and GCJK 2011, we instead looked at the fraction of overlap between offline encounters and online contacts (i.e., how many of the offline encounters between two users at that time also had online contact requests sent by either of those two users), and the transfer amount between offline and online. Given the online and offline interaction

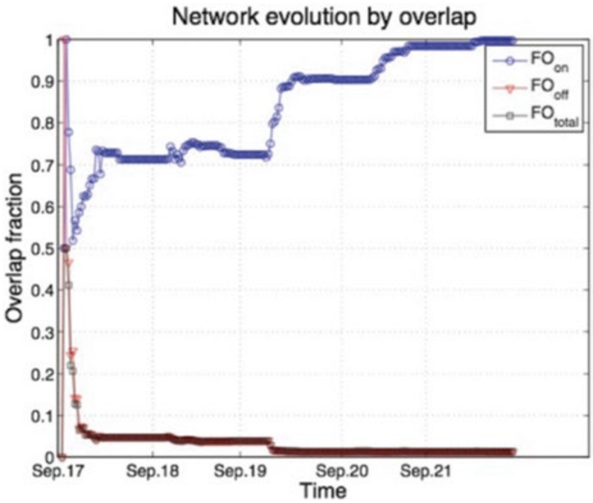


Fig. 3.21 Fraction of overlap between online contact and offline encounter networks from Find & Connect at the UbiComp 2011 conference

networks $G(V,E_{on})$ and $G(V,E_{off})$, we propose FO_{on} and FO_{off} to represent the fraction of overlap in online and offline interaction networks, respectively. Moreover, we use the Jaccard Coefficient to calculate FO_{total}, to compare the similarity of the two networks. The expressions are computed as follows:

$$FO_{on} = \frac{|E_{on} \cap E_{off}|}{|E_{on}|} \tag{3.1}$$

$$FO_{off} = \frac{|E_{on} \cap E_{off}|}{|E_{off}|} \tag{3.2}$$

$$FO_{total} = \frac{|E_{on} \cap E_{off}|}{|E_{on} \cup E_{off}|} \tag{3.3}$$

Figure 3.21 shows the fraction of overlap between online and offline interaction networks. The x-axis is the day of the conference and the y-axis is the overlap computed from the expressions above. We can see that all the overlap fraction metrics fluctuate greatly at the very beginning, and gradually reach steady state at the end of each day. The fraction of overlap in the online interaction network is always higher than 50 %, which is much higher than the offline overlap and total overlap. It means that the majority of people who have online interactions with others also had offline encounters. This is intuitive, because everyone wears the RFID badge, and are for the most part in the same room as others, since the conference is single session only.

For the transfer amount between offline and online, we define two transfer metrics: online to offline and offline to online transfer amounts. If the online interaction

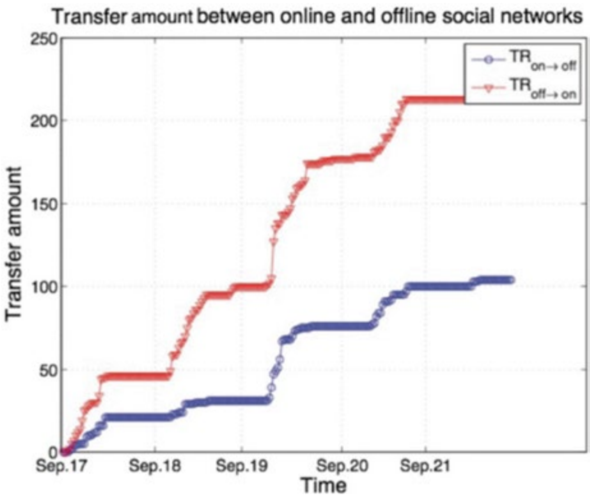


Fig. 3.22 The transfer amount between the online and offline interaction networks by cumulative amount in the UbiComp 2011 conference

happens before the offline interaction in the same node pair, it is included in the online to offline transfer amount, otherwise we count it to the offline to online transfer amount. Here $T_{i,j}{}^{on}$ and $T_{i,j}{}^{off}$ are calculated as the first time when the online and offline interactions happen between node i and node j. Here we assume $T_{i,j}{}^{on}$ and $T_{i,j}{}^{off}$ are both positive, and they are assigned with −1 only when there is no online or offline interaction between nodes i and j. Thus the O2O transfer amounts $TR_{on} \rightarrow_{off}$ and $TR_{on} \rightarrow_{off}$ can be calculated as follows:

$$Tr_{on\rightarrow off} = \sum_{i,j\in V,\, i<j,\, T_{i,i}^{off},T_{i,j}^{on}\geq 0} H\left(T_{i,j}^{off} - T_{i,j}^{on}\right) \tag{3.4}$$

$$TR_{off\rightarrow on} = \sum_{i,j\in V,\, i<j,\, T_{i,j}^{off},T_{i,j}^{on}\geq 0} H\left(T_{i,j}^{on} - T_{i,j}^{off}\right) \tag{3.5}$$

where $H(x)$ is the Heaviside step function.

Figure 3.22 shows the transfer amount between the online and offline interaction networks. Both the offline to online and online to offline transfer amounts increase during the conference, and tend to be stable after September 21 when everybody leaves the conference. The total transfer amount from offline to online is much higher than that from online to offline, and the difference increases dramatically with time. Since the encounters are detected automatically by our system but the online interactions are generated through user intervention, the offline interactions could be regarded as the background reference for users to choose their online communications. Therefore, the offline to online transfer flow holds the dominant position in this bidirectional O2O transition flow. Moreover, the offline to online transfer amount has the greatest increase on the day of September 19, which is the time of the first day of the main conference. The possible explanation is that at the end of the first conference day, people tend to review the offline encounters of the whole day displayed in our Find & Connect user interface, and then add others as contacts.

3.6 Discussion

From the three trials of Find & Connect at a research conference with parallel tracks (UIC 2010), business conference with single track (GCJK 2011) and research conference with single track (UbiComp 2011), we observe similarities and differences. Depending on the nature of the event, the properties of their social networks also change. From Table 3.1, we do find that there are differences between the three events. First, UbiComp 2011 has the most number of participants that used the Find & Connect system. This is because all participants carried an RFID conference badge and so their position was always tracked, and this conference had the most number of participants. GCJK 2011 and UIC 2010 conferences had fewer participants and used Wi-Fi, which required using a Nokia Symbian S60 phone since our Find & Connect client was only developed for Nokia Symbian S60 phones. Second, for the offline encounter network out of all the events, UbiComp 2011 has the most dense encounter network, with highest degree, highest density, and shortest average path length, due to it being a single-track conference where everyone is in the same room (for the most part) in every session during the conference. This is followed by the UIC 2010 conference and then the GCJK event. Since UIC/ATC 2010 is an international academic conference, people tend to meet more people that they did know but did not physically meet before. However, the GCJK event is an internal business meeting, so the attendees are not as interested to add others into their social network because they can be easily found in the company address book and the people already know each other. Despite that, the results are as expected, which confirms the utility and user behavior in Find & Connect.

The network properties of the contacts network in the UbiComp 2011 conference are fairly similar to the properties of the contacts network in the UIC 2010 conference and GCJK meeting. Even though the proximity technology used is different (UIC 2010 and GCJK used Wi-Fi, while UbiComp 2011 used RFID), nonetheless, for different conferences, it appears the user behavior in making contacts is the same, regardless of the conference, which for the most part is expected. However, in the UIC 2010 conference, attendees make the most number of contacts compared to GCJK and UbiComp 2011. Again, this is most likely due to the UIC attendees not knowing each other well, compared to a business meeting like GCJK or a tightly knit research community such as UbiComp 2011. In addition, UIC attendees probably are eager to try out the technology as this type of system was never deployed at UIC, whereas with UbiComp 2011, other systems such as IntelliBadge (Cox et al. 2003) were deployed before, so attendees did not find anything new behind our system, which was also reflected by attendee comments.

Interestingly, we discovered that for UIC 2010, attendees did not make use of the messaging features (where attendees could message each other, message a group, or post on LinkedIn), whereas for UbiComp 2011 (where we did not include messaging features or allow linking to online social networks such as Sina Weibo or Twitter) attendees wanted these features. The reason for the difference lies in that for UIC 2010, as indicated in our survey, not many attendees actively used online social networks such as Twitter (24 %), Facebook (48 %), or Sina Weibo (25 %, a Chinese version of Twitter) on a regular basis. On the other hand, for UbiComp 2011, most people actively tweeted on Twitter (69 %) or used Facebook (89 %) and frequently used them at least once a day (79 %). Even though both conferences were held in China and the majority of the attendees came from China, nonetheless the most influential people at UbiComp 2011 came from the US and Europe, who are the most active users of online social networking.

In addition, unsurprisingly, offline encounters do affect a user's probability of adding that person as a friend or contact. In the beginning before the friend/contact is added, social selection is at play where the encounters between two users are accumulated until the friend/contact request is initiated to add the friend/contact. Then social influence takes over after the request is accepted, as the users encounter each other more before reaching a certain time where they meet others (as demonstrated with UIC 2010 and GCJK 2011).

We also observed some interesting behavior with Find & Connect during the three events. For UIC 2010, where we encouraged users to use our system so that they could win a prize at the end of each day, many attendees used our system just for this purpose in order to become one of the top five users. Therefore, this could have biased the results. For GCJK 2011, we observe that users did not use much of the map to see where others were at, nor updated their status, but just used it to add others as contacts and exchange business cards virtually, which they found to be very useful for remembering people they had met. For UbiComp 2011, users did not use much of the add contacts feature, probably for the reason that they did not find any incentive to do so, since they could easily exchange real business cards and also add that person from that person's tweet. An interesting behavior is that we observe some users add others as contacts in our application when directly meeting with them.

Finally, we need to improve the system and the design. First, our method of calculating encounters using a central server may cause inaccurate positioning and spurious encounters, which could have been avoided if we had used other peer-to-peer technologies such as Wi-Fi, Bluetooth, or even peer-to-peer RFID to detect proximity rather than the absolute distance. Therefore, our computation of an encounter is meant to be more of a high probability that there may be an encounter rather than saying there is an actual encounter. In addition, identifying real encounters also requires the collection of other context such as audio and accelerometer, as well as the context data of others nearby who may be involved in the encounter. Therefore, more research is required in order to identify the minimum context required to reliably detect an encounter. Second, contact recommendations need to be improved, since from UbiComp 2011, only 2 % of the contact recommendations resulted in contact requests. A suggestion is to place the notification for contact recommendations near the top of the user interface of the client, so it is clearly visible. Third, as suggested by the attendees, our Find & Connect application needs to be more social, with more social features such as chatting directly with the attendees, and be integrated with online SNS, such as when adding a friend in Find & Connect, it sends a friend request to your online SNS, and also links your posts and activities directly to your online SNS.

3.7 Conclusion

In this chapter, we explained our concept of the ephemeral social network, and how it solves the problem of bridging the gap between offline and online. Our research problem is to find the relevant people and connect with them easily in a dynamic, ephemeral environment such as a conference. We want to be able to capture dynamic social networks at a particular point in time for a specific duration at a place, so that we are able to remember the event and the people that were there. Therefore, we built a system and application called Find & Connect based on this ephemeral social network concept, using social proximity and homophily in a conference context, to better help attendees find and connect with each other. We deployed Find & Connect in three different events, an academic research conference with parallel sessions (UIC 2010), a business meeting with a single session (GCJK 2011), and an academic research conference with a single session (UbiComp 2011). We have described the features and user interface of Find & Connect for each of the trials, to provide a glimpse as to how attendees use our system. Attendees enjoyed making contacts the most, although the most usage of our application was on the conference program and seeing where people were at the conference. Messaging was implemented at the UIC 2010 conference; however, it was not used frequently, whereas UbiComp 2011 attendees wanted this feature, thus demonstrating that social demographics of the attendees is important in deciding what features to include in the application.

From the trials, we conducted an analysis of O2O user behavior by quantifying the offline to online (and online to offline) networks created from the encounters, the follow, the friends, and exchanged contacts. The network properties of the contacts network in all the events are fairly similar, thus demonstrating that user behavior in making contacts is the same, regardless of the conference, which for the most part is expected. The encounters formed the most dense network, followed by the exchanged contacts, follow, and friends. From all the events, UbiComp 2011 had the most dense encounter network with the highest degree, highest density, and shortest average path length. In addition, most people used Find & Connect at the UbiComp 2011 conference because we allowed users to use their own mobile devices and separated the positioning technology from the mobile device. We discovered that offline encounters do affect the building of online contacts and friendship. Social selection or homophily is the factor for people to encounter each other through a session because of their similar interests, while social influence is the factor for people to continue to encounter each other after they have established an online connection as a friend or contact.

Even though people enjoyed using Find & Connect, the system was tailored for the specific details of a conference, and required infrastructure to set up the positioning system and the conference program. Many times, we meet and encounter people in other situations, where the encounter is not related to an organized event. This means that we will need to use peer-to-peer technologies such as NFC, Bluetooth, or RFID to detect proximity, as well as use sensors on the mobile phones to detect the type of context to determine whether an encounter is an actual encounter and the activity involved. Our future work involves generalizing Find & Connect to any event, whether it is organized or ad hoc, so that people can find and connect to each other easily. We also plan to create algorithms to detect ephemeral social networks other than sessions, and allow these ephemeral social networks to be synchronized into online social networks such as Facebook and LinkedIn. Ephemeral social networking is the next revolution in social networks, which will help to truly integrate and fuse offline activities into online.

Acknowledgments The author is grateful to the following people who contributed to this research: Xiang Zuo, Bin Xu, Chen Zhao, Xiaoguang Fan, Liang Wu, Xia Wang, Hao Wang, Hao Wang, Yongqiang Lyu, Dezhi Hong, Ying Wang, Lele Chang, Guwei Xu, and Lijun Zhu. The author also would like to acknowledge the engineers of the Find & Connect project: Fangxi Yin, Ke Zhang, Minggang Wang, Yilu Li, and Guanshang Wu. We would also like to thank all the participants who used Find & Connect in the trials and provided valuable feedback, as well as the financial and technical resources of Nokia.

References

Atzmueller, M., Benz, D., Doerfel, S., Hotho, A., Jaschke, R., Macek, B. E., et al. (2011). Enhancing social interactions at conferences. *Information Technology, 53*(3), 101–107.

Backstrom, L., Huttenlocher, D., Kleinberg, J., & Lan, X. (2006). Group formation in large social networks: membership, growth, and evolution. In *Proceedings of ACM SIGKDD, ACM* (pp. 44–54). New York.

Barrat, A., Cattuto, C., Szomszor, M., Van den Broeck, W., & Alami, H. (2010). Social dynamics in conferences: Analyses of data from the live social semantics application. In *Proceedings of the semantic Web workshop-ISWC2010* (pp. 17–33). Shanghai, China.

Bush, V. (1949). As we may think. *SIGPC Note, 1*(4), 36–44.

Cacioppo, J. T., Fowler, J. H., & Christakis, N. A. (2009). Alone in the crowd: The structure and spread of loneliness in a large social network. *Journal of Personality and Social Psychology, 97*(6), 977.

Cattuto, C., Van Den Broeck, W., Barrat, A., Colizza, V., Pinton, J. F., & Vespignani, A. (2010). Dynamics of person-to-person interactions from distributed RFID sensor networks. *PLoS One, 5*(7), e11596.

Chang, L., Chin, A., Wang, H., Zhu, L., Zhang, K., Yin, F., Wang, H., & Zhang, L. (2011). Enhancing the experience and efficiency at a conference with mobile social networking: Case study with Find & Connect. In *Proceedings of HumanCom 2011*. Lecture notes in electrical engineering, Vol. 102, (pp. 1–12). Enshi, China: Springer.

Chin, A., Xu, B., Hong, D., Wang, Y., Yin, F., Wang, X., Wang, W., & Fan, X. (2012). Using proximity and homophily to connect conference attendees in a mobile social network. In *Proceedings of the IEEE ICDCS'12 international workshop on PhoneCom 2012* (pp. 1–8). Macau, China: IEEE Press.

Choudhury, T., Borriello, G., Consolvo, S., Haehnel, D., Harrison, B., Hemingway, B., et al. (2008). The mobile sensing platform: An embedded activity recognition system. *IEEE Pervasive Computing, 7*(2), 32–41.

Cox, D., Kindratenko, V., & Pointer, D. (2003). IntelliBadge: Towards providing location-aware value-added services at academic conferences. In *Proceedings of ACM UbiComp* (pp. 264–280). Seattle, USA: ACM.

de Moraes, L. F. M., & Nunes, B. A. A. (2006). Calibration-free WLAN location system based on dynamic mapping of signal strength. In *Proceedings of the 4th ACM international workshop on Mobility management and wireless access* (pp. 92–99). New York: ACM.

Dey, A. K. (2001). Understanding and using context. *Personal Ubiquitous Computing, 5*(1), 4–7.

Dey, A. K., Abowd, G. D., & Salber, D. (2001). A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. *Human–Computer Interaction, 16*, 2–4. Taylor & Francis.

Eagle, N., & Pentland, A. (2009). Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences, 106*(36), 15274–15278.

Easley, D., & Kleinberg, J. (2010). *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge: Cambridge University Press.

Ekahau. (2011). Ekahau real-time location system. http://www.ekahau.com/products/real-time-location-system/overview.html. Accessed 6 Nov 2012.

Henricksen, K., Indulska, J., & Rakotonirainy, A. (2002). Modeling context information in pervasive computing systems. *Pervasive Computing Lecture Notes in Computer Science, 2414*, 167–180. Springer.

Huang, S., Proulx, F., & Ratti, C. (2007). iFIND: a peer-to-peer application for real-time location monitoring on the MIT campus. In *Proceedings of CUPUM*. Brazil.

Kermarrec, A. M., & Le Merrer, E. (2012). Offline social networks: stepping away from the Internet. In *Proceedings of the Fifth Workshop on Social Network Systems (SNS '12), 14*(2), New York: ACM.

Kostakos, V., & O'Neill, E. (2008). Capturing and visualising Bluetooth encounters. In *Adjunct proceedings of CHI 2008, workshop on social data analysis* (pp. 1–4). Florence, Italy: ACM.

Lane, N. D. (2012). Community-aware smartphone sensing systems. *IEEE Internet Computing, 16*(3), 60–64.

Lyu, Y., Hong, D., Wang, Y., Hou, Y., Yang, Z., Chen, Y., Shi, Y., & Chin, A, (2013). A scalable and privacy-aware location-sensing model for ephemeral social network service. In *International Journal of Distributed Sensor Networks* (11 pages). Hindawi

Mcpherson, M., Smith-Lovin, L., & Cook, J. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology, 27*, 415–444.

Mislove, A., Marcon, M., Gummadi, KP., Druschel, P., & Bhattacharjee, B. (2007). Measurement and analysis of online social networks. In Proc. of the 7th ACM SIGCOMM conference on Internet measurement (IMC '07). (pp. 29–42). New York: ACM. doi:10.1145/1298306.1298311

Ni, L., Liu, M. Y., Lau, Y. C., & Patil, A. P. (2004). LANDMARC: Indoor location sensing using active RFID. Wireless Network. (pp. 701–710). Netherlands.

Nokia. (2012). Mobile data challenge. http://research.nokia.com/page/12000. Accessed 6 Nov 2012.

Paulos, E., & Goodman, E. (2004). The familiar stranger: Anxiety, comfort, and play in public places. In *Proceedings of CHI 2004* (pp. 223–230). New York: ACM.

Vannoy, S. A., & Palvia, P. (2010). The social influence model of technology adoption. *Communications of the ACM, 53*, 149–153.

Wang, B., Bodily, J., & Gupta, S. K. S. (2004). Supporting persistent social groups in ubiquitous computing environments using context-aware ephemeral group service. In *Proceedings of PERCOM 2004* (pp. 287–296). IEEE Computer Society.

Wang, H., Chin, A., & Wang, H. (2011). Interplay between social selection and social influence on physical proximity in friendship formation. In *SRS 2011 workshop in conjunction with CSCW 2011*. Hangzhou, China.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge: Cambridge University Press.

Xu, B., Chin, A., Wang, H., Chang, L., Zhang, K., Yin, F., Wang, H., & Zhang, L., (2011). Physical proximity and online user behavior in an indoor mobile social networking application. In *Proceedings of IEEE CPSCom 2011* (pp. 273–282). Dalian, China.

Xu, B., Chin, A., Wang, H., & Wang, H. (2011). Using physical context in a mobile social networking application for improving friend recommendations. In *Proceedings of the PhoneCom 2011 workshop, IEEE CPSCom 2011* (pp. 602–609). Dalian, China.

Zhu, L., Chin, A., Zhang, K., Xu, W., Wang, H., & Zhang, L. (2010). Managing workplace resources in office environments through ephemeral social networks. In *Proceedings of UIC 2010* (pp. 665–679). Berlin: Springer-Verlag.

Zhuang, H., Chin, A., Wu, S., Wang, W., Wang, X., & Tang, J. (2012). Inferring geographic coincidence in ephemeral social networks. Machine Learning and Knowledge Discovery in Databases: Lecture Notes in Computer Science (Proceeding European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD)) 7524, (pp. 613–628). Bristol, UK.

Zuo, X., Chin, A., Xu, B., Hong, D., Wang, Y., Fan, X., & Wang, X. (2012). Connecting people from offline to online in a mobile social application. In *Proceedings of the IEEE international conference on cyber, physical and social computing (CPSCom'12)* (pp. 277–284). Besancon, France.

Chapter 4
Social Behavior in Mobile Social Networks: Characterizing Links, Roles, and Communities

Martin Atzmueller

Abstract Mobile social networks are enabled by the emergence of mobile and ubiquitous applications, providing social networking and social media functionalities in diverse contexts. This chapter focuses on social behavior in mobile social networks: We first discuss different aspects of mobile social networks. After that, we briefly describe exemplary systems. Finally, we summarize recent real-world analysis results, especially focusing on links and contacts between individuals, characterization of their roles, and dynamics of communities in mobile social networks.

4.1 Introduction

Mobile social networks are usually created using pervasive and ubiquitous applications. In this way, mobile social systems help to bridge the gap between physical and online worlds, utilizing context awareness, pervasive computing, distributed computing, and sensor networks.

Mobile devices, such as smartphones or RFID tags coupled with appropriate ubiquitous and social applications, enable an integrated approach for both physical and digital social interactions. In that way, mobile social networking specifically includes networking options that are only possible on mobile and ubiquitous devices. It is important that mobile social networks are not restricted to networks provided by smartphone technology. For example, an RFID sensor network is one prominent kind of mobile social network that enables the capture, analysis, and processing of offline interactions in order to support social networking and provide sophisticated social services in ubiquitous environments.

M. Atzmueller (✉)
Knowledge and Data Engineering Group, University of Kassel, Kassel, Germany
e-mail: atzmueller@cs.uni-kassel.de

A. Chin and D. Zhang (eds.), *Mobile Social Networking: An Innovative Approach*, Computational Social Sciences, DOI 10.1007/978-1-4614-8579-7_4,

This chapter describes social behavior in such mobile social networks. We focus on different aspects in mobile social networks, and discuss its implementation. With regard to this, we also briefly describe exemplary systems in real-world application contexts: We briefly discuss the Conferator (Atzmueller et al. 2011a) and MyGroup (Atzmueller et al. 2012a) systems for supporting social networking in conferencing and workgroup contexts. Both systems apply active RFID technology that allows us to detect real-world face-to-face contacts, i.e., offline contacts in addition to contacts in (online) social networks. In this context, we summarize a number of exemplary analysis results of social behavior in mobile social networks, focusing on the characterization of links, roles, and communities.

The rest of the chapter is structured as follows. The next section provides an overview on social behavior in mobile social networks, focusing both on the social and the behavioral aspect. After that, we briefly summarize exemplary systems enabling mobile social networks. The following section discusses analysis results in the application context of the Conferator and MyGroup systems. Finally, we conclude with a summary and outlook on future work.

4.2 Overview

In this section, we briefly provide an overview on mobile social networks (and corresponding) systems, and sketch analysis options for social behavior in such networks. A detailed discussion of these aspects is provided in the subsequent sections. First, we start with a brief introduction to social networks and social media.

4.2.1 Social Networks and Social Media

Social networks and social media are key concepts for the analysis of mobile social networks. From a technical perspective, social networks typically consist of a set of actors (nodes) connected by links (edges). Usually, these relations are modeled as a graph. For example, we could consider friendship relationships or the "Follower" relationship in Twitter. Extensions consider multiple types of connections (links or relations) modeling the network as an (extended) hypergraph. From an analytic point of view, we can then analyze the networks using a number of techniques from the area of social network analysis (SNA) (Wasserman 1994).

With regard to social media, Kaplan and Haenlein define social media as a set of internet-based applications that apply principles of the Web 2.0 (Kaplan and Haenlein 2010). In this context, user-generated content plays a prominent role, that is, content being produced and shared by users online. In this chapter, we adopt a similarly intuitive definition, and consider social media as Internet-based systems and services in the ubiquitous Web that use and provide all kinds of social data of human interaction and communication. Both social media and mobile devices play an important role for connecting people through the social media artifacts. Therefore, we especially focus on social media that are being enabled by social mobile devices.

This also includes data from sensor networks, as long as the data is created by real users. Essentially, mobile social networks are enabled by mobile devices; however, we do not constrain these to just networks on mobile, but include special kinds of networks and applications that are specifically created by mobile devices.

4.2.2 Mobile Social Networks

In this chapter, we aim to sketch a more comprehensive view on mobile social networks, in contrast to earlier work (e.g., Humphreys 2008) that mainly considered the social network on the mobile, i.e., the mobile device as just another technical device for accessing the social network. We specifically include networking options that are only possible on mobile and ubiquitous devices. This includes, for example, sensor networks, and ubiquitous computational systems enabled by RFID sensor networks as one prominent kind of mobile social network.

For the characteristics of mobile social networks, we therefore consider different devices which enable systems utilizing such networks. Basically, these include all devices for connecting people, for example, mobile phones, and RFID or Bluetooth tags. Specific studies for these include those using cell phones (Miluzzo et al. 2008) and Bluetooth-enabled devices and sociometric badges (cf., Chin et al. 2012; Eagle and Pentland 2006; Zhuang et al. 2012), as well as passive and active RFID tags, e.g., the proximity tags developed by the SocioPatterns consortium (http://www.sociopatterns.org), (e.g., Alani et al. 2009; Barrat et al. 2010). We describe these in more detail below; we especially focus on these sensors because they are utilized in the exemplary Conferator and MyGroup systems also described below. These systems provide attractive technical options for analyzing social behavior in mobile social networks: compared to the other discussed technologies, the active proximity tags are able to detect "real" contacts, that is, real face-to-face contacts between individuals. If the tags are worn on the front of the body, proximity of two tags indicates a face-to-face contact of the actors. In the analysis section below, we will consider exemplary analysis results in the context of such networks. In contrast, many technologies implementing mobile social networks, e.g., based on Bluetooth, mainly consider the "co-location" of the contacts. In an offline scenario, a network can then be created, for example, using frequent encounters – however, only as a proxy for contacts.

4.2.3 Analysis of Social Behavior in Mobile Social Networks

Mobile social networks are often created during certain events, for example, during conferences, at work, or other group-based activities. During a conference, for example, social contacts form an essential part of the experience of the participants. In general, the term "networking" is used for describing the inherent processes in such interactions. Typically, there are different (implicit and explicit) communities present at such events, defined according to interests or membership to certain tracks or special interest groups. In order to enhance the participating experience, ubiquitous computing approaches, e.g., based on RFID-tokens, can provide dynamic adaptation options. Mobile social networks can then be utilized, for example, for recommendations, or explanations of the current user context.

In the following, we focus on the analysis of key actors, roles, and communities. For the analysis, we can consider the static and dynamic behavior in mobile social networks. For the static behavior, we can simply analyze the network structure and derive basic statistics and characteristics from these. As an overall summary, such an analysis usually provides first insights into the link structure and behavioral characteristics. Different network types, for example, can then be distinguished by the diameter, contact distribution, degree distribution, or mean connectivity of the contained nodes. Identifying different "roles" of nodes, and finding so called "key actors", has attracted a lot of attention, ranging from different measures of centrality (cf., Brandes and Erlebach 2005) to the exploration of topological graph properties (Gaertler 2004; Wongchokprasitti et al. 2010) or structural neighborhood similarities (Lerner 2005). For an analysis of the dynamic behavior, we can consider how the behavior changes and evolves during the event, i.e., the life cycle of the mobile social network. We can analyze time-based slices of the networks, for example, that is, snapshots of the network constructed using different time-thresholds for constructing the edges.

The analysis of communities intuitively considers densely connected subgroups of actors, represented as nodes in the social network. While there exist different definitions of communities, the above definition includes the core of most – focusing on the density of the connections within the community (Atzmueller and Mitzlaff 2011; Fortunato and Castellano 2007; Girvan and Newman 2002; Lancichinetti and Fortunato 2009; Newman 2004). In general, usually not only the density within the community is assessed, but the connection density of the community is compared to the density of the rest of the network (Newman 2004). Then, cuts between communities are established in such a way as to maximize the community evaluation function. For community-based roles, a prominent metric measures how much a node connects different communities (cf., Scripps et al. 2007; Chou and Suzuki 2010). It can be based on initially given community structures or on a probabilistic model.

Furthermore, the analysis of links within contact networks can be analyzed on a static or dynamic level, for example, considering link prediction (Liben-Nowell and Kleinberg 2003; Murata and Moriyasu 2007) on new or recurring contacts, their duration, and points in time (Scholz et al. 2012). This includes both static and dynamic properties of the social (contact) network.

4.3 Systems Enabling Mobile Social Networks

With regard to the analysis of mobile social networks, there have been several approaches. For example, Hui et al. (2005) describe an application using Bluetooth-based modules for collecting mobility patterns of conference participants.

Furthermore, Eagle and Pentland (2006) present an approach for collecting proximity and location information using Bluetooth-enabled mobile phones. Cattuto et al. (2010) applied proximity sensing in the Sociopatterns project. Isella et al. (2011a) conducted further experiments on a variety of contact networks obtained via RFID technology. In addition, Alani and colleagues, e.g., (Alani et al. 2009), also added contact information from social online networks in the live social semantics experiments. In Barrat et al. (2010), the authors analyze social dynamics of conferences focusing on the social activity of conference participants in those experiments. They analyze, for example, their activity in social Web platforms such as Facebook, Twitter, and other social media, together with status and their research seniority. These experiments have also extended their focus from conferences to schools (Stehle et al. 2011) and hospitals (Isella et al. 2011b) using the SocioPatterns proximity tags. Chin et al. present mechanisms using WLAN and RFID positioning, e.g., (Chin et al. 2012; Zhuang et al. 2012) for their Find & Connect system. The system aims at connecting conference participants using location-based proximity as a proxy for face-to-face encounters and online social networks. Encounters based on the location can be estimated for constructing the mobile social networks, which enable flexible support of participants at conferences using localization and online network techniques.

The Conferator system (http://www.conferator.org), a social conference guidance system, uses the same technical basis (RFID tokens with proximity sensing) as the Sociopatterns project for connecting people. Essentially, Conferator and MyGroup – a similar system for working groups – are two applications (Atzmueller et al. 2012a; Atzmueller et al. 2011a) using social media for collective intelligence. Conferator offers conference participants the option to organize and manage their social contacts during conferences. For this purpose, active SocioPattern RFID tags are applied that allow to localize participants and to collect their face-to-face contacts. For these, the system allows the setup of a complete profile, the management of their own social contacts, social networking to other participants, and the management and personalization of the conference schedule. A similar application, MyGroup (Atzmueller et al. 2012), allows the support of social communication in the context of working groups using social-interaction-awareness by utilizing the same technology.

Conferator has been applied at a number of conferences, for example at the LWA 2010, LWA 2011, and LWA 2012 conferences of the German association of computer science, and at the ACM Hypertext 2011 conference. MyGroup is continuously running in the Knowledge and Data Engineering research group at the University of Kassel, and has also been applied at a number of different events, e.g., at a software development code camp to enhance social interactions and communication in developing software. The applied data mining methods are based on the community mining and key actor analysis techniques described above.

Using the SocioPatterns proximity tags, the system is able to detect proximity contacts. Due to the technical ability of the tags to detect contacts between each other, Conferator detects face-to-face contacts between actors, if the tags are worn on the front of the body. Therefore, in contrast to other approaches that apply passive RFID, Bluetooth, or WLAN, the active proximity tags allow a relatively accurate detection of the location of participants and of their face-to-face contacts for constructing mobile social networks.

4.4 Analysis of Mobile Social Behavior

This section focuses on social behavior in mobile social networks, and describes analysis options on different levels: It summarizes specific methods and techniques that can be used to characterize user and social behavior, for role mining and characterization, as well as dynamic community identification.

In the following, we primarily focus on the analysis of social behavior in mobile social networks, and include exemplary results obtained by the analysis of networks of the Conferator and MyGroup systems at different events.

For the static analysis, we mainly include common methods of social network analysis, providing overview statistics and centrality measures in the contact graph of the respective mobile networks. Furthermore, these can also be extended to a time-based analysis, considering different contact lengths of these networks. This bridges the path to more dynamic analysis, concerning the evolution of the networks. We consider individual behavior with regard to role characterization and provide examples of role patterns, before we describe an analysis of social behavior using link prediction methods. Furthermore, we discuss community mining, and provide examples of different community characterizations.

4.4.1 Characterization of Key Players

In social networks, key players are actors that are important for the network in terms of connectivity, number of contacts, and the paths that are passing through the corresponding node. There is a broad range of applications concerning the identification and characterization of key players. It can be applied for prestige and reputation mining, for identifying hubs in the network and for social monitoring. The assessment can happen on different layers. We can consider the network as a whole for discovering individual roles (Wasserman 1994). Additionally, we can consider roles in specific communities (Gaertler 2004). Finally, descriptive pattern mining and characterization can be applied for both layers, (e.g., Atzmueller et al. 2011b; Atzmueller and Mitzlaff 2011; Atzmueller and Puppe 2008; Macek et al. 2012).

For the first case, standard social network analysis (SNA) methods can be applied for analyzing the complete network structure (Wasserman 1994), for example for determining the mean path length between nodes, or for discovering the diameter of the network. Additionally, on the level of the whole network we can determine different centrality measures, in order to identify important nodes or hubs (Atzmueller et al. 2012b; Wasserman 1994). Examples are given by the degree centrality as the

Table 4.1 Exemplary results of characterizing non-organizers and PhD students at the LWA 2010 conference (Atzmueller et al. 2012b) using different centrality measures. The table shows the lift of the pattern comparing the fraction of non-organizers and PhD students covered by the pattern p compared to the fraction of the whole dataset, the size of the pattern extension (number of described non-organizers/PhD students), and the description itself. *Clo*, *eig*, *deg*, *bet*, and *str* denote the closeness, eigenvector, degree, betweenness, and weighted degree centralities, respectively with the values *low, medium, high*

Target	#	Lift	p	Size	Description
Non-organizer	1	1.06	0.88	51	*clo* = {*low*; *medium*}
	2	1.05	0.87	61	*eig** = {*low*; *medium*}
	3	1.04	0.86	59	*deg* = {*low*; *medium*}
	4	1.10	0.92	12	*clo* = {*low*; *medium*} AND *deg* = {*high*; *medium*}
	5	1.12	0.93	30	*clo* = {*high*; *low*} AND *eig** = {*low*; *medium*}
PhD-student	1	1.07	0.54	59	*bet* = {*high*; *low*}
	2	1.07	0.54	48	*str* = {*high*; *low*}
	3	1.14	0.58	26	*deg* = *high*
	4	1.31	0.67	12	*bet* = {*high*; *low*} AND *eig** = *high*
	5	1.38	0.70	20	*deg* = *high* AND *bet* = {*high*; *low*}
	6	1.58	0.80	10	*deg* = *high* AND *bet* = {*high*; *low*} AND *eig** = {*high*; *low*}

number of connections to the neighbors of the node, the betweenness centrality as the number of shortest paths of all node pairs that go through a specific node, or the closeness centrality that considers the length of these shortest paths. For the degree and betweenness centralities, high values indicate a higher importance, while the reverse is true for the closeness centrality (Wasserman 1994).

As an example, we can consider the contact network of the LWA 2010 conference that was obtained using the Conferator system (Atzmueller et al. 2012a, 2011a). In the network, we can observe higher centrality values for professors compared to other groups, for example, post-docs and students. An exception is given by the betweenness centrality, for which the post-docs showed the highest scores (cf. Atzmueller et al. 2012b). In these results, these findings indicate the potential influence in hierarchical relationships, for which the post-docs are contained in many shortest paths between the participants, and thus seem to have a very important function as bridges.

For a more detailed analysis, Table 4.1 summarizes exemplary analysis results focusing on the characterization of different status roles given a set of network properties (cf. Atzmueller et al. 2012b). Here, we focus on the application of descriptive pattern mining (Atzmueller et al. 2011b, 2012), that is, in identifying certain subgroups that are exceptional with respect to a certain property of interest. In our case, this property is given by a high share of a certain role, for example, “non-organizer” and “PhD student”. We observe, for example, that extreme values, i.e., sets of high and low centrality values, are also very significant for distinguishing PhD students. As expected, the combination with other strong influence factors increases the precision of the patterns (indicated by the lift parameter). The most descriptive factors for the non-organizer role are given by the closeness, eigenvalue centrality, and the

Table 4.2 Exemplary role influence patterns (Scholz et al. 2012) for the Hypertext 2011 conference measuring the mean duration of recurring contacts. The column "Lift" shows the relative increase of the mean of the subgroup described by the factors in the "Description" column compared to the mean of all attendees

#	Lift	Mean	Size	Description
1	2.10	5,944.17	6	PhD AND low affiliation
2	1.52	4,297.15	26	Low affiliation
3	1.09	3,089.00	6	Session chair AND professor
4	1.08	3,038.67	21	PhD candidate
5	1.06	3,003.93	14	PhD
6	0.87	2,461.25	8	Session chair
7	0.82	2,326.18	11	Professor

degree, for which lower values than those of the organizers are measured. However, if we consider combinations of factors, we observe, that there are subgroups with regard to the non-organizer role, for which extreme values, e.g., the closeness together with the eigenvalue centrality, yield a significant increase in characterization power.

4.4.2 Characterizing Roles and Links

A related analysis to the identification of key players and according "status" roles is the characteristic behavior when establishing links. This is an analysis relating to link prediction, i.e., the prediction of new links between nodes in a network. Fundamental work in this area has been done by Liben-Nowell and Kleinberg (2003), considering standard network proximity measures, which has also been extended to weighted variants (Murata and Moriyasu 2007).

For analyzing influence factors for link prediction in mobile social networks, we consider the prediction of new and recurring links. In the following, we specifically consider human contact networks obtained using the Conferator system, and summarize results concerning the analysis of contacts patterns in those social networks, and their underlying mechanisms (cf. Scholz et al. 2012). Homophily (Rosvall and Bergstrom 2007), for example, is a classic topic of social network analysis. Similar to the RFID-based setting, Eagle and Pentland (2006) and Zhou et al. (2009) presented an analysis of proximity information collected by Bluetooth-based devices, similar to Xu et al. (2011), relating this to online social networks.

Table 4.2 shows exemplary results (cf. Scholz et al. 2012) of a descriptive link prediction analysis at the Hypertext 2011 conference in the form of characteristic patterns (Atzmueller et al. 2009; Atzmueller and Puppe 2008) that describe certain subgroups of the conference participants. The table shows combinations of influence/role factors that are significantly correlated with the duration of recurring contacts at the conference. In the table, we observe, for example, that people with a low affiliation, i.e., participants that are new to the conference, are still very active after the first day, i.e., on day 2 or 3 of the conference. Furthermore, being a session chair and being a professor increases the mean duration of contacts by 10 %, while considering the single factors separately decreases the duration (by 13 % and by 18 % respectively).

4.4.3 Characterization of Communities and Roles

Community mining (Lancichinetti and Fortunato 2009; Leskovec et al. 2010; Newman 2004; Newman and Girvan 2004) in social networks aims at discovering and analyzing (cohesive) subgroups, clusters, or communities that are "densely" connected with each other in a network. Standard techniques for the mining of communities include graph-based approaches (Girvan and Newman 2002), clustering according to features of the nodes, or pattern-mining techniques for optimizing a community evaluation function (Atzmueller and Mitzlaff 2011). The core idea of the evaluation function is to apply an objective evaluation criterion; for example, the number of connections within the community compared to the statistically "expected" number based on all available connections in the network, and to prefer those communities that optimize the evaluation function locally (Atzmueller and Mitzlaff 2011), or globally (e.g., Girvan and Newman 2002; Lancichinetti and Fortunato 2009).

In the context of mobile social networks, the discovered communities can then be applied, for example, for recommendations (Boratto et al. 2009; Gargi et al. 2011; Farzan and Brusilovsky 2007) or for personalization of intelligent systems (Boratto et al. 2009; Farzan and Brusilovsky 2007; Wongchokprasitti et al. 2010). In such cases, the community assignment can be made explicit, or the information about other members of the community can be used implicitly for the adaptation of the application. We focus on patterns characterizing specific community-oriented roles (Scripps et al. 2007), that is, how well certain participants are able to bridge communities, for example. Before that, we first focus on a time-based analysis considering how the social behavior measured by the length of contacts evolves during the LWA 2010 conference, using data obtained by the Conferator system. Both techniques can be applied in a recommendation setting, for example, for recommending people in interest-based communities, or for identifying influential bridges between different communities – based on the mined patterns.

For the time-based analysis, we utilized the MOSES algorithm (McDaid and Hurley 2010), and considered different time slices corresponding to different minimal contact lengths. Summarizing the setting in Atzmueller et al. (2012), we were able to use four different communities corresponding to special interest groups present at the conference, as a ground truth of interest-based communities. For the analysis, we considered their interactions with regard to different minimal conversation lengths (i.e., contact length between participants for their face-to-face conversations). In the following, we discuss exemplary results in this setting. Concerning our mobile social network of conference participants, the community results in

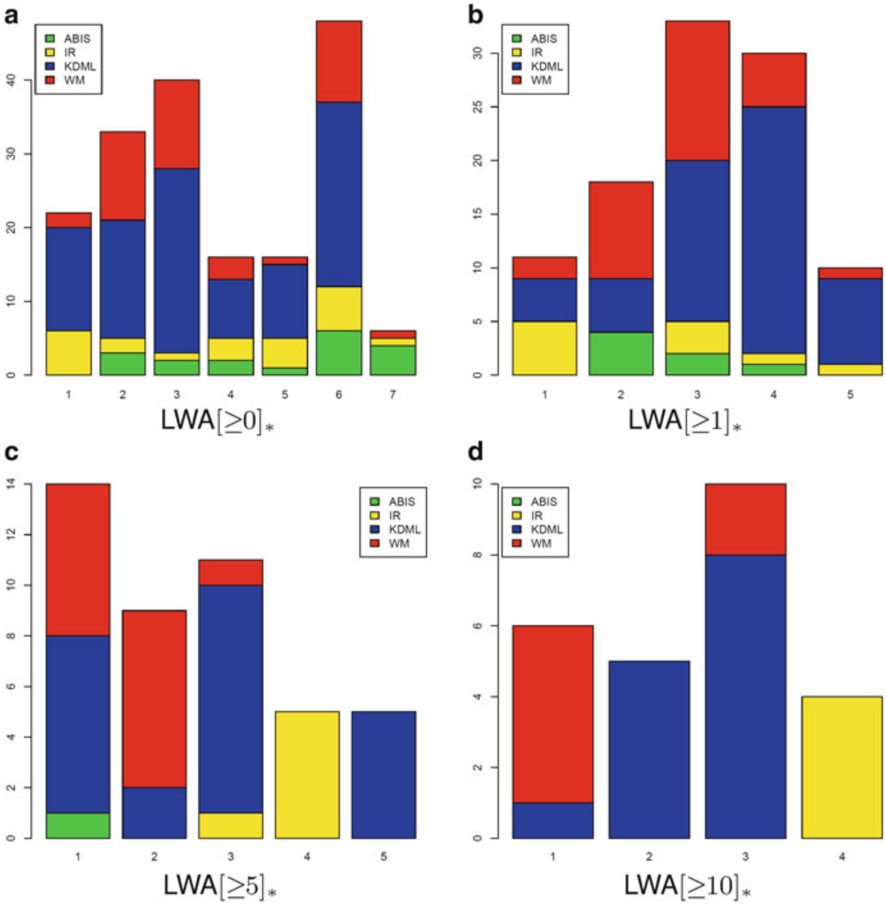


Fig. 4.1 Exemplary community detection results (Atzmueller et al. 2012), for different minimal conversation lengths (in minutes), using the MOSES algorithm (McDaid and Hurley 2010). The different communities are colored according to their special interest track distributions

Fig. 4.1 suggest that the communities tend to focus more on the special interest groups with an increasing minimum conversation length threshold. The communities start with a mixture of different interest groups, but concentrate more and more on special sub-communities. These findings suggest the trend that more specialized sub-communities of the special interest groups are mined. This is especially significant for higher minimal conversation length thresholds, see Fig. 4.1. In summary, this community-oriented analysis suggests, that participants actually tend to interact more frequently with members of their own special interest group for longer conversations.

Role mining with regard to communities mainly considers the relations between the communities for a specific actor. Scripps et al. (2007) present a method for

Table 4.3 Exemplary results (Macek et al. 2012) of pattern analysis at the Hypertext 2012 conference for describing individual roles using pattern mining and subgroup discovery. The patterns are described by (combinations) of properties of the participants, e.g., being session chairs or having a strong affiliation to the Hypertext conference

Min. contact length: 180 sec					
#	Target	Lift	Share	Size	Pattern
1	Ambassador	1.47	0.63	8	SessionChair = true
2	Ambassador	0.98	0.42	12	Affiliation = strong
3	Bridge	1.05	0.29	7	Country = Netherlands
4	Bridge	1.83	0.50	6	SessionChair = true AND Affiliation = strong
5	Bridge	1.53	0.42	12	Affiliation = strong
6	Bridge	1.38	0.37	8	SessionChair = true

assessing roles considering the membership in the communities and the potential to bridge or to connect different communities. In this way, different actor profiles concerning their centrality prestige and their community importance can be derived. Chou and Suzuki (2010) present a similar method considering a set of given communities for such a community-oriented analysis.

While the above methods mainly focus on the network and community structure, a simple characterization or description is usually not provided by standard methods for role and community mining. To this end, the characterization of actors and their roles is provided by descriptive pattern mining techniques (e.g., Atzmueller et al. 2011a, b; Atzmueller and Lemmerich 2012) that utilize different centrality measures and allow the characterization of role-specific nodes given different centrality measures. In this way, roles can get an intuitive interpretation given the network characteristics as depicted in Table 4.3, where we see exemplary results from the Hypertext 2011 conference (Macek et al. 2012). The table describes the community-oriented roles by different conference-oriented properties. In both cases, the VIKAMINE tool (Atzmueller and Lemmerich 2012) for pattern mining and subgroup analytics was applied for obtaining these characteristic patterns. They can then be inspected in a detailed context, for example, or they can be described by representative instance prototypes (cf., Atzmueller and Puppe 2008).

4.5 Conclusions

Pervasive and ubiquitous applications, in addition to the advances in technology, have created a number of options for mobile social networking. This chapter has presented different aspects of mobile social networks, and the emerging social behavior in those, especially focusing on roles, the creation of links, and social behavior in communities. We have summarized different exemplary systems, and discussed related analysis results. Then, ultimately, the analysis and understanding of such phenomena bridges the gap to collective intelligence (Malone et al. 2009;

Mitchell 2009), and provides different options for making use of the collected information (and derived knowledge) for integrating it into smart ubiquitous and social systems.

References

Alani, H., Szomszor, M., Cattuto, C., den Broeck, W. V., Correndo, G., & Barrat, A. (2009). Live social semantics. In *Proceedings of international semantic Web conference* (pp. 698–714). LNCS 6497, Springer Verlag, Heidelberg, Germany.

Atzmueller, M., & Lemmerich, F. (2012). VIKAMINE – Open-source subgroup discovery, pattern mining, and analytics. In *Proceedings ECML/PKDD 2012: European conference on machine learning and principles and practice of knowledge discovery in databases*. Lecture notes in computer science, Vol. 7524, pp. 842–845. Heidelberg: Springer

Atzmueller, M., & Mitzlaff, F. (2011). Efficient descriptive community mining. In *Proceedings of the 24th international FLAIRS conference* (pp. 459–464). Palo Alto, CA, USA: AAAI Press.

Atzmueller, M., & Puppe, F. (2008). A case-based approach for characterization and analysis of subgroup patterns. *Journal of Applied Intelligence, 28*(3), 210–221.

Atzmueller, M., Lemmerich, F., Krause, B., & Hotho, A. (2009). Who are the spammers? Understandable local patterns for concept description. In *Proceedings of the 7th conference on computer methods and systems* (pp. 151–156). University of Krakow, Krakow, Poland.

Atzmueller, M., Benz, D., Doerfel, S., Hotho, A., Jäschke, R., Macek, B. E., et al. (2011a). Enhancing social interactions at conferences. *it – Information Technology, 53*(3), 101–107.

Atzmueller, M., Benz, D., Hotho, A., & Stumme, G. (2011b). Towards mining semantic maturity in social bookmarking systems. In *Proceedings of the 4th international workshop on social data on the Web.*

Atzmueller, M., Becker, M., Doerfel, S., Kibanov, M., Hotho, A., Macek, BE., Mitzlaff, F., Mueller, J., Scholz, C., & Stumme, G. (2012a). Ubicon: Observing social and physical activities. In *Proceedings of the 4th IEEE international conference on cyber, physical and social computing (CPSCom)* (pp. 317–324). Washington, DC, USA: IEEE Press.

Atzmueller, M., Doerfel, S., Mitzlaff, F., Hotho, A., & Stumme, G. (2012b). Face-to-face contacts at a conference: Dynamics of communities and roles. In *Modeling and mining ubiquitous social media*. Lecture notes in computer science, Vol. 7472, pp. 21–39. Heidelberg: Springer.

Barrat, A., Cattuto, C., Szomszor, M., den Broeck, W. V., & Alani, H. (2010). Social dynamics in conferences: Analyses of data from the live social semantics application. In *Proceedings of international semantic Web conference*. Lecture notes in computer science, Vol. 6497, pp. 17–33. Heidelberg, Germany: Springer

Boratto, L., Chessa, A., Agelli, M., & Clemente, M. L. (2009). Group recommendation with automatic identification of user communities. In *Proceedings IEEE/WIC/ACM international joint conferences on Web intelligence and intelligent agent technologies,* Vol. 3, pp. 547–550. Milan, Italy.

Brandes, U., & Erlebach, T. (Eds.) (2005) Network analysis: Methodological foundations. In U. Brandes, T. Erlebach (Eds.) *Network analysis*. Lecture notes in computer science, Vol. 3418. Heidelberg: Springer.

Cattuto, C., den Broeck, W. V., Barrat, A., Colizza, V., Pinton, J. F., & Vespignani, A. (2010). Dynamics of person-to-person interactions from distributed RFID sensor networks. *PLoS One, 5*(7), 1–9.

Chin, A., Xu, B., Wang, H., Chang, L., Zhu, L., & Wang, H. (2012). Connecting people through physical proximity and physical resources at a conference. To be published in ACM Transactions on Intelligent Systems and Technology.

Chou, B. H., & Suzuki, E. (2010). Discovering community-oriented roles of nodes in a social network. In *Data warehousing and knowledge discovery*. Lecture notes in computer science, Vol. 6263, pp. 52–64. Heidelberg: Springer.

Eagle, N., & Pentland, A. S. (2006). Reality mining: Sensing complex social systems. *Personal Ubiquitous Computing, 10*(4), 255–268.
Farzan, R., & Brusilovsky, P. (2007). Community-based conference navigator. Socium: Adaptation and Personalisation in Social Systems: Groups, Teams, Communities. Workshop/UM 2007.
Fortunato, S., & Castellano, C. (2007). *Community structure in graphs*. Encyclopedia of complexity and system science. Springer. arXiv:0712.2716, 42 pages.
Gaertler, M. (2004). Clustering. In *Network analysis: Methodological foundations* (pp. 178–215). Berlin: Springer.
Gargi, U., Lu, W., Mirrokni, V., & Yoon, S. (2011). Large-scale community detection on YouTube for topic discovery and exploration. In *Proceedings of the 5th international AAAI conference on weblogs and social media* (pp. 486–489). Palo Alto, CA, USA: AAAI Press.
Girvan, M., & Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences, 99*, 7821–7826.
Hui, P., Chaintreau, A., Scott, J., Gass, R., Crowcroft, J., & Diot, C. (2005). Pocket switched networks and human mobility in conference environments. In *Proceedings 2005 ACM SIGCOMM workshop on delay-tolerant networking* (pp. 244–251). WDTN'05, New York: ACM.
Humphreys, L. (2008). Mobile social networks and social practice: A case study of Dodgeball. *Journal of Computer-Mediated Communication*, *13*(1), 341–360.
Isella, L., Stehle, J., Barrat, A., Cattuto, C., Pinton, J. F., & den Broeck, W. V. (2011a). What's in a crowd? Analysis of face-to-face behavioral networks. *Journal of Theoretical Biology, 271*(1), 166–180.
Isella, L., Romano, M., Barrat, A., Cattuto, C., Colizza, V., den Broeck, W. V., Gesualdo, F., Pandolfi, E., Rava, L., Rizzo, C., & Tozzi, A. E. (2011b). Close encounters in a pediatric ward: Measuring face-to-face proximity and mixing patterns with wearable sensors. CoRR abs/1104.2515.
Kaplan, A., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons, 53*(1), 59–68.
Lancichinetti, A., & Fortunato, S. (2009). Community detection algorithms: A comparative analysis arxiv:0908.1062. Physical Review, E80.
Lerner, J. (2005). Role assignments. In U. Brandes & T. Erlebach (Eds.), *Network analysis* (Lecture notes in computer science, Vol. 3418, pp. 216–252). Berlin/Heidelberg: Springer.
Leskovec, J., Lang, K. J., & Mahoney, M. (2010). Empirical comparison of algorithms for network community detection. *In Proceedings of the 19th international conference on world wide web (WWW '10)* (pp. 631–640). New York: ACM.
Liben-Nowell, D., & Kleinberg, J. M. (2003). The link prediction problem for social networks. In *Proceedings of CIKM* (pp. 556–559).
Macek, B. E., Scholz, C., Atzmueller, M., & Stumme, G. (2012). Anatomy of a conference. In *Proceedings of Hypertext 2012* (pp. 245–254). New York, NY, USA: ACM Press.
Malone, T. W., Laubacher, R., & Dellarocas, C. (2009). Harnessing crowds: Mapping the genome of collective intelligence. In *MIT Center for Collective Intelligence*. MIT, Boston, USA.
McDaid, A., & Hurley, N. (2010). Detecting highly overlapping communities with model-based overlapping seed expansion. In *Proceedings of the 2010 international conference on advances in social networks analysis and mining*. ASONAM'10, (pp. 112–119). Washington, DC: IEEE Computer Society.
Miluzzo, E., Lane, N. D., Fodor, K., Peterson, R., Lu, H., Musolesi, M., Eisenman, S. B., Zheng, X., & Campbell, A. T. (2008). Sensing meets mobile social networks: The design, implementation and evaluation of the CenceMe application. In *Proceedings SenSys'08* (pp. 337–350). New York: ACM Press.
Mitchell, T. M. (2009). Mining our reality. *Science, 326*, 1644–1645.
Murata, T., & Moriyasu, S. (2007). Link prediction of social networks based on weighted proximity measures. In *Proceedings of IEEE/WIC/ACM international conference on Web intelligence* (pp. 85–88).
Newman, M. E. J. (2004). Analysis of weighted networks. http://arxiv.org/abs/condmat/ 0407503.
Newman, M. E., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics, 69*(2), 026113.1–15.

Rosvall, M., & Bergstrom, C. (2007). An information-theoretic framework for resolving community structure in complex networks. *Proceedings of the National Academy of Sciences, 104*(18), 7327.

Scholz, C., Atzmueller, M., & Stumme, G. (2012). On the predictability of human contacts: Influence factors and the strength of stronger ties. In *Proceedings of fourth ASE/IEEE international conference on social computing (SocialCom)* (pp. 312–321). Boston: IEEE Computer Society.

Scripps, J., Tan, P. N., & Esfahanian, A. H. (2007). Exploration of link structure and community-based node roles in network analysis. In *Proceedings of the 7th international conference on Data Mining* (pp. 649–654*)*. Washington, DC, USA: IEEE Press.

Stehle, J., Voirin, N., Barrat, A., Cattuto, C., Isella, L., Pinton, J. F., Quaggiotto, M., den Broeck, W. V., Regis, C., Lina, B., & Vanhems, P. (2011). High-resolution measurements of face-to-face contact patterns in a primary school. CoRR abs/1109.1015.

Wassermann, S., & Faust, K. (1994). *Social network analysis*. Cambridge: Cambridge University Press.

Wongchokprasitti, C., Brusilovsky, P., & Para, D. (2010). Conference navigator 2.0: community-based recommendation for academic conferences. In *Proceedings of workshop social recommender systems*. IUI'10, Hong Kong.

Xu, B., Chin, A., Wang, H., Chang, L., Zhang, K., Yin, F., Wang, H., & Zhang, L. (2011). Physical proximity and online user behavior in an indoor mobile social networking application. In *Proceedings of the 4th IEEE international conference on cyber, physical and social computing* (CPSCom 2011) (pp. 273–282). Washington, DC, USA: IEEE Press.

Zhou, T., Lu, L., & Zhang, Y. (2009). Predicting missing links via local information. *The European Physical Journal B, 71*(4), 623–630.

Zhuang, H., Chin, A., Wu, S., Wang, W., Wang, X., & Tang, J. (2012). Inferring geographic coincidence in ephemeral social networks. In *Proceedings of European conference on machine learning and principles and practice of knowledge discovery in databases (ECML PKDD)*. Lecture notes in computer science, Vol. 7524, pp. 613–628. Heidelberg, Germany: Springer.

Chapter 5
Mobile Social Service Design for Special Context

Huanglingzi Liu, Wei Wang, Dong Liu, Hao Wang, and Ying Liu

Abstract This chapter introduces two cases about mobile social service design for special context. Unlike the general social network service through which people interact with each other mainly for social purposes, the social service for special context supports a temporary social network to conduct some activity together. Although people in the special social setting are mainly motivated by individual benefits, the relationships are also influenced or enhanced by the social activities in the special context. The two design cases discussed in this chapter are large-scale exhibition service design and local group buying service design. Based on the two case studies, we summarize a mobile social service design framework.

H. Liu (✉)
China Mobile, Beijing, China
e-mail: liuhlz@gmail.com

W. Wang
Adwo Mobile Media Technology, Beijing, China
e-mail: wwang04@gmail.com

D. Liu
University of Science and Technology of China, Hefei, China
e-mail: dong.e.liu@gmail.com

H. Wang
Babytree Inc., Beijing, China
e-mail: wanghao@babytree-inc.com

Y. Liu
Intel, Beijing, China
e-mail: newllyy@gmail.com

A. Chin and D. Zhang (eds.), *Mobile Social Networking: An Innovative Approach*,
Computational Social Sciences, DOI 10.1007/978-1-4614-8579-7_5,

5.1 Background and Introduction

In recent years, mobile services have attracted a lot of research attention and industry interests. Social software has seen a tremendous growth in user penetration over the past few years, and social software has already been integrated into mobile devices (Counts and Hofte 2006). Mobile social services take advantage of the nearly constant physical proximity of devices to their owners, to enable a wide range of new social interactions. Popular mobile social services, such as Foursquare and Loopt, allow users to share their location information with friends. However, in this chapter we introduce two case studies about mobile social service design for special context, instead of general social network service design. Unlike the general social context in which people gather together mainly for social interaction purposes, such as drinking coffee in a café, the special context refers to the social setting in which people gather together mainly for individual benefits, but at the same time the relationships are also influenced or enhanced by the social activities in the special context (Stock & Zancanaro, 2007; Guo et al, 2011).

The first case is the design for a large-scale exhibition. The exhibition industry in China has witnessed a rapid increase in the last few years. The annual growth rate of the Chinese exhibition industry is around 20 %, and the figure in the European exhibition industry is 2–3 % (Guo 2007). In order to improve the effects of the exhibition, a technology-enhanced visitor experience has started to gain more attention. For example, Expo 2010 Shanghai provided an interactive Internet platform to support visitor's remote experience.

Many studies have been done on the utilization of mobile technologies to enhance social interaction among visitors in the museum context. Stock et al. (2007) proposed to design technologies to construct a structured group and foster their social activities. Groupware is then used. In 'The Fire and The Mountain' exhibition, hybrid interactive artifacts (i.e., installations that support visitors manipulating and interacting with physical and digital exhibits) are utilized to enhance social interaction (Garzotto and Rizzo 2007). Some researchers use social awareness to help visitors create social experiences. For example, in the Imprints system (Boehner et al. 2005), visitors are supported to leave personalized marks at exhibits. They use an icon to represent themselves and attach that icon to museum exhibits so that each can seek out the traces of other visitors. In the Artlinks system (Cosley et al. 2008), visitors are allowed to see other visitors, their reactions to an exhibit, and connections among visitors through these reactions. The demographic and museum-going information allows people to make connections based on similarity of circumstance. People might, for instance, feel closer to someone who is about the same age, or who has visited the museum for similar reasons (Cosley et al. 2008). Awareness of social cues in physical information space enables 'social navigation.' The benefits of social navigation not only give users a sense of social presence and of not being alone in the space, but also provide users with a relative guide based on the knowledge accumulation (Svensson et al. 2005).

Compared with the museum context, the exhibition is more complex, and it would be interesting to explore and see if mobile technologies can be applied to promote social interactions in such a context and how to make it happen. Moreover, mobile technologies provide a good platform for the further development of social interactions because of their high penetration in end users.

There are different kinds of exhibitions. In this design case, we mainly focused on those that are open to the public, large in scale, and organized systematically so that exhibitors can promote their products or services. Social contexts in exhibitions are more complex than those in the museum. Exhibitors would like to attract the attention of visitors by various means, which, on the other hand, might distract visitors from the communication with the surrounding people. From the exhibitor organizer's point of view, the social interaction among visitors might be regarded as an unfavorable safety factor. A mobile social interaction service may not be accepted and used by end users if it only facilitates social interactions without being aware of its user's context. The user-centered design process was followed in the whole concept design process to fully understand social activities in the exhibitions.

The second case is the design and development of a mobile group buying service. Group buying is a business model where people with the same merchandise interests form a group and conduct the purchase together to achieve a discount. There are three types of participants in group buying: the merchant, the initiator, and the joiners. The initiator first negotiates with the merchant, promising to bring a certain number of customers in exchange for an appealing deal from the merchant. The initiator then distributes the information to gather a group of like-minded people and conduct the purchase. The joiners benefit from a lower price which is unavailable to the individual buyer. The merchant benefits from bulk sale and word-of-mouth advertisement. Group buying can be initiated by one of the joiners to benefit others as well as themselves, as well as a professional proxy, who coordinates the group buying for a certain commission fee from the merchant, or the merchant itself as a marketing campaign.

In the Internet era, most of the online group buying websites involve the "daily deal", for example, *groupon.com* or *t.dianping.com*. They provide a limited number of appealing deals every day on their websites so that users can easily choose from them. Group buying is open to any user online. The joiner is insensitive to the location of the merchant and other joiners (Tan & Tan, 2010). We call this *online* group buying. Online group buying attracts many more joiners than offline group buying, and thus usually has a better discount. On the other hand, joiners are usually total strangers. They have no information of each other except that they are interested in the same product. Their social interactions are therefore limited to the sharing of buying experiences such as comments and ratings, which are used to assist the buying activity. Explicit social benefits are hard to find in online group buying.

In addition, there exists local group buying, where the joiners, the initiator, and sometimes even the merchants are in the same local community. Such locality induces some interesting characteristics in group buying, which remain largely unexplored in the research community. In this case, we designed and developed a mobile service called "HappyGo" that supports local group buying with mobile devices. In order to reveal users' behaviors in group buying within the local context, and study the underlying factors of user behaviors during the local group buying, we conducted a trial involving more than 300 users from a company office.

Based on the two case studies, we summarize a mobile social service design approach that organically embraces both user-centered design principles and technology-oriented design principles.

5.2 Case 1: Design for Large-Scale Exhibition

In this case, the objective is to provide some social services in the exhibitor's mobile device to enhance their interaction with the exhibition context and promote the exhibitors or related products. The project scope is based on the available mobile technologies, for example, the ability of capturing user context, or recording their visiting history, or processing natural language, etc. However, user studies need to be done to understand the visitor's real experiences in the exhibition context, so that designers and researchers can further consider the possibility of mobile technology design for exhibitions.

5.2.1 User Study Methodology

We conducted individual interviews with participants to collect narratives through open questions, and evaluations on social service requirements through an interview-guided questionnaire. Then the shadowing method (Millen 2000) was carried out in exhibition context. Participants' real exhibition experiences were observed.

For recruiting, we asked help from a marketing research firm. Twenty-nine participants were recruited and interviewed in this study. Among the interviewees, two are exhibition staff working in a professional exhibition design company, four are exhibitors, four are exhibition organizers, and 19 are visitors who have visited at least two open-to-public exhibitions in the recent half year and at least one exhibition which exceeded 100,000 visitors.

Three categories of exhibition are covered in this user study. The first category is culture exhibition, such as travel, education, tea, or books exhibition. The second is an industrial exhibition, such as lighting/sound/music instruments, energy/water saving instruments, motor equipment, or information/communication technology exhibition. The third is trade exhibition, such as import and export exhibition.

All the participants come from Beijing and Shanghai, two major exhibition cities on the mainland of China.

Before the interview procedure, 19 visitors filled in the pre-questionnaire so that the demographic characteristics and the general issues relative to the exhibition visit could be known. Among the 19 visitors, ten were male and nine were female, aged between 17 and 60 years (mean = 36, median = 33).

The visitors were asked several open questions mainly designed to explore their real exhibition experiences. After the open interview, a five-point questionnaire was filled in by all the interviewees to collect both their qualitative and quantitative remarks on the necessity of various social interaction features.

5.2.2 User Study Findings

Exhibition visitors do have some motivation to use mobile social services, but at the same time, the barriers to adoption are obvious also. Factors influencing visitors' social engagement are analyzed, and implications for new mobile social service designs in exhibitions are presented in this section.

5.2.2.1 Exhibit Oriented Social Interaction

For general visitors, their goal for visiting the main exhibition is to know where, what, and how are the exhibits. No matter what kind of social interaction happens, the motivation for social engagement mainly lies in personal requirements on the knowledge of exhibits.

The most common social interaction for an exhibition visitor is to discuss with the responsible interpreter on some exhibits. Visitors who happen to feel interested in some exhibit gather around the exhibitor and ask for more information, by which a temporary discussion group is formulated and implicit social interactions among visitors are experienced.

In an exhibition environment, visitors are absolutely influenced by the unfamiliar people around them. They may imitate other's behavior, or modify their visiting plan based on the observation results. For example, visitors observed that some take photos of some exhibits, and they were motivated to take a photo of that exhibit also. If they saw that some exhibition booth was surrounded by people, even if they could not push in through the crowd, they were curious and eager to know what was happening at the popular place.

Although there are quite a lot of implicit interactions happening among unfamiliar visitors besides observing what others do, such as overhearing what others say, feeling other people's moods and the whole atmosphere of an exhibition booth, etc., visitors think that an exhibition visit is mainly a personal experience rather than a social activity.

During the interview session, they expressed the opinion that exhibits are their primary focus, and they have no interests in the unfamiliar people surrounding them. In fact, our observations proved that they told us one aspect of the truth. Visitors seldom discussed with each other concerning their own initiatives. And they usually had no idea of other people walking around them or looking at similar things around them.

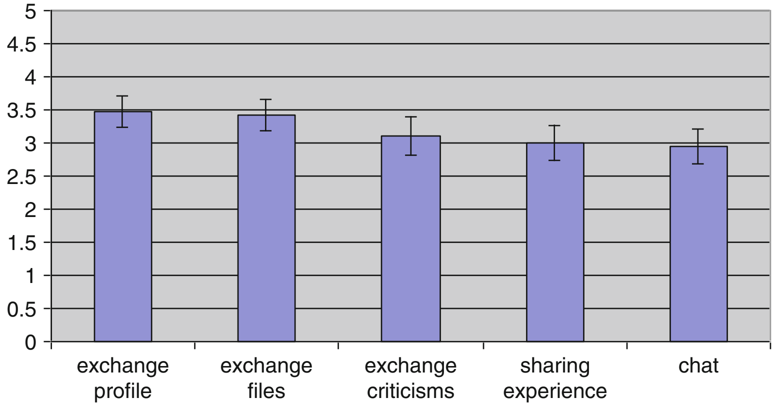


Fig. 5.1 Visitor remark on necessity of 'connection with other visitors'

A male visitor said: *I buy the entrance ticket to see the exhibits, not the visitors such as us…Why should I know what they are doing, what they are thinking? They have nothing to do with us…If I need more knowledge about the exhibits, I could ask the exhibitors there.*

In the questionnaire survey session, we designed five questions relative to the requirement of the connection with unfamiliar people.

1. Is it necessary to provide you a special service/device so that you can chat with other visitors on site?
2. Is it necessary to provide you a special service/device so that you can exchange exhibit introduction files with other visitors on site?
3. Is it necessary to provide you a special service/device so that you can exchange business cards or profile information such as interests with other visitors on site?
4. Is it necessary to provide you a special service/device so that you can exchange criticisms or remarks relative to exhibits/exhibitors on site?
5. Is it necessary to provide you with a special service/device so that you can share your visiting experiences, such as visiting routes, with other visitors?

Participants described the necessity on a five-point Likert scale. Figure 5.1 shows the average requirement level of 19 visitors on the above five aspects.

The average scores of the necessity on the five social features are between 2.5 and 3.5; which means social interaction among visitors in an exhibition situation is an uncertain demand. Some visitors may think it would be ok to have such kind of support but some think it is useless, or most people feel it is difficult to make a judgment on the necessity since they cannot imagine the use scenario once the supporting technology is there. However, participants respond differently to the five questions. If the social interaction requirement is described in a more detailed way or the social activity is exhibit/exhibitor-oriented, it seems people are more likely to accept it.

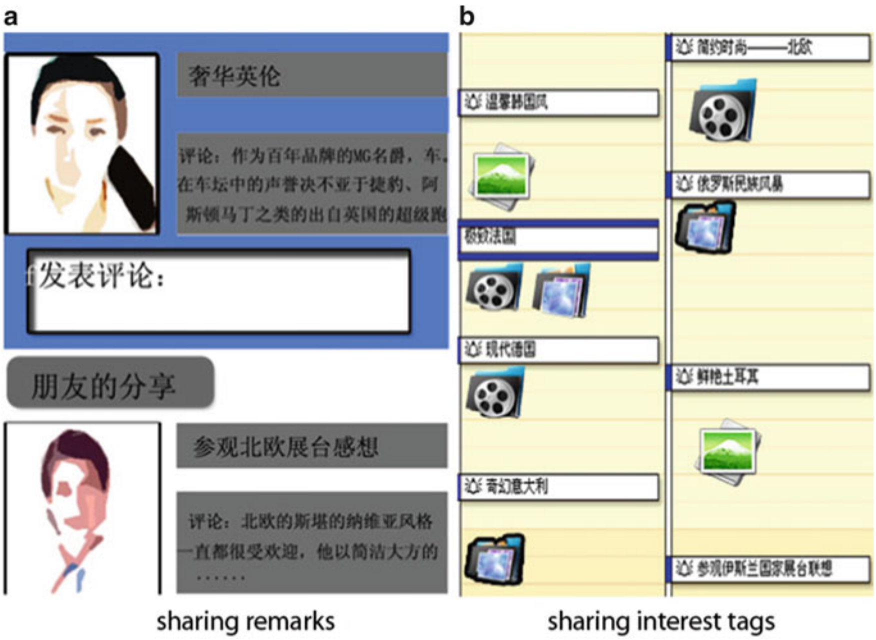


Fig. 5.2 Two UI solutions for experience sharing: solution (**a**) allows each visitor to publish and share their impressions or remarks in a conversation thread, solution (**b**) allows each visitor to express their impressions in some keywords and attached photos or videos

5.2.2.2 High Requirement on Time Cost of Social Interaction

'Time critical' is important for exhibition visitors. All 19 participants expressed the view they experienced time pressure during an exhibition visit.

A mechanics professional said: *When visiting an excellent exhibition, I'm eager to know as much as possible about the exhibits. As for some special exhibition, I would like to spend more time. But half a day, even 1 day, is not enough for me… Why not visit the same exhibition twice? Oh, you are not born to be a visitor. Other things are there for you to do.*

In the questionnaire, except for the questions relative to the requirements of the connection with surrounding people, we also designed questions to measure visitors' requirements for connection with their company. Before answering these questions, many participants require the interview moderator to describe the connection solutions in a much detailed way. They showed great concerns with regard to the cost efficiency of technology-supported social interaction.

> How to chat with others…How to exchange remarks…Is it possible?…Is it easy to do?…
>
> What kind of comments are you talking about?…How can I take these comments down?… It's too boring if it needs a lot of input…

Figure 5.2 shows two UI solutions for sharing a visiting experience which we also tested in the questionnaire survey. Fifteen visitors preferred solution (b).

The main reason for their choice was that it is '*easier to practice sharing since there is no need to input a lot and read a lot on the exhibition site*.'

If the information consumption based on social interaction is time consuming, there is no real advantage in having access to surrounding people and getting information through mobile technology, because they are already located in an information-rich environment.

To reduce time cost of social interaction among visitors, context-aware UI is required. Explicit input from visitors may interrupt the relationship with the physical world. Context could reduce the input cost and make the communication much more efficient (Hong et al. 2005). With the help of context-aware computing, the ways that people engage in social interaction in exhibitions could adapt to several situations involving user locations, nearby exhibits, personal interests, social activity history, etc.

5.2.2.3 Influence from Exhibitors on Social Interaction

As the designer and organizer of an exhibition, exhibitors would like to dominate and foster mutual communication with visiting customers so that they can spread and promote their products/techniques. For exhibitors, knowing visitors and being known by visitors is their main target.

In order to promote brands or advertise their exhibits, exhibitors aim to attract as much of the visitors' attention as possible. They usually do a lot of preparation work before the exhibition, for example, designing an exhibit booth, displaying the product introduction poster, brochure, and videos or interactive media, and so on. Except for this design work, exhibitors also hold some on-site activities such as gift distribution, on-site lectures or performance shows, etc. to access a group of approaching visitors. However, those gathering activities are not very satisfying compared to the exhibit's display design work. Figure 5.3 shows the percentage of satisfied visitors on the exhibition activities discussed above. Most people are satisfied with the distribution of printed exhibit introduction material; and the lowest number of people are satisfied with the gift distribution. During interview, most visitors complain about the gift distribution activity.

The following are some of the citations from visitors:

> Most are cheap gifts…Exhibitor mainly wants to collect our business cards…
>
> Usually useless gadgets…I seldom try to get these things…Those who queue in line to get the gifts are usually not the target customers, I think…

Faced with those challenges, exhibitors during the interview explained their embarrassments:

> We'd like to distribute good gifts to those potential customers…However we don't know whether they are potential customers before we meet them…
>
> Of course gifts are small…Could you imagine a visitor carrying big pieces and going around the exhibition?

Obviously, exhibitors tend to gather people to enhance the promotional effects, but they have problems categorizing people. If exhibitors could classify visitors into different groups, pertinent on-site social activities for different categories of visitors could be held to expand or strengthen exhibition influence.

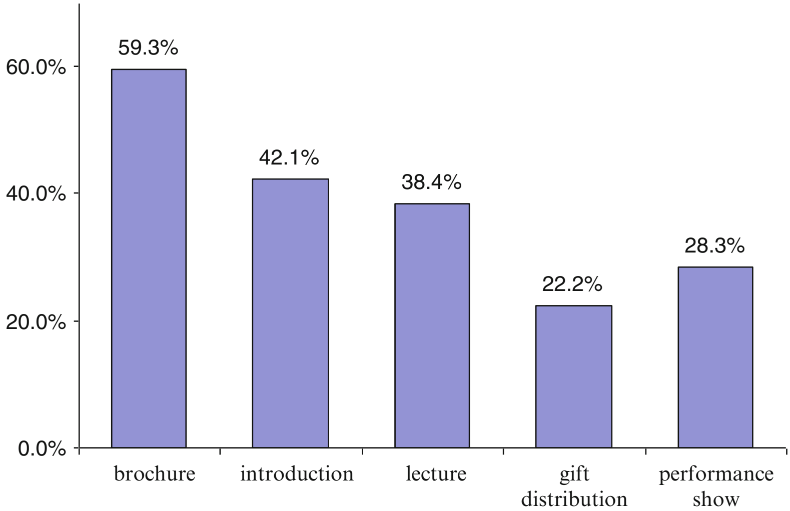


Fig. 5.3 Percentage of satisfied visitors with regard to exhibition activities

5.2.2.4 Connection Within Real and Long-Term Based Social Network

Petrelli and Not (2005) report that only 5 % of visitors come to a museum alone, while 45 % come in organized groups, 20 % with friends, and 30 % with children. Our questionnaire reached similar findings. Nineteen participants who have attended exhibitions at least twice in a recent half a year reported 44 exhibition visits all together, and 79 % visits have been with a company. Based on the interview and observation result, we can generally categorize visitor groups into three main categories according to their on-site activities:

1. The “always together” group in which members would like to go everywhere together
2. The “linked group” in which members may separate some of the time during the visiting procedure, but try to have each other in their field of vision or be aware of each other’s location from time to time
3. The “leave-together group” in which members visit the exhibition separately and make an appointment to leave together

When they are asked why they go to an exhibition with someone else, two main reasons can be deduced:

1. Better effects, such as better memory and a better impression can be achieved from an exhibition
   A visitor to an ordinary supplies exhibition made this comment: *We discussed the various exhibits, compared their strongpoints and weakpoints during the visiting procedure…A wiser decision could be made based on a full discussion of different opinions*.

2. The exhibition itself could be a social platform for visitors and their companions A visitor of a motor exhibition remarked: *It won't be lonely if someone comes with you to attend an exhibition…It's also a good opportunity to talk about common interests…*

However, these connections to companions are not encouraged by the exhibition environment. Firstly, crowding is a major problem for most exhibitions in China. It's not easy to maintain the "always together" group. Once the group members are separated by the flood of people, it will be difficult to 'see' where each other is. So "the linked group" may find that they have lost the link to their company. Even worse, making a phone call in the crowded and noisy environment to connect with friends is absolutely not an easy task. Secondly, as for the 'leave-together group,' although they are supposed to interact with the exhibits or exhibitors individually, they still would like to share the visiting experience with their mates from time to time, even to discover more about exhibits together under some condition. For example, one visitor mentioned *when I saw something really outstanding, I would like to recommend this to my mates immediately, even to friends who didn't come… If I know my mate is also interested in my recommendation, I may wait for him and exchange comments on site, then visit individually….*

In the questionnaire, we designed four questions relative to the requirements of the connection with their friends. They are:

1. Is it necessary to provide you a special service/device so that you can locate your friend(s) on the exhibition map?
2. Is it necessary to provide you a special service/device so that you can recommend an exhibit/exhibition stall to your friend(s) on site?
3. Is it necessary to provide you with a special service/device so that you can send appointment (including venue, time) to your friend(s) on site?
4. Is it necessary to provide you a special service/device so that you can exchange comments or impressions with your friend(s) on site?

Participants described the necessity on a five-point scale. Figure 5.4 shows the average requirement level of 19 visitors on the above four aspects.

People tend to accept the intention to enhance the connections with their mates in an exhibition environment if the cost of interaction is not very high. As for 'locate friend',' recommend friends exhibits', and 'make appointment with friends in the exhibition', all of the three average scores of necessity are higher than 3.5, which means that participants can easily point out the use scenarios and describe the benefits that the technology may bring them. For example, one male visitor described: *with the service (locate friend), I can find my mates in the crowded exhibition… I could be very relaxed during the whole exposition since I know where my friend is and we can meet somewhere if we want to visit together in the midway… If I find something interesting, I can recommend it to my friend or let him join in my visit…*

Compared to being connected to people they do not know, visitors show more interest in being connected with their mates and remote friends. Before going to an exhibition, some visitors may communicate with people online to get some information, or contact people who possibly have similar interests and also go there.

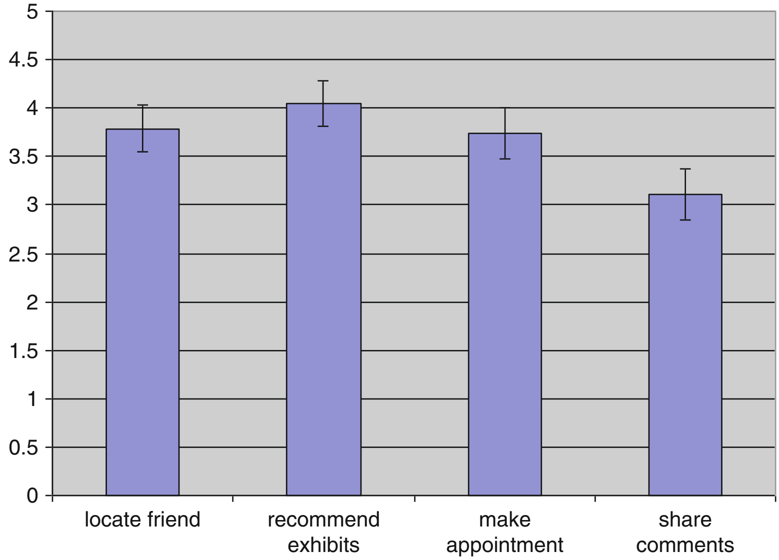


Fig. 5.4 An average visitor's remark on the necessity of 'connection with mates'

After the exhibition, visitors may talk about what is interesting within their social networks, publish comments, and/or publish photos on their Blog or BBS, which means the sociality is mainly experienced by visitors before and after the exhibition. In the exhibition, visitors require more connection with their mates or remote friends than the unfamiliar people surrounding them. Incentive or reputation structures should be built toward the transformation from temporary social gathering to real, long-term based social networking.

5.2.2.5 Multi-channel Interaction Environment

According to our observations, visitors keep interacting with various physical tools such as indoor signs, product brochures, printed posters, public displays, touch sensitive screen, and so on.

Among these physical tools, the printed brochure is one of the most important communication media between visitors and exhibitors. Exhibitors may write down supplementary information on the printed sheet according to a visitor's requirement. Exhibitors and visitors may exchange information face-to-face based on the printed material. Visitors can take notes on the printed sheet and keep it as long as they like. Public display or huge posters also play an important role in advertising products or information guidance. Exhibitors may play an advertisement video through a big display to attract approaching visitors. Exhibition organizers may post a big floor map of the exhibition hall at the entrance.

Fig. 5.5 Part of the storyboard created from brainstorming

Visitors usually take their mobile phone, camera, business card and paper notebook with them during the exhibition visit. They are kept busy with photo and note taking, looking, and listening.

These interactions with the physical exhibition should not be interrupted by the technology-supported social interactions and should not be an obstruction to social interaction, which could be a big challenge for designers. Therefore, environmental design is required. One way might be that the newly designed mobile social service could be coordinated with other exhibition facilities and personal articles, which means the new design is a component of the whole exhibition service ecosystem rather than taking the place of all existing facilities such as a printed visiting guide, guiding signs, public digital display, product brochures on exhibition site, etc.

5.2.3 Concept Design

In this section, we turn to our design thread and undertake a series of design exercises to help outline our findings and implications. The goal is not just "needfinding" but rather to gain further insight on constructive solutions. Our user-centered design process starts from image-based brainstorming to storytelling and scenario building (Fig. 5.5). After review and evaluation, one of the concepts, 'Sharing Exhibition Visiting Route', emerged from those exercises for further study. Its scenario is described as follows:

> John is visiting an auto exhibition with his friend David. They are fascinated by a new sports car model and would like to tag it so that more information about the model can be obtained.

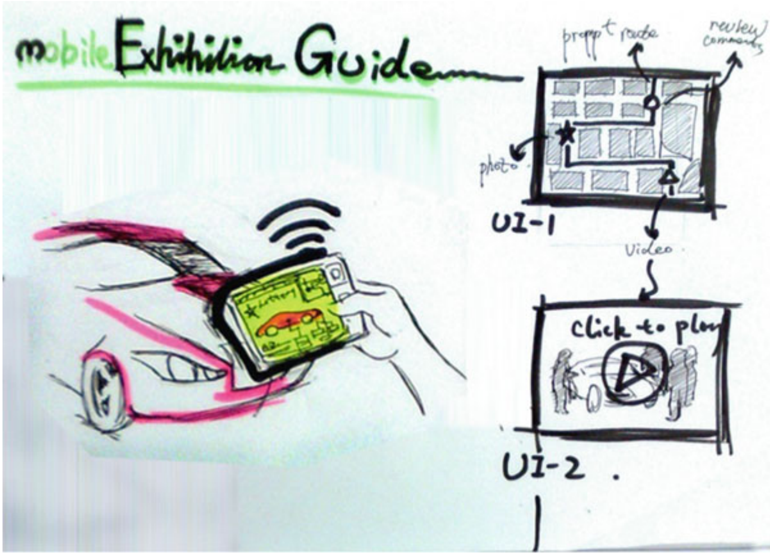


Fig. 5.6 Concept sketch of exhibition visiting and route sharing

> John opens the Exhibition Guide from his mobile handheld and the nearby exhibits list is displayed. He tags the model as 'interesting' and notices that there are some lottery-attached questionnaires and other visitors who also are interested in this model. The name "Lucy" and her remarks about the car model appear, since she is supposed to have a lot in common with John and she is willing to share her visiting routes with others (UI-1 in Fig. 5.6). John decides to follow her route to get better navigation round the exhibition, for example to see a video shot by Lucy (UI-2 in Fig. 5.6). During the visiting procedure, when a difference appears between them, John chooses another route to follow.
>
> Public displays in the exhibition show the popularity of each exhibition booth (number of tagged times) and the crowding status. John sometimes changes his route according to the information. Sometimes the public display shows a game registration ad which encourages visitors to vote for something or compete with other participants using mobile handhelds. John wins a digital lunch coupon because he posts an advertising slogan for a car model and gets the most votes.
>
> David has no interest in following other people's routes. He creates his own interests, tagging 'interesting' exhibits and adding relative comments to it. Sometimes he recommends some exhibits to John. If John likes it, the current followed route might be modified so that John can visit the recommended exhibit on his route.
>
> After the exhibition, John reads all the comments on his PC that he followed and created, while David is happy to find that his route followers are increasing constantly.

In this concept prototype, features reflected in the above five implications, which are combined to build a mobile visitor community, include:

- Personal requirements on the knowledge of exhibits are route information with point of interests shared among visitors having similar interests.
- Time cost: context information is utilized and users need to make some clicks or use simple text input to complete route sharing and following activities.
- Exhibitor's requirement on information distribution and collection are service-registered visitor's basic information, personal interests, and feedback about

exhibits, which can be captured by survey questionnaires and social promotion activities such as games which can be done in a digital way.
- Maintenance of temporary social network: social engagement could be very flexible and personal since the social network is dynamic and users could participate in any activity without careful preparation. Furthermore, visitors on site and off site are connected.
- Coordination with the exhibition environment: mobile handhelds are not the only channel for visitors to engage in social interaction, public displays in the exhibition also play an important role in augmenting the social atmosphere.

5.3 Case 2: Design for Group Buying

In this case, the objective is to provide group buying services in the local context with the support of mobile Internet. There are already lots of group buying services. For example, some popular group buying sites such as GroupOn or Dianping in China present group buying deals and allow customers to buy the deals. In addition, lots of information on how a customer initiated and organized a group buying activity can be obtained through various Chinese University online forums. For example, the forum of Tsinghua University (www.newsmth.org), one of the most popular online forums in China, sets up a sub-forum specialized for group buying issues. People can post a group buying requirement, conduct a survey to see the potential buyers of some product, initiate a group buying activity, or enroll them in a group buying activity. In addition, there exist local group buying activities where the joiners, the initiator, and sometimes even the merchants are in the same local community. For example, some local group buying is organized on community online forums, which are originally designed for user-generated posts (we can call it as offline group buying). But the online forum does not well support group-buying-related actions such as group publishing, management, and statistics for initiators, and following, joining, and quitting for joiners. On the other hand, with the increasing processing power and decreasing cost of ubiquitous connectivity, people start to access Internet on-the-fly from their mobile devices rather than from personal computers (PC) at fixed places. Following this trend, mobile Internet services are taking off, which inherently have two advantages over PC-based Internet services. First, information can reach end users faster, since people usually carry their mobile devices with them. Second, mobile devices capture the context of users, especially the personal profile and location, which can be utilized to provide more personalized user experiences. The advantages of mobile Internet can be taken to benefit local group buying services. However, all these potential benefits provided by mobile group buying services are only hypotheses. We need to answer the following questions before publishing and operating it for mass use.

- Is a local group buying service usable and valuable to end users?
- How is the mobile group buying service used by local people if compared with *offline* and *online* group buying services?

- Which factors influence user behaviors?
- Will a mobile solution enhance the user experience, especially compared with online forum-based group buying?

We developed a mobile service called "HappyGo" that supports local group buying, and conducted a trial involving more than 300 users from a company office. Based on our findings, we believe that local group buying complements online group buying by creating a "local" economic circle, while also providing users with social benefits. Factors influencing the mobile group buying experience are also discussed and analyzed.

5.3.1 HappyGo System Design

The HappyGo system adopts a client–server architecture, where clients are installed as mobile applications into the mobile devices of all users, and connects wirelessly to a centralized server to report the locations of users and the HappyGo service.

The HappyGo service concentrates on facilitating the spread of deal information as well as communications across users to support mobile group buying. Accordingly, it provides users with the ability to create and maintain groups, browse groups and receive push notifications of new groups, customize viewing by subscriptions, view the online message board, and join a group to participate in group buying.

The main features of HappyGo include:

1. Group buying information subscriptions and push notifications
   Users subscribe for shopping categories they are interested in; whenever a new group falls into one of the subscriptions that are created, notifications will be sent by the server to subscribed users.

   User can specify their requirements concerning the attributes of a shopping group, i.e., category of the merchandise and discount (as Fig. 5.7 shows). For the category of merchandise, a tree-structured taxonomy is provided for users to select from, e.g. "Electronic/Digital" – "Mobile phone" – "3G mobile phone" is a level-3 category in the taxonomy. For discount, options include "no discount" and from "10 % off" to "90 % off."

   When adding a subscription, the user is encouraged to make it visible to other people. Accordingly, users can find out all publicly visible subscriptions. This design is intended for potential group initiators, who can check for the trends of people's shopping interests and make a rough estimation of the scope of the group. The requirements on category and discount are shown without identification of the subscriber, to protect privacy.

   The HappyGo server deploys a matching module to handle the subscriptions. Whenever a match is found, the group information will be pushed from server to the client that is associated with the subscribed user. Such push notification requires the server to actively contact the client, which can be realized by long-lived network connection between server and client. The client will then notify

Fig. 5.7 The UI for making subscription of merchandise in HappyGo

Fig. 5.8 The UI for detailed group information and potential actions

the user appropriately using, for example, an icon flash, sound, or pop-up message. If the user is offline, the HappyGo server can also send short messages to notify users of the group information, given that users have registered their mobile phone numbers.

Push notifications will also be utilized by the server to deliver important or urgent information, such as someone joining or quitting a group (sent to the group initiator), the initiator announcing completion or cancellation of a group (sent to all group joiners), and so on. Such information can also be sent via short messages if the target users are offline.

2. Group buying actions
   Figure 5.8 shows the UI for detailed group information as well as potential actions for users, which will be different for individual users – group initiator can "complete" or "cancel" their group, whereas other users would "join" or "quit" the group. They can enter a chat room associated with each group to discuss any group-related issues.

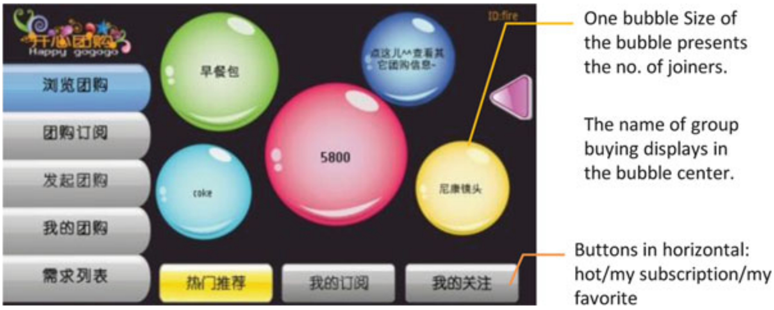


Fig. 5.9 Browsing groups in HappyGo

3. Browsing groups
   As shown in Fig. 5.8, users can browse active groups in three ways: the most popular groups, the groups matching the user's subscriptions, and the groups that the user follows. Following a group means to receive latest news of that group, and users may choose to join it at a later time (Fig. 5.9).

5.3.2 Field Trial of HappyGo

We deployed HappyGo in a Nokia office building in Beijing, China which has about 2,000 employees. Several stores, such as restaurants and groceries, are located in the office building. During a 1-month pilot test from May 2010 to June 2010, 318 users had installed and tried out this service. We analyzed the activity log from the server after pilot test, together with an online questionnaire survey and focus group study.

An online questionnaire and a small-scale focus group were used to study the participants' shopping experiences, their main concerns during the trial of the service, acceptance level of the HappyGo features, and overall satisfaction level and open comments.

There are plenty of variables which may affect participants' HappyGo experiences. In order to make the after-trial questionnaire study more focused, before the formal collection of subjective data, we firstly conducted a special mobile survey, which we called an event-triggered mobile survey.

HappyGo users who just logged out or completed a group buying activity would get a SMS with a link to a 1-min questionnaire survey invitation. Two questions were asked: (1) what score will you give to the most recent use experience? The score is rated on a 10-point system. The higher the score, the better the user experience. (2) Why do you give such a score? To avoid spamming users, the SMS is sent out to random users.

One hundred and eight responses were collected, and 92 effective textual entries were extracted for further bottom-up content analysis (Table 5.1). Nine categories of factors affecting users experiencing HappyGo were found as shown in Table 5.2.

Table 5.1 Factors analyzed in the event-triggered mobile survey

Factors	Example	Percentage of total comments (%)
Utilitarian motives (for cheaper product or effortless procedure)	Very happy to get the cheaper product…	16.3
Hedonistic motives (for fun/curiosity/enjoyment)	It's really fascinating when buying something with the colleagues together	4.35
Risk concern on product	I worried about the quality of the food… Too little information provided…	17.39
Trust concern on initiators/joiners	Some didn't get his/her products even he/she confirmed the participation of group buying	10.87
Risk concern on group formation	Fewer participants…	8.70
Usability problem on communication	I failed to be updated with some important information submitted on group message board	27.17
Usability problem on mobile input/output	Too troublesome to describe the product information with the mobile device…	8.70
GPRS cost	I didn't upload the product photo since I don't know how much it may cost me	7.6
System robustness	No response when I clicked…	8.70

Table 5.2 Overall usage statistics for the trial use

Overall usage statistics	Mean
Frequency of group browsing usage	134.23
Frequency of all-subscriptions browsing usage	5.01
Frequency of message board usage	9.23
Number of joined groups	2.67
Number of created subscriptions	0.64
Number of sent messages	1.65
Number of created groups	0.35

The above nine factors were further investigated in the final questionnaire. A five-point Likert scale is used for the subjective evaluation.

The online questionnaire was sent out to HappyGo users by e-mails, to which 179 users responded. Among the responses, 125 were females, 45 were males, and nine did not mention their gender. From another perspective, 40 were inactive users who logged in to HappyGo several times but never joined or initiated any group. One hundred and thirty nine were active users, who joined or initiated some groups at least once.

In order to compare with online group buying, we also crawled through and analyzed more than 30,000 daily deals from hundreds of online group buying websites in a 3-month period, which included the most popular group buying websites in China, such as dianping.com, ju.taobao.com, lashou.com, nuomi.com etc. Altogether 33,944 groups were crawled through and divided into ten categories.

Table 5.3 Buying group property

Group property	Mean	Std. deviation
Pre-defined buyers no.	6.519	5.184
Duration (hours)	121.412	103.303
Prices	162.667	342.361
Discount	0.785	0.179
Final buyers no.	19.593	20.905

As discussed before, group buying forums, especially those mainly serving people living in the same community, support local group buying in a similar manner to HappyGo. Both of them are user-initiated and have a fixed price. In order to further justify the benefit of local group buying with mobile device support, we compared the data collected from HappyGo trial use and the data collected from a community group buying forum. We chose the Huilongguan group buying forum (http://jc.hlgnet.com/hot.php) as the comparative data source. Huilongguan is one of the largest living communities in Beijing, China. The community occupies an area of 8.5 million square meters, and the residential population is around 230,000. The Huilongguan forum is very popular with the residents. The number of registered users is around 50,000. All the forum notes posted from 6th May to 5th June 2010 were collected and analyzed. We choose the same period during which we tested HappyGo. One hundred and twenty-three group buying activities were organized by 75 registered users on the forum during 1 month. The following buying group information was also extracted for further comparison study: initiator, duration, discount, and enrollments.

5.3.3 Findings

5.3.3.1 Overview of HappyGo Usage

As previously mentioned, 318 users participated in the field study according to the log data. Two hundred and thirty-seven users joined or initiated group buying activities at least once. Among them, 52 users created 112 group buying activities. The average number of joiners per group was 7.6 (std. deviation 1.25), and the largest group had 77 joiners. Two hundred and one users had joined buying groups at least once. The other 81 users logged in to browse only.

Table 5.2 shows the average usage statistics. As for the information consumption, a participant browsed the group buying activities 134 times, browsed the subscriptions 5 times, used the message board 9 times, and joined 2.67 buying groups. For information production, a participant created 0.35 buying groups, 0.64 subscriptions and 1.65 messages. Given the 1-month period, we conclude that HappyGo had been actively used by the participants. Table 5.3 summarizes the properties of 37 successfully completed buying groups.

Questionnaire analysis shows that the overall satisfaction level is 3.71 (std. deviation 0.9), which is nearly ‘satisfied.’ Ninety-seven percent of participants showed willingness to use HappyGo in the future (the average Likert score given by them is 4.5). Eighty-eight percent of participants showed willingness to subscribe for the group buying products offered by the service in which they were interested.

5.3.3.2 HappyGo Is Effectively Used by Three Types of Initiators for Financial Benefits

In offline group buying, most groups are initiated ad hoc by one of the joiners, whereas proxy is the only type of initiator in online group buying. In HappyGo, we find all three types of initiators (i.e., merchant, proxy, and joiner).

There are three strawberry groups initiated by a user whose family owns a small strawberry plantation. He has sold strawberries to colleagues for several years. Before HappyGo, he sold to colleagues in nearby teams. He acknowledges that HappyGo helped to increase the sales volume by a large margin. Another *merchant* initiator uses HappyGo to advertise her online shop. She says, "*Actually I don't care how many people finally join the group … They noticed my online shop, that's good enough*".

We observe that such merchants are mostly individual or family-owned small businesses. They serve a small number of customers due to limited resources, and rely almost solely on word-of-mouth advertisement. Thus they usually target returning customers from the same local community. Online group buying websites are not suitable for them as their capacities and visibilities are far below the requirement of those websites. HappyGo pools together the targeted customers of those small businesses, so as to attract them to initiate group buying in HappyGo. Another benefit is the saving of logistic cost as "*I can finish all the transactions in one delivery as all the customers are in the same building",* as quoted by the initiator of the strawberry groups.

Proxy initiators also appear in HappyGo. The proxy is a normal HappyGo user who has a close relationship with a merchant. For example, one initiator organizes group buying on "Chinese dried dates" in HappyGo. The goods are provided by her roommate. This group successfully attracted 45 joiners, and the initiator received monetary reward from her roommate afterwards. As usual, a proxy initiator is always motivated by the commission fee to explore possibilities to connect merchants with new customers. So it is not surprising to see proxy-initiated group buying in HappyGo.

Joiner-initiated groups are also common in HappyGo. For example, many users initiate lunch groups to enjoy a 25 % discount from the company cafeteria if the group size reaches four. Another user-initiated group conducts purchases in an online clothes store to enjoy the free shipping after the sales volume reaches a threshold.

We see many joiner-initiated groups in HappyGo, for two reasons. From the joiners' perspective, they feel comfortable to join the groups initiated by other joiners. All participants share a common background, i.e., colleagues. It leads to a certain level of trust among users. From an initiators' perspective, they have some knowledge of other users in terms of the shopping interests because they share the same context. Taking group lunch as an example, almost everyone in the company prefers to have lunch within the office building due to limited break time at noon. Before a user initiates a lunch group, she knows there is a good chance to attract enough joiners. Such pre-knowledge reduces the perceived risk of the initiators to encourage more joiner-initiated groups.

We observed that, as an open platform, HappyGo attracts all three types of initiators to pursue financial benefits. It builds an economic circle in the community to benefit local merchants, proxies, and normal users.

5.3.3.3 HappyGo Is Used as a Group-Buying-Oriented Social Medium

In addition to financial benefits, joiners are also found to apply HappyGo for social benefits. We observed four patterns.

First, emotional communications among participants are observed in HappyGo. The joiners often express their feelings towards the initiators in the message board during the group buying session. For example, both compliments such as: *You're so generous to help,* and criticisms such as *Just close the group if you are not serious* … are posted. It is a natural social interaction, as the joiners know the initiator is a person. In online group buying, joiners perceive the initiator as a virtual entity so that such emotional expressions are rare.

Second, offline interactions are promoted by HappyGo. One focus group participant recalled how she organized a group buying activity through HappyGo. *The first time when I proposed group buying at HappyGo, I was worried that not enough persons would join. So I encouraged people around my seat to register* … When some users picked up the delivered products, they were noticed by others, who then approached to ask. Word-of-mouth promotion happens. One interviewee who actively joined group buying activities mentioned: *When I took the strawberries in the elevator and shuttle bus, there were many questions asked such as 'where did you get that?', 'what's the price?', 'how does it taste?' …I feel like an agent.*

Third, HappyGo is used by some joiners purely as a social platform. For example, someone proposes a group for "testing how many people know me in this building". The initiator asks others to join if they know him. Both the initiator and the joiners treat it as a fun experience.

Fourth, HappyGo offers a new opportunity for joiners to extend social networks. Merchandise posted in HappyGo were categorized and compared with those in online group buying websites (see Table 5.4). While some categories such as food and electronic devices appear frequently in both HappyGo and online group buying websites, we observe one category which is unique, and also the top category in HappyGo: *lunch together*. Thirty-seven groups for lunch together were created, which constituted 1/3 of all groups. During the 1-month trial period, at least one lunch-together activity happened every working day.

Lunch together is a social activity for most people to exchange information and enhance social bonds. People usually have lunch with close friends and colleagues. People seldom *extend* their social network during lunch. However with HappyGo, offline social gatherings, such as having lunch together, can be linked with group buying; since all joiners of a lunch-together group share the same economic purpose

Table 5.4 Categories of merchandise in HappyGo and online group buying websites

Category	HappyGo		Online	
Maternal infant care	2	1.79 %	3,372	9.93 %
Home article	2	1.79 %	1,916	5.64 %
Living service	3	2.68 %	2,388	7.04 %
Entertainment	3	2.68 %	2,375	7.00 %
Clothes	8	7.14 %	1,171	3.45 %
Cosmetics	8	7.14 %	6,079	17.91 %
Electronic devices	11	9.82 %	3,524	10.38 %
For fun	17	15.18 %	0	0
Food	21	18.75 %	5,509	16.23 %
Lunch together	37	33.04 %	0	0
Restaurant	0	0	4,531	13.35 %
Others	0	0	3,079	9.07 %

(for a 25 % discount at the company cafeteria), they now have a good "starting point" to meet new people. Many interviewees recognize lunch together helps them get to know new people and make new friends. "*It's very natural to build a relationship with people working in the same company…*" "*We have a lot to discuss during lunch even if we did not know each other before…*"

During the trial, we imposed little limitation on how HappyGo could be used. The above observations suggest that group buying can be used as a social medium in the local context. *Socializing itself becomes the purpose*.

5.3.3.4 Factors Affecting HappyGo Usage

We compared the subjective evaluation on the nine factors between active users and inactive users. Statistical analysis shows that the inactive users give higher scores on the following factors:

1. Product risk
   When assessing the risks perceived during the trial use of HappyGo, the inactive users perceived higher level risks towards the product than the active users.
2. Drive for cheaper products
   As to the description 'I use HappyGo mainly for cheaper products', the inactive users gave lower scores than the active users, which means that the inactive users are less likely motivated by cheaper products.

One most encouraging benefit of group buying is good economy. However, those inactive users are not driven so much by this benefit. At the same time, they showed higher perceived risks towards the product. According to the theory of planned behavior (TPB) (Ajzen 1991), an unfavorable attitude gives rise to weak use intention. So it's easy to understand that lower motivation and higher perceived risks are the main factors affecting inactive users' acceptance.

Table 5.5 Differences between active users and inactive users

		Active users	Inactive users
Risk concern level on product	Mean	3.98	4.82
	T-test value	1.902 ($p<0.001$)	
Trust perception on initiators/joiners	Mean	2.83	2.40
	T-test value	1.973($p=0.144$)	
Risk concern level on group formation	Mean	3.92	3.88
	T-test value	1.976($p=0.932$)	
Usability problem on communication	Mean	4.22	3.03
	T-test value	1.994 ($p<0.001$)	
Usability problem on mobile input/output	Mean	4.15	3.42
	T-test value	1.987($p=0.024$)	
Risk concern level on GPRS cost	Mean	3.83	3.53
	T-test value	1.980 ($p=0.487$)	
System robustness problem	Mean	3.25	3.37
	T-test value	1.970 ($p=0.380$)	
Motivation level on buying cheaper products	Mean	4.07	3.43
	T-test value	1.981 ($p=0.050$)	
Motivation level on fun/unique experience	Mean	3.53	3.47
	T-test value	1.986($p=0.816$)	

To reduce the perceived risks, social comments and a reputation system could be provided to the inactive users, so that they have more valuable information to make better decisions. With the community-based reputation system, all the users can give credibility scores or comments to the group initiators, and satisfaction scores or comments to the buying groups. They also can recommend some buying groups to others directly. The initiator can also set reputation requirements towards potential joiners, and leave feedback to the joiners after the completion of the group. The feedback from the social community can offer valuable information for all potential buyers. To improve the low motivation level, more explorative designs need to be done to meet users' potential mobile group shopping needs.

Communication and mobile input/output are the two main usability factors identified by HappyGo users. Table 5.5 also indicates a surprising finding, which is that the active users showed a higher level of agreement on the two usability problems than the inactive users. The result somewhat conflicts with the technology acceptance model (TAM) (Bagozzi et al. 1992), which suggests that perceived usefulness and perceived ease-of-use affect acceptance. One reason may be that those inactive users have little knowledge of system use, since they never created a group, or seldom communicated with others.

We then further analyzed the correlation between the subjective evaluations on the two usability problems and the number of buying groups which the users joined in. The result showed that number of groups was negatively correlated with the subjective evaluation on usability problems ($p=0.027$, 0.031). Those active users who joined or initiated fewer buying groups recognized the two usability problems more seriously. Usability problems on communication and mobile input/output are the main factors affecting active users' acceptance.

Table 5.6 Comparison between online forum and HappyGo

	Online forum	HappyGo	*T*-test (sig)
Number of groups	123	112	N/A
Number of initiators	75	52	N/A
Average duration (days)	207	6.6	N/A
Average discount	0.53	0.80	8.935 (0.000)
Average enrollment (persons)	10.8	7.6	0.998 (0.324)

When interviewing some initiators and active joiners in the post-pilot focus group, they expressed clearly that the limited mobile input and information browsing ability hindered them a lot.

> It's troublesome inputting the product information with the mobile device… and inputting more explanations of the product in the message board from time to time.
>
> It's really troublesome to contact others, including the group initiators…
>
> There are so many questions proposed by the participants, it's troublesome to answer the questions with my mobile input.

Better communication among joiners during group formation and after the group formation is clearly proposed by the focus group participants. The current chat room for each group cannot satisfy all the communication needs at various stages of the group buying. For future design, HappyGo needs to strengthen the community functions, such as adding the news/update feed, setting up the special 'after-group' session for initiators to distribute notifications and communicate other issues, even mobile payment solutions.

There are no significant differences between active users and inactive users on the other factors. However, both of the two groups of users show high level perceived risk on group formation and GPRS cost. Relatively high average duration (around 5 days) also reflects that several hours to form a buying group is regarded as a high-risk task by initiators, even though the group buying information can reach the potential joiners in several minutes. The relatively low requirement on buyer numbers (6.5) if compared with the final joiner number (19.6) also reflects the fact that the initiators would like to assure the possibility of successfully completing the group buying by setting the lower threshold.

5.3.3.5 Comparison Between Mobile Group Buying and Online Forum Group Buying

Table 5.6 shows the overall comparison between online forum and HappyGo on some group buying attributes.

There are no significant differences toward number of groups and average buyers per group. The registered users of the online group buying forum is around 50,000, which is much higher than the potential buyers using the HappyGo service, since only 401 users installed HappyGo in the office building. So the result of differences on number of groups and average buyers indicates that the participation degree of

HappyGo users is absolutely higher than that of online forum users. The Hawthorne effect (Jones 1992) may be one reason for the high participation rate of HappyGo. In addition to this, the trial use context of HappyGo, an office building, is also another important factor; since participants are colleagues, the high level of trust towards each other enhances their participation intentions.

One big difference between group buying with online forum and HappyGo is that initiators did not set a required number of joiners. We went through the group descriptions and related messages on online forums carefully, and found most initiators were professional brokers or sellers. They usually have physical shops near the community and have different customer resources, so that they don't need to set the required enrollment. While most initiators of HappyGo are consumers, they themselves are the buyers of the products, so they have to keep the agreement with the sellers or brokers towards the pre-negotiated required buyer number.

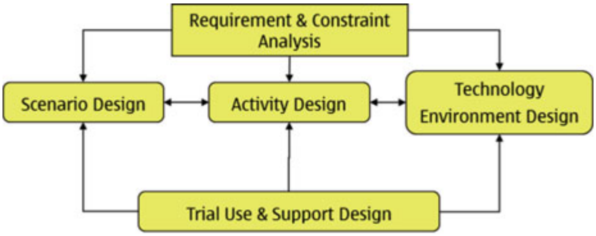
Another outstanding difference is that the duration of group buying activity with online forum is much longer than the one with HappyGo. The HappyGo initiators set an average enrollment period as 6.6 days, while online forum initiators set an average 207 days. The difference of initiators may lead to the difference of group buying duration also. Most HappyGo initiators are product buyers themselves; they have their own requirements with regard to waiting time. However, for most online-forum initiators, the main concern is the quantity of buyers. For them, longer duration may mean more enrollments. But another reason for this big difference may be that the mobile group buying service allows initiators to find more potential buyers in a relatively shorter period. With the HappyGo service, the updated group buying information can be sent to relevant users' mobile devices in an immediate way. Furthermore, it is easier for users to keep an eye on the group buying activities with a handy mobile device than with a fixed PC.

5.4 Mobile Social Service Design Framework

There is an international standard that is the basis for many UCD methodologies. This standard (ISO 9241–210:2010: Human-centered design for interactive systems) defines a general process for including human-centered activities throughout a development life cycle, but does not specify exact methods.

As we discussed in the former two cases, mobile social service design should absorb strategies and principles of user-centered design, and rely on the features and advantages of mobile and wireless technology. Human, technological, and social activity are the three interwoven issues during the service design process. To guide the practice, we tried to embed the key components and thoughts in a general framework for mobile social service design (Fig. 5.10). The core aspect of the framework is the social activity design, which is not only based on requirement and constraint analysis (influence factors analysis) and refines the results of the use scenario design, but also specifies the technology environment design and the trial use and related consumer support services design.

Fig. 5.10 Design framework of mobile social service



5.4.1 Social Activity Design

All social services for some special context involve some form of activity, for example, exhibition information collection and sharing, and group buying. The social activity supported by the mobile technology can be defined as specific interactions of users with other people, using wireless and mobile technology-enhanced tools and resources oriented towards specific outcomes. The social activity design is to clarify those mutually inter-dependent elements: user objectives, tasks, process, activity rule, form of organization, resources, or tools. There are mainly three guidelines in designing the social activities.

1. Base the activity on categories of goals
   In requirement analysis or scenario design, the scopes of user goals are determined to some degree, which influences directly the tasks and process design. Social scenario and requirement analysis may be ineffective if it has no connection or not been validated by a social activity.
2. Foster and maintain motivations for social activities
   Compared to traditional mobile information service, mobile social service requires more readiness of users to use. In this way, mobile users also require more stimuli to activate and encourage persistence.
3. Align the social activity with the supported technology role
   Social activity significantly influences the technology used; at the same time, basing our activity design on the understanding of the technology role, designers can achieve informed trade-offs.

5.4.2 Requirement and Constraint Analysis

From a procedural point of view, all the items of requirement analysis can be done at two levels: general level and concrete level.

The analysis tasks at the general level include:

1. Common features of the social activity
2. Current state and development trend of ICT

3. Potential end users and existing applications/services
4. Motivations and expectations from the project team

After the general requirement analysis, the designers can build some presumptions and a future user study plan. In concrete requirement analysis, questions such as who are the users, what are their needs, how can these needs be satisfied with certain support activities, etc. should be answered to some degree. So the analysis tasks at the concrete level include:

1. Possible situation, environment, and influencing factors
2. Social–cultural features of user organization or group
3. Characteristics of potential users
4. Tasks, motivations, and possible difficulties or barriers
5. Users' attitudes, skills, experiences, and use patterns toward mobile and wireless technology

In the design process proposed by Mike Sharples et al. (2002), constraints can be mixed up with requirements, since requirements provide constraints on the design; for example, in the general requirement analysis, time and budget available for the project could be regarded as constraints.

5.4.3 Scenario Design

A scenario (Blomberg et al. 2003) is a narrative story about specific users, their activities, and the environment or situation, which has the following characteristic elements: setting, agents/actors, goal/objectives, events. A social activity scenario has to describe how users with certain characteristics in certain settings carry out various activities to achieve their specific goals. A scenario can help designers and developers imagine and understand for whom, how, and where the mobile activity and the supported technology are to be provided.

5.4.4 Technology Environment Design

The technology environment is the conditions such as community platforms and related tools, networks, etc., that support and sustain the social activities.

5.4.5 Trial Use and User Support Services Design

Trial use can help designers find the potential problems encountered by our target users and their acceptance procedure. The user support service means a range of services enabling users to overcome difficulties, and to develop competencies and confidence in adopting the new social services.

5.5 Conclusion

Two cases about mobile social service design for special context are discussed. Unlike the general social network service through which people interact with each other mainly for social purposes, the social service for special context supports a temporary social network to conduct some activity together. Although people in the special social setting are mainly motivated by individual benefits, the relationships are also influenced or enhanced by the social activities in the special context. The two design cases discussed in this chapter are large-scale exhibition service design and local group buying service design.

Many studies have been done on the utilization of mobile technologies to enhance social interaction among visitors in the museum context. Compared with the museum context, exhibition, especially those which are open to the public and usually large scale, are more complex. User data from interviews, questionnaires, and field studies have been analyzed. Exhibition visitors do have some motivation to use mobile social services, but at the same time, the barriers to adoption are obvious also. Concept design based on the insights gained from the user study was conducted to explore how the mobile technologies can be applied to promote social interactions in such a context.

As for local group buying, the joiners, the initiator, and sometimes even the merchants are in the same local community. Such locality induces some new characteristics into group buying. A mobile service to conduct local group buying, HappyGo, was developed, and a field study of HappyGo service combined with comparison with an online group buying forum was conducted to evaluate the service and explore the factors influencing user experiences.

Based on the two case studies, we summarized an activity-oriented design framework for mobile social service design, which mainly includes five elements. The core element is the activity design, which determines directly the experience in the special context and includes the specification of all the activity components. The element of 'requirements and constraints analysis' mainly investigates relative users, context, business environment. The third one includes creation and repeated evaluation, refinement of social scenarios. The fourth element mainly translates the understanding of activity and expectation of technology role into a realistic technology solution that supports the whole experience. Finally, the user support services utilize various means to bridge the distance between individual goals and activity supported by mobile and wireless technology with limitations.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179–211.

Bagozzi, R. P., Davis, F. D., & Warshaw, P. R. (1992). Development and test of a theory of technological learning and usage. *Human Relations, 45*(7), 660–686.

Blomberg, J., et al. (2003). An ethnographic approach to design. In J. A. Jacko & A. Sears (Eds.), *The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications* (pp. 964–986). Mahwah: Lawrence Erlbaum Associates.

Boehner, K., et al. (2005). Imprints of place: Creative expressions of the museum experience. In *Proceedings of CHI 2005 extended abstracts* (pp. 1220–1223). Portland, Oregon, USA.

Cosley, D., et al. (2008). ArtLinks: Fostering social awareness and reflection in museums. In *Proceedings of CHI 2008*. Florence, Italy.

Counts, S., Hofte, H. (2006). Mobile social software: Realizing potential, managing risks. In *CHI2006 workshop*, 22–27 April. Montréal, Canada.

Garzotto, F., & Rizzo, F. (2007). Interaction paradigms in technology-enhanced social spaces: a case study in museums. In *Proceedings of the 2007 conference on Designing pleasurable products and interfaces*. Helsinki, Finland.

Guo, J. (Ed.). (2007). *The report on development of China's convention & exhibition economy in 2006–2007*. Beijing: Social Science Literature Press.

Guo, S., Wang, M., & Leskovec, J. (2011). The role of social networks in online shopping: Information passing, price of trust, and consumer choice. In *Proceedings of ACM EC 2011*. San Jose, CA, USA.

Hong, D., Chiu, D.K.W., & Shen, V. Y. (2005). Requirements Elicitation for the Design of Context-aware Applications in a Ubiquitous Environment, *In Proceedings of the 7th international conference on electronic commerce (ICEC)* (pp. 590–596). Xi'an, China: ACM publications.

Jones, S. R. G. (1992). Was there a Hawthorne effect? *American Journal of Sociology, 98*(3), 451–468.

Liu, H., et al. (2012). HappyGo: A field trial of local group buying. In *Proceedings of CSCW2012*, Seattle.

Millen, D. R. (2000). Rapid ethnography: time-deepening strategies for HCI field research. In *Proceedings of the 3rd conference on designing interactive systems: processes, practices, methods, and techniques* (pp. 280–286). New York.

Petrelli, D. and Not, E. (2005). User-centered design of flexible hypermedia for a mobile guide: Reflections on the hyperaudio experience. *User Modeling and User-Adapted Interaction, 15*, 303–338.

Sharples, M., Corlett, D.,& Westmancott, O. (2002).The design and implementation of a mobile learning resource. *Personal and Ubiquitous Computing, 6* (3) pp. 220–234.

Stock, O., & Zancanaro, M. (Eds.). (2007). *PEACH-intelligent interfaces for museum visits (Cognitive Technologies)* (pp. 269–288). Secaucus, NJ, USA: Springer-Verlag New York.

Svensson, M., Hook, K., & Coste, R. (2005). Designing and evaluating Kala: A social navigation system for food recipes. *Transactions on Computer–Human Interaction, 12*, 374–400.

Tan, W. K., & Tan, Y. J. (2010). Online or offline group buying? In *Proceedings. of IEEE FSKD* (pp. 2853–2857). Yantai, Shandong, China.

Chapter 6
Exploiting Personal and Community Context in Mobile Social Networks

Daqing Zhang, Zhiyong Yu, Bin Guo, and Zhu Wang

Abstract Mobile social networks (MSNs) are believed to be more user-friendly and intelligent than online social networks. In this chapter, we first extend the definition of mobile social networks by classifying MSNs into four categories, and define two important terms, e.g., *personal context* and *community context* in the emerging field of mobile social networks. We then present the context model and the related taxonomy of personal context and community context. We further divide the life cycle of MSNs into four phases — *discovery, connection, interaction, and management* — and elaborate how personal context and community context facilitates the process in each phase. Three major data sources for deriving personal and community context in MSNs are identified, e.g., *sensor-rich mobile and wearable devices, Internet applications and Web services, and static infrastructure*. Leveraging the three data sources, techniques ranging from data representation, data cleansing, and data anonymization to clustering techniques and inference techniques are presented for inferring personal and community context. Finally, future research directions and challenges are identified, in order to shed light on next-generation MSN development from the context-aware perspectives.

D. Zhang (✉) • Z. Yu
Institut Mines-Telecom/Telecom SudParis, 9, rue Charles Fourier,
91011 Evry Cedex, France
e-mail: daqing.zhang@it-sudparis.eu; zhiyong.yu@it-sudparis.eu

B. Guo • Z. Wang
School of Computer Science, Northwestern Polytechnical University, Xi'an, China
e-mail: guobin.keio@gmail.com; transitwang@gmail.com

A. Chin and D. Zhang (eds.), *Mobile Social Networking: An Innovative Approach*,
Computational Social Sciences, DOI 10.1007/978-1-4614-8579-7_6,

6.1 Introduction

Mobile social networking (MSN) is rapidly becoming a new research domain and killer application to showcase the power of merging social networking and mobile computing. It is believed that MSN will not merely be a simple extension of online social networking, it will *revolutionize social networking* by enabling *anytime anywhere social interaction* and *a higher degree of intelligence*. The former is achieved by the inherent property of feature-rich mobile devices via wireless communication, while the latter is made possible by exploiting the comprehensive users' *context*, extracting information from rich data sources such as mobile/wearable devices, social networking sites and services, Internet applications, and surrounding sensor networks (Zhang et al. 2011).

Although online social networking and social network services have gained great success in attracting hundreds of millions of users in such a short time to socialize in the virtual world, none of these current social networking sites and services do actually tab into the huge amount of context data and information provided by the community users who move and interact constantly in the physical world and digital world using sensor-rich smart phones (Menkens 2009). Motivated by the observation that the explosive growth of social networks such as Facebook and Twitter, the mushrooming popularity of mobile smartphones such as iPhone, and the rapid evolution of sensor networks provide an unparalleled opportunity to achieve a more comprehensive understanding of the context surrounding a user or user community in nearly any given environment (Beach et al. 2010, Zhang et al. 2010), this chapter intends to investigate what *new facets* of context are crucial in mobile social networking, and how context-awareness can help in shaping the future mobile social networking paradigm.

This chapter begins with a simple taxonomy of MSNs to broaden the definition of mobile social networks. By classifying MSNs into four categories, mobile social networking taking place in both the *physical world* and the *virtual world*, for either *spontaneous* or *long-term interaction,* is included, leading to new perspectives about future mobile social networking. By merging pervasive computing and mobile social networking, the most significant benefit is to bring *awareness* of context about people, places, resources, and services to social networking. Specifically in the context of mobile social networking, this work explores what context is crucial in facilitating the *creation, discovery, interaction*, and *management* of the mobile social community, as the new facets of context for MSNs. In particular, we present two terms, i.e., *personal context* and *community context,* the two corresponding context taxonomies, as well as how those contexts are related to each phase of an MSN's life cycle in the section entitled "Personal Context and Community Context". Following the definition and detailed elaboration of both *personal context* and *community context* in that section, in the section entitled "Context-aware Discovery, Connection, Interaction, and Management in MSNs" we proceed to explore how personal and community context can be used in four different phases of the MSN life cycle, i.e., entity discovery, entity connection, community interaction, and management. In the remaining sections of this chapter, we further study what the possible data sources that can be leveraged to infer the relevant personal and community context are, what the context management framework for MSNs looks like, what the popular techniques that can be used to extract those context are, and what the possible future MSN technology and research challenges are. In summary, this chapter attempts to answer the following questions:

1. What is MSN? Why is context-awareness important for future MSN?
2. What roles do personal context and community context play in MSNs?
3. What is personal context and community context? What are the new facets of context in MSNs?
4. What are the data sources that can be leveraged to extract the relevant personal and community context?
5. What are the popular techniques for inferring personal and community context?
6. What are the research challenges in enabling future context-aware MSNs?

6.2 Mobile Social Networks and Social Community

6.2.1 Mobile Social Networks: Definition and Classification

The advent of online social networking applications and services have radically changed the way people interact. More and more people are regularly using online social networking services such as Facebook, LinkedIn, etc. to form communities and socialize in virtual spaces. With the quick penetration of mobile phones, especially sensor-equipped smart phones, a current trend for online social networking services, such as Facebook, is to create mobile applications to give their users instant and real-time access from their devices. In turn, native mobile social networks have been created such as Foursquare and Gowalla, building communities around mobile functionality. More and more, the line between mobile and web is being blurred as mobile applications use existing social networks to create native communities, and web-based online social networks take advantage of mobile features and accessibility. This has led to the following two trends:

1. Online social networks being extended for mobile access through mobile browsers and smartphone applications
2. Mobile social networks leveraging sensed user activities, location, profiles and generated contents by online social networking services

On one hand, people can form a social network anytime, anywhere, which removes the constraints in space, and people do not need to be bound to desktop PCs. On the other hand, the computing and sensing capability in mobile devices and environments makes it possible to capture the context and situation of both individual and group, which can be leveraged to facilitate the creation of social communities and interaction among mobile community users. In such a way, the mobile social networking can be enabled virtually and remotely just like online social networks; it can also take place to support spontaneous and face-to-face interaction.

According to the definition in Wikipedia, mobile social networking is defined as "*social networking where individuals with similar interests converse and connect with one another through their mobile phone and/or tablet*". Considering the existing online social networking services such as Facebook, LinkedIn, etc. at one extreme, and mobile ad hoc social network at another extreme, we can simply classify MSNs into the following four categories according to the time span of the community and distance among community members:

- *Short-term short-range MSNs*. This category corresponds to the MSNs in physical proximity for a short time duration. Those mobile social networks, also called spontaneous MSN (Mani et al. 2009), can be built based on short-range ad hoc wireless communication protocols such as Wi-Fi and Bluetooth; they can take place in a room, building, bus stop, public space, or even on a moving ship or train. In this case, the life span of such a mobile social network is short and the distance among mobile community members is short. The objective of this type of MSN is to facilitate face-to-face interactions or information sharing/collaborative work in an event or certain physical space (for both friends and strangers).
- *Short-term long-range MSNs*. This category corresponds to the MSNs formed by remote community members for a short time duration. Those mobile social networks are formed to fulfill a specific task for a short time duration (for instance, remote volunteer support for a disaster event). People are pulled together no matter where they are, members are already linked through an online social network. In this case, the life span of such a mobile social network is short and the distance among mobile community members could be big. The objective of this type of MSN is to facilitate information sharing/collaborative work in the virtual world and among connected people to fulfill a task.
- *Long-term long-range MSNs.* This category corresponds to the current online social networking services (such as Facebook) extended to mobile devices. In this case, the life span of such a mobile social network is long and the distance among mobile community members could be big. The objective of such extended online MSNs is to facilitate anytime, anywhere information sharing/collaborative work in the virtual world.
- *Long-term short-range MSNs*. This category refers to the MSNs confined to a fixed group of people in a specific physical environment. This is a special form of mobile ad hoc social networks for a fixed physical space. In this case, the life span of such a mobile social network is long, and the distance among mobile community members is short. The objective of this kind of MSN is to facilitate information sharing/collaborative work among community members familiar with each other (people belong to one family, one company).

As we can see from the above-mentioned MSN taxonomy, the mobile social community can be either created from *scratch* or from an existing bigger community, based on a common goal or common interests of mobile users. For example, a short-term short-range MSN needs to be created from scratch, it helps to facilitate face-to-face encounters or spontaneous interaction with people who are *not necessarily part of their social network*. While the short-term long-range MSNs promote interactions among remote users to form communities for a specific task (such as finding an expert to solve a problem), this mobile social community needs to be detected or created within an existing social community, as in the case of Facebook or working groups in a company. In both cases, the community creation needs to understand each user's personal context, be it user location, user interest, or expertise. After the mobile community is created, other members can choose to join through community discovery and matchmaking mechanisms.

As for long-term long-range MSNs and long-term short-range MSNs, the mobile social community is already formed among existing contacts. The major difference between these two categories of MSNs lies in the fact that they are suitable for supporting different type of social interactions. For example, a long-term short-range MSN can support interactions in both the physical and virtual worlds, it might have a special privacy policy during information sharing and interaction due to the high trust level among close community members (among family members or colleagues). While a long-term long-range MSN is more a mobile extension of current online social network. In both cases, user interaction and resource sharing can be enhanced by exploiting each member's personal context, such as user availability, schedule, or resources owned.

6.2.2 Mobile Social Community and Life Cycles

The term "community" describes groups of people who identify themselves with a common goal (often reflected in common interests) and who have the means to communicate with each other and collaborate around the common goal. The community formed by mobile users for social networking is called a *mobile social community*; it offers a context for people to meet and communicate in a physical and/or virtual space. As presented in the last section, a mobile social community can be either created manually or automatically based on certain predefined conditions from scratch; it can also be formed within a bigger social community according to certain context or by an entity, while users can either create a mobile community, or are invited to participate a community, or automatically become community members by meeting the matchmaking criteria.

As the life cycle of *mobile social community* consists of community creation, operation (interaction, merging, splitting), and termination, inspired by the proposed three-step processes (i.e., discover, connect, and organize) of community management in the EU FP7 Societies project (Roussaki et al. 2012), this chapter proposes to divide the major operation of mobile social community into four phases: *discover, connect, interact,* and *manage*. The key functionalities of each phase are described as follows:

- *Discover*: discovery of people, devices, networks, and services for forming new communities or for matchmaking with requirements of existing communities; discovery of existing communities for joining, merging, splitting; discovery of context data sources for context extraction.

- *Connect*: connecting individuals to form communities; connecting individual or communities to communities; connecting individual or community to devices, networks, and services; connecting context data sources to context inference services.
- *Interact*: direct interaction among community members via text, voice, social media; indirect interaction by tagging the same set of photos, commenting on the same objects or places, visiting the similar places or following similar trajectories, etc. Propose activity or topic for interaction, recommend users or resources for interaction.
- *Manage*: adding individuals to communities or removing individuals from a community; forming, merging, splitting, and deleting communities; managing interactions among community members; managing devices, networks, and services in a community; managing current and historic context of individuals and communities.

Apparently, in order to support intelligent decision-making in each phase, personal context and community context are the key to automatically enable the process.

6.3 Personal Context and Community Context

6.3.1 Context-aware Computing

Context-aware computing was developed as an important research branch of pervasive computing starting in the late 1990s (Schilit et al. 1994, Dey et al. 2001). Its objective is to make the pervasive systems and services more intelligent by considering relevant information which was not taken into account in the system design. Early efforts in context-aware computing research focused mainly on context modeling and middleware support for specific applications; examples include context-aware mobile tour guide (Abowd et al. 1997), context-aware smart home (Zhang et al. 2003), and context-aware healthcare (Zhang et al. 2005). The context model proposed is usually about *individuals* in *smart spaces* for *certain applications*. The state-of-the-art context model is the two-tier ontology-based model (Wang et al. 2004a) which was originally designed to characterize the situation about mobile users and the relevant entities in a smart space, where the high-level context entities include *user, location, computing entity*, and *activity*. Properties of these entities, as well as the relationships between them, form the skeleton of a general contextual environment.

With the development of socially aware computing proposed by Alex Pentland (2005), the social aspect of context caught the attention of the research community. The focus of context-aware computing was shifted from understanding user information such as where they are, what they are doing, etc., to recognizing who they are with, how they interact with each other, etc. With the increasing popularity of social networking websites and social media, the notion of social context is extended to include the set of information arising out of direct or indirect interactions among people in the social networking services. The social aspect of context typically includes user social preference information, user social relations, and user's social role and interactions, as well as user's social situation. This work intends to exploit the social aspect of personal context to help in forming new communities and enhancing interactions among users of existing communities.

With the main process of *mobile social community* dividing into four phases — *discover*, *connect*, *interact*, and *manage* — it is obvious that the first two phases are related to mobile community creation, discovery, or detection; and the last two phases are related to operations after the mobile community is formed. In the community creation stage, the social aspect of the personal context is needed and used to form relationships among users; such context could be location, proximity, historic relationship, common goal or interest, etc. If the discovery stage refers to detecting existing communities to join, or looking for individuals to invite, then the community context or personal context is needed in order to filter out the irrelevant search. In the interaction stage, the objective is to facilitate effective interactions among community members; in this case, both personal and community context can play a crucial role in making the interaction efficient and intelligent. For example, if one member has a certain behavior pattern, then this pattern can be used to choose the right time for interaction; if two members are working on the same document, they might be reminded to have a direct communication channel for discussion. For the management of a mobile social community, both personal and community context are important for decision-making. For example, if a member is detected to be inactive and located in a different place for a certain period of time, she might already have left the community; and when two communities are found to have similar goals with a big number of overlapping members, these two communities could be advised to merge into one community. In the following subsections, both personal context and community context will be modeled and presented in detail, all from the perspective of effectively supporting different phases of MSNs.

6.3.2 Personal Context Taxonomy

The personal context of a mobile community user is defined as all the relevant information that can be accessed to characterize the situation of an individual. In this work, personal context is perceived as part of a process the mobile user is involved in. So the personal context of a mobile user includes all information that can be acquired and inferred at the current point in time, as well as all historical context data. As some future states of a user can be predicted from current and historic observations, we can thus classify personal context into *past context*, *present context*, and *future context* along the time dimension, in consideration of the differences in their acquiring and storing methods. In each phase, the personal context can be further classified into two categories: *static context* which doesn't change (or changes slowly) and *dynamic context* which changes often. Therefore, taking the present personal context as an example, a taxonomy is given as follows:

Static Context

- Identity: the user identity refers to the information that identifies who a person is, i.e., name, age, gender, nationality, address, education, occupation, etc., as well as their biological characteristics, such as weight, height, finger print, iris.
- Preference: personal preferences could be conceived of as an individual's view about specific objects, typically reflected in an explicit decision-making process. What we care about is their preference in social activities, e.g., a person likes a certain restaurant, dislikes a certain movie, etc.
- Resource: personal resource refers to devices, network, and services that a person can use. Device capability includes screen resolutions, processing power, graphical representations, audios, input style, and so on. Network parameters include bandwidth, delay, access technologies, network addresses, etc. Services subscribed to or used by a user are also important resources.
- Affiliation: this information represents details about a user's role, as well as the list of communities a user is involved in. Examples could be their role in a university, such as lecturer, professor, student, visitor, etc.
- Contact list: usually a person has a contact list in the mobile phone, SNS, e-mail, etc. It contains addresses of their family members, friends, classmates, colleagues, acquaintances, and so on.

Dynamic Context

- Time: time is the most important information associated with dynamics. The key time attributes include starting time, ending time, and the duration of an event. It can be at different granularities, such as month, week, date, minute, second, and with different semantic meaning, such as birthday, break time, etc. And there are other time-related factors such as time zone.
- Location: personal location refers to where an individual is. Location also can be at different granularities with different semantic meanings. For example, user A is in Paris, user B is playing soccer in a field located at the coordinate (48.6338, 2.4346). There are also attributes associated with a certain location, such as the weather, temperature, light conditions, noise level, etc.
- Physiological condition: this refers to the condition or state of the bodily functions of a user, e.g., heart beat rate, blood pressure, body temperature. Those parameters can be used to infer their activity or emotion.
- Behavior: behavior is a single action performed by an individual in a short time. For example, walking, sitting, typing, calling, laughing, holding an object, etc.
- Activity: personal activity, which is performed for a relative longer time, is the high-level abstraction of a person's behaviors, e.g., working, sleeping, meeting, dinner. Their situation in terms of such things as emotion (e.g., happy, sad) or presence (e.g., busy or not) defines their willingness and ability to engage in a certain activity.
- Intention: Personal intention is what a user wants to do in the near future. It can be viewed as their temporary mental state for a task or activity.

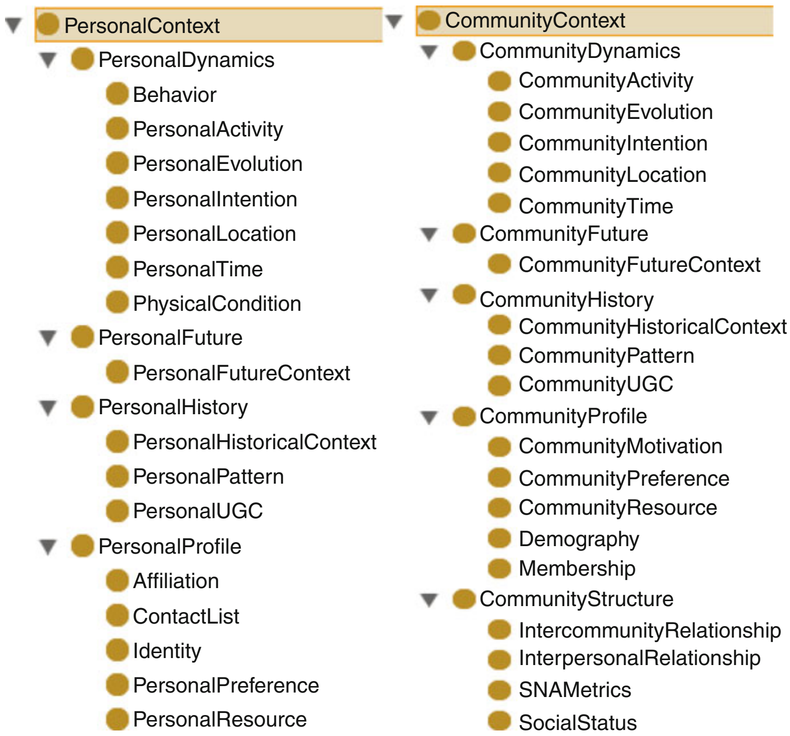


Fig. 6.1 Personal and community context taxonomy

While dynamic personal context changes from time to time, the static personal context such as a person's identity, preference, resource, affiliation, and contact list also evolves along with time. The personal context taxonomy can be found in the left panel of Fig. 6.1.

6.3.3 Community Context: Definition and Taxonomy

Generally speaking, "community context" refers to the information about or description of a community. In early years of this research, community context was mainly used as a terminology in the organizational psychology field (McArthur and Bruza 2001; Mohammed and Dumville 2001), and referred to shared mental models whose values remain stable, thereby providing the secure foundations on which the community could be built, e.g., rituals, social conventions, language, policies, rules, and protocols. Recent work about community context is mostly in the CSCW field for content recommendation or knowledge sharing (Klamma et al. 2006; Kleanthous 2007), and community context is seen as community preferences, tasks, activities, location, time, device capabilities, and so on.

In order to support the community lifecycle (discover, connect, interact, and manage) in MSN, *community context* is defined as *any information about the whole community and subset of members* which helps communities to function effectively by understanding and exploiting the interests, behaviors, and relationships of community members and the community as a whole (Koch et al. 2002). Again based on the temporal and volatile nature of context, we classify the community context into the following categories: *static* (includes profile and structure), *dynamic, historic, present*, and *future* context (see Fig. 6.1). In the right panel of Fig. 6.1, we show the community context taxonomy which includes community profile, community structure, community dynamics, and historical context as well as future context.

Community Profile

- Community motivation: this refers to the common goal, purpose, or reason for forming the community. For example, a soccer-loving community would aim to share news about soccer matches, players, coaches, etc. A discussion community would intend to discuss matters related to the members.
- Membership: this is to describe who belongs to the community, i.e., nodes of the MSN. Each community has its size. There are two kinds of membership: *clearly defined ones*, such as a research team, a class of students, and *coarsely defined ones*, such as a group of friends.
- Demography: this is obtained by simply computing the statistics of the community members, e.g., the composition or distribution of members' age, gender, nationality, education, occupation, income, etc. For example, a student community.
- Community preference: community preference is about the attitude of the group in social activities. While each member could have his/her own preference about a restaurant, a movie, and so on, the community preference is not a simple average of all members' preference. When the members differ in personal preference, there needs to be a way to identify the community preference.
- Community resources: as each community member owns some personal devices and services, and some members are willing to share part of these resources inside the community, community resources refer to those that are shared among community members.

Community Structure

- Interpersonal relationship: this describes links of the MSN in different forms, i.e., Boolean, type, strength, orientation. Boolean relationship is the most popular one, which indicates two members do or do not have a relationship. Typed relationship might imply that member A and B are from the same university, while A and C are from the same hometown. The strength of relationship characterizes how close the relationship is between two connected nodes. The orientation of relationship is used in asymmetric MSNs such as Twitter, e.g., A follows B, but B does not follow A.
- SNA metrics: SNA (social network analysis) can be adopted to exploit many useful metrics, e.g., connection, distribution, or segmentation. There are some overall measurements of links, e.g., network closure, centrality, density, distance, clustering coefficient, and cohesion.

- Social status: this characterizes the social influence or the role of a member in the community, such as that member A is the leader of the group, B and C are peers. We can see that social status is an organizational feature; sometimes it is represented by links but sometimes not.
- Intercommunity relationship: It is the structural information among multiple communities, i.e., whether two communities are hierarchical or overlapping. For example, a university contains many colleges and a college consists of several departments, and they are hierarchical. The soccer lover community and a class of students might have some members in common, so they are overlapping.

Community Dynamics

- Time: time records when a personal activity or community event occurs. To characterize the relationship between personal events and community events, time synchronization is needed.
- Community location: most of the time, community location is where most of the members stay. For example, the soccer lovers play soccer on the field. However, for a long-range community, the members might not be geographically close to each other; the community location can be expressed by the distribution or fusion of the members' locations.
- Interaction: interaction refers to information exchange among community members or communication initiation from one member to another (others). It can last for seconds to minutes depending on the community activities engaged. Interaction-related information includes involving parties (who is interacting with whom), modality (e.g., face-to-face talk, phone call, e-mail, instant message, SNS, online or physical item exchange), objectives or activities (e.g., propose an opinion, agree, like, acknowledgement).
- Community activity: community activity is the high-level abstraction of a sequence of interactions among community members for a relatively longer period of time, usually ranging from minutes to hours. e.g., meeting, dinner, party. The related information about community activities include members' presence (i.e., busy or not), the active level of members (e.g., member A seldom speaks in meetings), leaders (e.g., two debate for a long time), participation strategy (e.g., selfishness or altruism), emotion of the community (e.g., happy, sad).
- Community intention: community intention is what the community is formed for. It can be the common goal or motivation of the community based on each member's motivation or preference. For example, the objective of group of music lovers is to enjoy music together, and their preference contains pop music and rock music; they all agree to listen to a song by Michael Jackson given a list of candidate songs.
- Community evolution: communities evolve with time, due to a change in members' profile, community structure, motivation, people, demography, preference, resource, interpersonal relationship, SNA metrics, social status, and intercommunity relationship. It is interesting to understand the dynamics of long-term communities, lasting for months or years.

Community Context History

– Historical context: this refers to the history of previous profile, structure, and dynamics. Information in the past can be used to understand users' current and future context.
– Pattern: pattern is derived from historical contexts. It might include interaction pattern (such as frequent interaction modality, e.g., infrequent face-to-face talk while writing many e-mails, frequent sequence, e.g., always make a call before a face-to-face talk), personal or collective behavior pattern (e.g., always do the same thing or go to the same place), personal or community evolution pattern (e.g., whenever the activity frequency decreases, some members will quit the community).
– UGC (user-generated content): every activity will generate certain digital traces or contents, such as pictures, videos, message records, meeting minutes, comments, blogging, and wikis. All these contents might be utilized to support community lifecycle management in a specific way.

Future Community Context

– Future context: this describes what personal or community profile, structure, and dynamics would be like in the future. Usually these contexts are predicted based on historical and current contexts.

6.3.4 Sensed Raw Context, Inferred Context, and Historic Context

For both personal context and community context, they can be either static or dynamic, can be directly sensed from various data sources or indirectly inferred by different rules and algorithms. Usually simple context items such as location, temperature, available networks and services, etc. are sensed or acquired directly, but complex context such as a user's intent and the relationship among community members need to be inferred with algorithms or rules.

As individuals and communities generate large amount of data continuously, the data streams are gathered as the real-time context data source. With real-time data streams, raw context is obtained and stored in the context knowledge base for personal and community context query. With the passage of time, the accumulated record of data and context becomes historic context, and the complex context is generally inferred through processing both the real-time context and historic context. In particular, a complex community context needs to be inferred through processing the real-time context of relevant individuals and communities as well as the historic context of individuals and communities. In such a way, both the personal context and community context, either real-time sensed context or inferred context, are stored in the context knowledge base for query and further inference, and this context information becomes historic context with the passage of time. In the following sections, we will explore what the possible context data sources are for extracting personal and community context, what the popular techniques are for inferring context, and what the context management framework is which supports context acquisition, aggregation, reasoning, prediction, and query.

6.4 Context-aware Discovery, Connection, Interaction, and Management in MSNs

6.4.1 Context-aware Discovery

In mobile social networking, in order to create/form a mobile social community or detect a community for joining, the first task is to discover the relevant people, communities, devices, network, and services, crossing the boundary between the physical and virtual world. In the case of community creation, the objective is to find the right people with the appropriate devices and services in a shared networking environment; in the case of community detection and discovery, the objective is either to find a group of people with certain relationships who share a similar goal or common interests to form a mobile community, or to discover existing social communities using a service discovery mechanism for joining purposes.

Existing systems already allow the discovery of people in a physical environment via a wearable RFID tag worn by an individual, a mobile device with a permanent MAC address, or a search in the web or social networking sites (virtual environment). There are also device and service discovery protocols such as Bluetooth, Wi-Fi, UPnP, Jini, etc. which support discovery of user devices and virtual services (including a community advertisement service). However, these systems currently do not fully exploit the wide variety of context information in order to discover people across the boundary between virtual and physical spaces for social networking purpose. This work intends to make use of various contexts for people, resources, and community discovery, with the objective of filtering out non-relevant entities and locating merely those which are relevant. As the mobile social community is pulling people and resources together around a common intent, co-location, or similar interests, thus the most significant context should include individual's current or future location, personal interest and preference, intent or activity, community's location, objective, and profile. For example, we can discover people based on static personal preferences in social network services such as Facebook, LinkedIn, Google+, etc.; we can search a community of people with a common goal or in a similar situation regardless of whether they use social networks or not. Thus, the complex context inference techniques play a very important role in deriving a user's intent, logical location, social relationship, a community's interest and preference, future mobility pattern, and availability.

6.4.2 Context-aware Connection

Building on the selected list of relevant people, devices, services, and communities discovered, the next task is to either select and connect relevant entities across the physical and virtual worlds creating a new mobile social community, or to select and connect to the appropriate social communities participating in the community activities they are interested in. This allows for the establishment of communication mechanisms between these entities, enabling each of them to capitalize on the capabilities of the entities it is connected to, in order to share whatever resources and services each entity might want to use. Thus, groups of users and entities that demonstrate commonalities connect, creating mobile social communities.

There are many context-based criteria that can be used to form new mobile social communities or connect to appropriate existing communities (Roussaki et al. 2012), such as:

- Sharing the same geographic location (current location, or location/area of residence or work, etc.), including groups of persons that are often co-located (e.g., family members, colleagues, fellow students, classmates, residents of the same building, neighbors)
- Having the same or similar preferences
- Sharing common interests or features (personal, business-related, service-related, hobbies, etc.)
- Sharing a common belief, idea, intent, or goal
- Sharing common experiences or background or knowledge
- Having ties of friendship, kinship, communication/interaction, authority/hierarchy, trust, shared membership, or other forms of social relationships, etc.

6.4.3 Context-aware Interaction

Once the relevant entities (people, devices, services, networks, etc.) have been discovered and connected in a mobile social community, one key task is to facilitate personalized and effective interactions among community members. In this phase, each member's personal context and the community context play an essential role in enabling intelligent information exchange and resource sharing.

The greatest challenge in the *interaction phase* is to identify the situations and events that should trigger interactions among a subgroup of members, between two sub-groups, or just between two members. The situations and events vary greatly, depending on the requirements arising from possible interaction patterns. A simple event can range from two members being available for a chat at a certain instant to several members who are all working on a certain document, while a complex event could be to propose a local tour to a group of people based on their availability and interests. In general, as interactions involve at least two parties, thus the relationship (similarity, commonality) between individuals, between an individual and a group, or between groups inside the community is something that needs to be monitored all the time. Even future interaction situations and events are also key things which need to be predicted and estimated, so that proactive actions can be taken to facilitate future interactions.

Another challenge is to identify the *collective behavior of a community* based on each individual's behavior. This is a very challenging and tricky problem, as different strategies have to be applied in different situations. For example, a community's preferred temperature in an air-conditioned room could be the average value of each community member in the room in most cases, but when one member catches a cold, the recommended room temperature for the community might be the preferred temperature of the sick member. Community collective behavior can be very tough to present when members have conflicting interests in certain situations, which is also the case for human beings in a real-world setting. As for the prediction of community dynamics, it is even a tougher problem to tackle, as the evolution and structure of a community depends on many factors, ranging from social to economic to behavior background of each individual member, all affecting the interaction pattern of a community and the members inside the community.

6.4.4 Context-aware Management

After the mobile social communities are created or formed or created, another key task is to manage the lifecycle of the established mobile communities in the *management phase*. This includes introduction of new members to the community, through additional discover/connect cycles, and removal of members who are no longer relevant or no longer willing to be part of the community. Similarly, the great challenge in the management phase is also to monitor and detect situations from various context data sources that would cause the change of the community status, such as new member joining, old member leaving, one community splitting into two, several communities merging into one, …; again the ability to predict the behaviors of users & communities is another challenge in effectively managing the resource usage of users and communities, affecting also the life cycle management of the mobile communities.

6.5 Data Sources for Personal and Community Context

The three main data sources for extracting personal and community context are: *Internet and Web services, static sensing infrastructure,* and *mobile devices and wearable sensors*. The three sources have different features and strengths. Both personal and community context can be directly sensed from various data sources, or indirectly inferred by rules and algorithms. Here we mainly elaborate the raw data sources for basic contexts. In the section entitled "Techniques for Extracting Personal and Community Context", we will describe how to get inferred contexts from basic contexts.

6.5.1 Internet Services and Web Applications

This data source refers to all the information we can get from the Internet. The information itself might be a collection of data from user manual input, static sensing, or mobile sensing, but it ends in the form of information appearing on the Internet. The concrete data sources of this kind include:

- Calendar: online calendars such as Google calendar and to-do lists in Doodle contain information about personal or community intention, and can be used to predict their future activities and presence.
- E-mail: for reasons of privacy, access is only allowed to specific information, and the inferred context can only be used to serve individual needs. For example, a user's contact list can be used to recommend receivers when the user wants to do a certain activity; it can also indicate when and how often a user contacts others.
- Instant messenger: in addition to sending messages instantly such as e-mail, instant messengers such as Skype, MSN, and QQ can reveal other personal information such as online state (offline, online, busy, absent), interaction modality (text, audio, or video), etc.
- Social network site: SNSs such as Facebook, MySpace, Google+, Linkedin, Netlog, etc., contain rich information about the community structure and interrelationship among members. Various types of interactions among users can be found and extracted from these sites, i.e., inviting, connecting, posting, sharing, visiting, and replying. Furthermore, users might create or join certain groups with a specific motivation.
- Blog and micro-blog: Twitter focuses more on news propagation than general SNSs. As leading and following exhibits an asymmetric relationship, thus Twitter is more suitable for investigating one's social influence.
- Check-in: location-based social networks such as Foursquare record and share user's location at any place in real-time through a so-called check-in process. The check-in information contains a user's visit history and location preference during the visit.
- Comment: comments collected from social networking sites such as Yelp also imply the user's preference about commented entities. Those comments could be clicks on a post, ratings on an application, tips on a location, or opinions on many kinds of item, e.g., book, music, movie, restaurant, hotel. This is a good source for inferring personal or community preferences.
- Internet forum: a forum such as BBS (Bulletin Board System) usually attracts users with similar interest. From text messages left in BBS, we can understand users' opinions, interests, and interaction patterns.
- Media sharing site: social media such as pictures (Flicker), videos (YouTube), and bookmarks (Delicious) usually reflect users' preferences. They may be also tagged with keywords and/or geographic information related to the user.
- Wiki: A wiki site has a set of authors who can collaborate on the online content editing. We could know the membership of the community, activity level, and who are dominant in the community.

6.5.2 Static Sensing Infrastructure

This refers to the sensing infrastructure that is installed in houses, buildings, or public areas. This data source is best suited to infer indoor and outdoor activities of individual or group, as well as the environment context about a certain place. The popular static sensing infrastructure includes:

– Wireless local area network: in wireless LAN, we can get the SSID of each Wi-Fi access point and the radio signal strength from each access point. When a device receives Wi-Fi signals from multiple access points, the location of the device can be calculated by triangulating signals from access points with known locations. Given the IP address of the mobile device, the user's geographical location can be roughly estimated.
– Cell tower: this records the cell ID for every mobile phone in its range, thus user ID can be used to localize users nearby.
– Sensor network: this includes surveillance cameras, microphones for sounds, RFID readers for smart cards, fingerprint readers for user identities, pressure-based sensing floor mats for user's indoor locations, and wireless sensor networks for sending environmental temperature, humidity, etc. Once we get videos and sounds of a person or a community, multimedia processing technologies can be used to extract their activities, such as who are present, what are they doing, etc.

6.5.3 Mobile Sensing

This refers to those information we can get from mobile phones and wearable sensors. As the mobile and wearable devices accompany people most of the time, this is a great data source to infer interpersonal physical interactions, personal location and activities, and public environment contexts. The popular data sources include:

– Accelerometer, gyroscope, and compass: the data obtained from those mobile phone sensors can be used for inferring personal activity, e.g., sitting, walking, running, etc.
– GPS: this is widely used for outdoor localization.
– Bluetooth: Bluetooth is not only used for interpersonal interaction (transfer files), but also for sensing proximity.
– Phone call and SMS: these are data sources that record daily interpersonal interactions.
– NFC: near field communication is usually used in object identification, localization, and e-payment. It is best suited to recognize one's activities such as shopping, travelling, etc.
– Camera and microphone: just like their counterparts in static infrastructure, cameras and microphones in mobile phones can record one's pictures, videos, and sounds as well as the surrounding multimedia information. The obtained information also reflects one's interests.

– Wearable health monitoring device: this can monitor an individual's physiological parameters such as heartbeat, pulse, temperature, blood pressure, etc. The information acquired can be used to monitor their health status as well as physiological response in different events and situations.

6.6 Personal and Community Context Framework for MSNs

The Context Management (CM) framework aims to infer and manage individual and community context information, in order to facilitate mobile social community creation and detection, user interaction, and community management. The framework is responsible for acquiring raw data from mobile devices, social networking sites, Internet applications, environment sensors and devices, etc., modeling the collected data and maintaining current and historic context information in appropriate data repositories. Additionally, inference techniques are provided, enabling the extraction of high-level information from raw context data. Management of context information also includes enabling context consumers to retrieve and update current and historic context. Finally, synchronization of the personal and community context repositories is also needed. As shown in Fig. 6.2, the context management framework consists of six layers, which are described below in detail.

6.6.1 Pervasive Sensing Layer

The pervasive sensing layer involves the three major data sources identified in the last section: *mobile and wearable sensors*, *static infrastructure*, and *Internet and Web applications*. The three sources have different attributes and strengths:

– Mobile and wearable sensors are always user-centric; they are good at sensing individual activities, interpersonal interactions, and significant locations.
– Static infrastructure enables the detection of user location and indoor user activities.
– Internet applications and Web services are a major source for extracting user profile information and the relationship among users in a community.

Due to their diverse features, aggregation and fusion of data from these three different sources provide unique opportunities for personal and community context extraction in MSNs.

6.6.2 Data Anonymization Layer

As privacy is a major concern for both personal information exchange and community data sharing, our proposed framework incorporates an anonymization layer

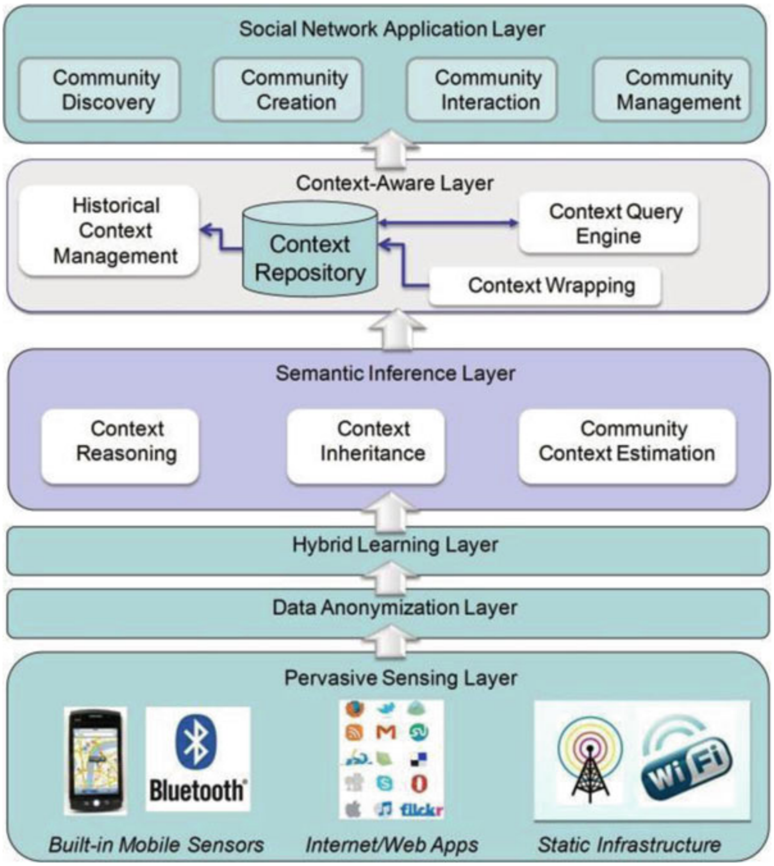


Fig. 6.2 Personal and community context management framework

before the releasing and processing of acquired user data. All the data released must be anonymized to protect individual privacy; different anonymization algorithms can be applied for privacy protection (Campbell et al. 2008).

6.6.3 Hybrid Learning Layer

The hybrid learning layer applies diverse machine learning and data-mining techniques to convert the low-level single-modality sensing data into high-level features or micro-context; the focus is on mining frequent data patterns in order to derive the individual's behavior and single space context, before extracting the high-level personal and community context.

6.6.4 Semantic Inference Layer

The semantic inference layer is needed when different features or micro-context need to be aggregated using logic-based inferences; it is complementary to the statistical learning approach. The following components are involved:

- Context reasoning. This uses rule-based reasoning engines to infer high-level contexts from low-level contexts, leveraging experts' domain knowledge.
- Context inheritance. This provides the necessary algorithms in order to estimate currently unavailable context based on information available from parent communities.
- Community context estimation. This is responsible for inferring community context based on personal context of each community member. Various models can be used to represent and estimate the community context values, e.g., aggregation of individual user context of each community member, average values, maximum or minimum value of the community members, etc. (De Silva et al. 2009).

6.6.5 Context-aware Layer

This is the middleware layer that manages the context obtained from lower layers and provide efficient retrieval support for MSN services. A set of components are included.

- The context repository is the core component that can represent and store context in a unified form (e.g., RDF, OWL) (Wang et al. 2004b, Guo et al. 2011).
- The historical context management module is responsible for the consistent storage of historical context data for both individuals and communities.
- The context query engine provides a unique interface for MSN services to access the context needed.
- Context wrappers are responsible for converting the learned/inferred context into proper format based on the formal modeling language, and inserting them into the context repository (Yu et al. 2008).

6.6.6 Social Networking Application Layer

The application layer includes a variety of potential MSN services that can be enabled by the availability of individual and community contexts, including community creation, interaction, management, and so on.

6.7 Techniques for Extracting Personal and Community Context

In order to extract personal and community context, various technologies could be deployed and developed. This section investigates what existing technologies can be used to empower personal and community context extraction, and more importantly, what new methods have to be developed to bridge the semantic gap.

6.7.1 Existing Technologies for the Extraction of Personal and Community Context

There is a spectrum of technologies which are able to partly meet the data representation, data processing, and inference requirements of personal and community context. We first give a summary of existing technologies as shown in Table 6.1, and then present them in more detail.

1. Context Representation and Anonymization Techniques
   Effective data representation techniques are the basis of extracting personal and community context. Ontology and Semantic Web are two widely adopted techniques for context representation. Ontology is a formal representation of a set of concepts within a domain, and the relationship among those concepts. It has been used in artificial intelligence, the Semantic Web, system engineering, etc. as a means of knowledge representation. The key benefits of ontology include: sharing common understanding of information among people or agents, reusing of domain knowledge, making domain assumptions explicit, separating domain knowledge from operational knowledge, and supporting logic-based inference. For instance, an ontology-based user profile model is proposed in Stan et al. (2008) to allow users to have a situation-aware mobile social network through handhold devices, by controlling how reachable they are for specific categories of people in a given situation.

   The Semantic Web was developed to make it possible for machines to understand the semantic meaning of information on the World Wide Web. The concept of Semantic Web applies methods beyond linear presentation of information (Web 1.0) and multi-linear presentation of information (Web 2.0) to make use of hyper-structures, leading to entities of hypertext. For example, Semantic Web is used to build agent-based system architecture for developing mobile social networking applications (Karistiansson et al. 2009).

   In addition to context representation, another important technology for personal and community context is data anonymization techniques. Data anonymization refers to technology that converts clear text data into a non-readable and irreversible form to protect privacy. One method for privacy protection is to use hash or encryption techniques to hide sensitive information. For example, in the MIT Reality Mining dataset, Hash processing (MD5) has been performed to

Table 6.1 Existing personal and community context related technologies

Category	Description	Data sources	Techniques
Individual context	Individual behavior pattern (temporal, spatial and social behavior, e.g., mobility pattern)	Smart phone (e.g., GPS, Wi-Fi, Bluetooth, sensors) Smart card (e.g., transport card) Internet service (e.g., Tweeter, 4sq)	Classification methods (e.g., supervised learning: maximum entropy, naive Bayes classifier, decision trees, SVM; semi-supervised)
	Inter-user interaction pattern (virtual interaction, physical encounter)	Smart phone (e.g., Wi-Fi, Bluetooth, call records) Internet service (e.g., Facebook, Tweeter, 4sq)	Classification methods (e.g., supervised, semi-supervised)
	User-place interaction pattern	Social networks (e.g., Flicker, 4sq) Static infrastructure (camera, Wi-Fi, Bluetooth, RFID, cell tower)	Classification methods (e.g., supervised, semi-supervised)
	Relationship between behavior patterns in the virtual world and the physical world	Online social networks (e.g., Flicker, 4sq) Event-based social networks (EBSN, e.g., Meetup)	Association rules
Community Context	Community collective behavior (community preference, community intent)	Social networks (e.g., Facebook, Flicker, 4sq, Meetup) Static infrastructure (camera, Wi-Fi, Bluetooth, RFID, cell tower) Smart phone Smart card	Inference techniques (e.g., rule-based, case-based, hybrid)
	Community lifecycle pattern (community formation, evolving and dissolving pattern)	Social Networks (e.g., Facebook, Flicker, 4sq, Meetup) Static infrastructure (camera, Wi-Fi, Bluetooth, RFID, cell tower) Smart phone Smart card	Classification methods (e.g., supervised, semi-supervised)
	Multi-granularity community structure (community detection)	Online social networks Offline social networks	Clustering methods (e.g., unsupervised: k-means clustering, hierarchical clustering)

hide user identity (Eagle et al. 2009). Another method for privacy protection is to ignore individual information and focus instead on a coarse-grained description of the system, such as the aggregated motion of individuals (Gonzalez and Barabasi 2007). Other privacy-preserving methods can also be employed, ranging from sharing only statistical summaries of the individual data set, to inserting random perturbations into individual data records before sharing them (Mitchell 2009).

2. Clustering/Classification Techniques
Data clustering/classification is an effective way to extract high-level personal and community context, such as human behavior patterns, community structures, and community dynamics. In general, clustering/classification techniques can be classified into three categories, i.e., supervised learning, unsupervised learning, and semi-supervised learning.

   Supervised learning is the machine-learning method of inferring a function from supervised training data. A supervised learning algorithm analyzes training data and produces an inferred function, which is called a classifier (if the output is discrete) or a regression function (if the output is continuous). The inferred function should predict the correct output value for any valid input object. There are two different kinds of classifiers, which are static classifier and temporal classifier. Static classifiers include SVM, naïve Bayes, Bayesian network, and decision tree, etc. The temporal classifier includes hidden Markov model (HMM), conditional random fields (CRF), and dynamic Bayesian networks (DBN), etc. For example, SVM can be used to identify human interaction patterns in a smart meeting scenario (Yu et al. 2012).

   Unsupervised learning addresses the problem of finding hidden structure in unlabeled data. Since the examples given to the learner are unlabeled, there is no error or reward signal to evaluate the potential solution. Approaches to unsupervised learning include clustering and blind signal separation. Clustering methods include k-means, mixture models, k-nearest neighbors, hierarchical clustering, etc. For instance, hierarchical clustering is used to discover hierarchical community structures of various networks in Ahn et al. (2010), e.g., uncovering mobile social networks based on mobile phone call records; a variant of k-means is leveraged to detect overlapping communities of online social media in Wang et al. (2010).

   Semi-supervised learning combines both labeled and unlabeled examples to generate an appropriate function or classifier. On one hand, labeled instances are often difficult, expensive, or time consuming to obtain, as they require the efforts of experienced human annotators. On the other hand, unlabeled data may be relatively easy to collect, but hard to use. Semi-supervised learning addresses this problem by using large amount of unlabeled data, together with the labeled data, to build better classifiers. Because semi-supervised learning requires less human effort and gives higher accuracy, it is of great interest both in theory and in practice. Self-training is a commonly used technique for semi-supervised learning, which is applied to collectively identify human behavior and context based on smart phone data in mobile social networks (Miluzzo et al. 2010).

3. Context Reasoning Techniques
   In general, there are three different kinds of context reasoning techniques, which are *rule-based reasoning*, *case-based reasoning*, and *hybrid reasoning*.

   Rule-based reasoning (RBR) complements the statistical learning approach, and uses explicit rules to effectively associate low-level information with the expected personal or community context on the basis of expert domain knowledge. For example, RBR is used to analyze human behavior models in Shi et al. (2009).

   Case-based reasoning (CBR) aims to solve new problems based on the solutions to similar old problems. Case-based reasoning is a kind of analogy making. The advantages of CBR include minimal effort required for knowledge acquisition and maintenance, high problem-solving capability based on knowledge reuse, and automatic evolution over time. For instance, in Sadeh et al. (2009) CBR is used to identify the user's privacy preference when using mobile social networking applications.

   Hybrid reasoning is an inference method that leverages multiple reasoning techniques, such as rule-based, case-based, et al. It combines the advantages of different methods, and is able to achieve better performance. For example, COSAR leverages hybrid reasoning to recognize human activities through the use of ontologies and ontological reasoning combined with statistical inference (Riboni and Bettini 2011).

6.7.2 Future Technologies Empowering Personal and Community Context

While there are already quite a number of techniques and tools that are capable of supporting personal and community context processing, they are still not able to fully empower personal and community context-awareness in MSNs. In this section, we will discuss what technologies and tools are expected to empower context aware MSNs.

Community models: central to the extraction of community context will be the development of community models — computational models of the social networks that bind individuals, mine the various links among individuals, recognize different community types comprising tightly connected people, and determine the collective characteristics of communities (Lane 2012). Meanwhile, due to the complexity of human behaviors and diversity of interaction patterns, different community models might be needed in different circumstances. Community models will influence how community context should be extracted, guiding a variety of system operations. To this end, to facilitate community context, it is necessary to investigate and develop proper community models according to the specific requirements of applications.

Participatory sensing technology: participatory sensing is emerging with the prevalence of mobile smart devices to collect large-scale data about people and society: however, accuracy, frequency, completeness, and coverage cannot be assured for the data acquired. A more realistic strategy is to leverage human judgment and assistance in supervising the use of an individual's electronic products (i.e., smart mobile phones equipped with camera, microphone, GPS, compass, accelerometer, etc.) for data collection, to record and share their surrounding information. The shared information is not raw data from sensors but labeled data with annotation from users. This strategy is called participatory sensing (Shilton 2009), which is a promising technique for personal and community context extraction.

Large-scale multi-modal data fusion and integration: huge amounts of data are being produced in mobile social networks every day, such as tweets published over micro-blog, photos and videos shared through Flickr or YouTube, sensor data collected by smart phones, leading to severe data fusion and integration challenges. In particular, context-aware MSNs need to deal with real-time, multimodal data collected on city or society scale; existing data-mining techniques fall short of capabilities to model, fuse, integrate, and manage the multi-model, heterogeneous data. Thereby, effective fusion and integration tools and methods are needed to handle large-scale multi-model data (Hu et al. 2011).

6.8 Research Topics and Challenges

6.8.1 Data Quality

The data quality from the user-contributed contents in MSN can be very different, ranging from authorized to inaccurate and even fake ones. The data quality from mobile phones and infrastructure also varies a lot. On one hand, due to limited resources, many of the sensors found in today's mobile devices produce data that has low resolution and limited precision. On the other hand, people place their phones in varied ways. For example, some people put their mobile phones in the pocket, some put in the handbag (Sun et al. 2010). Thus for the same user activity such as walking or running, the data quality is very different. When analyzing human behaviors from raw sensor data, it is better to train different classifiers that work in different contexts. However, both data collection and context identification are challenging issues.

6.8.2 Data Uncertainty

The participation of users introduces uncertainty to MSN systems. To mine community contexts, we often need to collect data from many anonymous participants. If there is a lack of the control needed to ensure the source is valid and information is accurate, this can lead to data trust issues. For example, Twitter data is sometimes unreliable; mobile phone users may send incorrect or even faked data to the data centre. Therefore, trust and abnormal data detection methods should be developed to ensure the trustworthiness and quality of the collected data.

6.8.3 Heterogeneous Data Management

Different mobile devices have different capabilities; they might have different levels of accuracy in sensing the physical and virtual world. Integrating information from diverse data sources adds difficulty to community context mining. Raw data from different sensor sources need to be transformed to the same metrics and represented by a shared vocabulary/ontology to facilitate the learning and inference process (Wang et al. 2004a).

6.8.4 Data Fusion

Data from independent sensing sources should (1) be associated, integrated, and fused to infer high-level contexts, and (2) be cross-checked to allow trustworthy information inference. For instance, if a body-worn accelerometer detects a "sitting" activity and the GPS system detects that the user is moving, the activity can be more properly corrected as "driving/taking bus".

6.8.5 Data Visualization

To better understand data produced by mobile social networks, visualization techniques and tools should be developed. These should be designed to demonstrate the sensed data and learned contexts at multiple levels of granularity. Situvis (Clear et al. 2010) represents an important step in this direction.

6.8.6 Security and Privacy

In addition to the privacy issues faced by traditional online social networks, sharing and revealing personal data in MSNs are exposed to extra privacy issues that are unique to mobile environments. Compared with personal data (e.g., user profile, IDs), data gathered in community can reveal much more information about individual and organizational behaviors. For example, your location might reveal your interests; the health data about an organization might suggest environmental problems for the staff. The impact is obvious: if personal data cannot be anonymized and under the control of data owners, people may be less likely to share their data. The problem becomes more important in short-term communities: the lack of centralized control and the anonymous-participation nature pose additional security challenges.

User control is very important in personal data sharing, as it is about what one wants to reveal and to whom one allows the system to reveal. For example, you might want to track your heart rate each day, but there is no reason to share that information with anyone but your doctor. Researchers in this field are exploiting methods that enable users to manage their data by tailoring access-control and data-management tools (Shilton 2009).

6.8.7 Community Management

In comparison with individual users, community information is more difficult to manage due to its complex nature, such as different forming methods (ad hoc, common goal-based), membership (different user roles), granularities (there are parent and children communities), lifecycle (short-term or long-term), and so on. We should study the social science and domain knowledge to provide effective tools for community management.

6.8.8 Context-Awareness in MSNs

Extraction of individual contexts has been well studied in the field of pervasive computing. Research on community contexts is still at an early stage. Challenges include how to extract community preferences, how to mine the underlying structure of MSNs, how to identify and store historical and live contexts in ad hoc or long-term communities, and so on.

6.8.9 Social Concerns

MSN services are built on positive social properties such as friendships, communities, contacting/encountering, etc. There is also a need to pay attention to negative social features, such as social selfishness (Guo et al. 2011), to ensure data sharing in delay-tolerant MSNs.

6.9 Conclusion

Mobile social networks are still in their infancy, but have the potential to revolutionize the field of social networking by incorporating social and community awareness in MSNs and fusing together the mobile, sensor, and social data. In this chapter, we first extend the definition of mobile social networks by classifying MSNs into four categories. With the new taxonomy of MSNs, mobile social networking taking place in both *physical world* and *virtual world*, for either *spontaneous* or *long-term* interaction, is included. By examining the new facets of context-awareness in the emerging field of mobile social networks, we present two important terms: *personal context* and *community context*, which play crucial roles in enabling future intelligent mobile social networking. We further present the model and detailed description of *personal context* and *community context*.

Corresponding to the life cycle of the mobile social community, the major functional phases of MSNs are divided into four: *discovery, connection, interaction, and management*. In each phase, how personal context and community context is used to facilitate the process is elaborated.

With the merging of pervasive computing and social computing, we attempt to identify the possible data sources which can be leveraged to derive social and community context; these data sources include *sensor-rich mobile and wearable devices, social networking and web services*, and *infrastructure-based networked sensors*. Leveraging the data sources, techniques ranging from data representation, data cleansing, and data anonymization to clustering techniques and inference techniques are presented for deriving personal and community context. Finally, various research challenges are summarized in order to shed light on possible future research directions about context-awareness in mobile social networking. In summary, with context awareness, MSN is not merely a simple extension of online social networking, it goes beyond its counterpart with many new perspectives and much more intelligence.

Acknowledgments This work is partially supported by the EU FP7 Project SOCIETIES (No. 257493) and EU FP7 MONICA Project (No. 295222).

References

Abowd, G., et al. (1997). Cyberguide: A mobile context-aware tour guide. *Wireless Networks, 3*, 421–433.

Ahn, Y. Y., Bagrow, J. P., & Lehmann, S. (2010). Link communities reveal multiscale complexity in networks. *Nature, 466*(7307), 761–764.

Beach, A. et al. (2010). Fusing mobile, sensor, and social data to fully enable context-aware computing. In *The eleventh workshop on mobile computing systems & applications* (pp 60–65). New York: NY, USA.

Campbell, A. T., et al. (2008). The rise of people-centric sensing. *IEEE Internet Computing, 12*(4), 12–21.

Clear, A. K., Holland, T., Dobson, S., Quigley, A. J., Shannon, R., & Nixon, P. (2010). Situvis: A sensor data analysis and abstraction tool for pervasive computing systems, *Pervasive and Mobile Computing, 6*(6), 575–589.

Dey, A. K., et al. (2001). A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. *Human–Computer Interaction, 16*, 97–166.

De Silva, H., Moessner, K., & Carrez, F. (2009). Group knowledge management for context-aware group applications and services. In *Proceedings of PIMRC* (pp 2851–2855). Japan: Tokyo.

Eagle, N., Pentland, A., & Lazer, D. (2009). Inferring social network structure using mobile phone data. *Proceedings of the National Academy of Social and Community Intelligences, 106*(36), 15274–15278.

Gonzalez, M. C., & Barabasi, A. L. (2007). From data to models. *Nature Physics, 3*, 224–225.
Guo, B., Zhang, D., & Imai, M. (2011). Toward a cooperative programming framework for context-aware applications. *Personal and Ubiquitous Computing (PUC), 15*(3), 221–233.
Hu, D. H., Zheng, V. W., & Yang, Q. (2011). Cross-domain activity recognition via transfer learning. *Pervasive and Mobile Computing, 7*(3), 344–358.
Klamma, R., Spaniol, M., & Cao, Y. (2006). Community aware content adaptation for mobile technology enhanced learning. In *Proceedings of the first european conference on technology enhanced learning* (pp. 227–241). Crete.
Kleanthous, S. (2007). Semantic-enhanced personalised support for knowledge sharing in virtual communities. *In Proceeding of the 11th international conference on user modeling, UM 2007, 4511* (pp. 465–469). Berlin: Springer.
Koch, M., Groh, G., & Hillebrand, C. (2002). Mobile Communities – Extending Online Communities into the Real World, Proc. Americas Conf. on Information Science (AMCIS2002), Dallas, TX, USA.
Karistiansson, J., Hallberg, J., & Synnes, K. (2009). An architecture for mobile social networking applications. In *First international conference on computational intelligence, communication systems and networks* (pp 241–246). Idore.
Lane, N. D. (2012). Community-aware smartphone sensing systems. *IEEE Internet Computing, 16*(3), 60–64.
Mani, M., Ngyuen, A.-M. & Crespi, N. (2009). What's up: P2P spontaneous social networking. IEEE International Conference on Pervasive Computing and Communications (pp 1–2). Galveston: TX, USA.
McArthur, R., & Bruza, P. (2001). The ABC's of online community. In *Web intelligence: research and development. Lecture notes in computer science, 2198*, 141–147. Springer-Verlag: London, UK.
Menkens, C. (2009). Towards a context aware mobile community application platform. In *Sixth international conference on information technology* (pp 1504–1509). Las Vegas: NV, USA.
Miluzzo, E., Cornelius, C.T., Ramaswamy, A., Choudhury, T., Liu, Z., & Campbell, A.T. (2010). Darwin phones: The evolution of sensing and inference on mobile phones. In *Proceedings of the 8th international conference on mobile systems, applications, and services* (pp. 5–20). New York.
Mitchell, T. M. (2009). Mining our reality. *Science, 326*(5960), 1644–1645.
Mohammed, S., & Dumville, B. C. (2001). Team mental models in a team knowledge framework: Expanding theory and measurement across disciplinary boundaries. *Journal of Organizational Behavior, 22*(2), 89–106.
Pentland, A. (2005). Socially aware, computation and communication. *IEEE Computer Society, 38*(3), 33–40.
Riboni, D., & Bettini, C. (2011). COSAR: Hybrid reasoning for context-aware activity recognition. *Personal and Ubiquitous Computing, 15*(3), 271–289.
Roussaki, I., et al. (2012). Context-awareness in wireless and mobile computing revisited to embrace social networking. *IEEE Communications Magazine, 50*(6), 74–81.
Sadeh, N., Hong, J., Cranor, L., Fette, I., Kelley, P., Prabaker, M., & Rao, J. (2009). Understanding and capturing people's privacy policies in a mobile social networking application. *Personal and Ubiquitous Computing, 13*(6), 401–412.
Schilit, B., Adams, N., & Want, R. (1994). Context-aware computing applications. *IEEE workshop on mobile computing systems and applications (WMCSA'94)* (pp. 89–101). Santa Cruz.
Shi, P., Liu, F., Yang, M., & Wang, Z. (2009). A fuzzy rules-based approach to analyzing human behavior models. In *The 11th international conference on computer modelling and simulation* (pp 346–351). Cambridge: UK.
Shilton, K. (2009). Four billion little brothers? Privacy, mobile phones, and ubiquitous data collection. *Communications of the ACM, 52*(11), 48–53.

Stan, J., Egyed-Zsigmond, E., Joly, A., & Maret, P. (2008). A user profile ontology for situation-aware social networking. In J. C. Augusto, D. Shapiro, & H. Aghajan (Eds.), *3rd workshop on artificial intelligence techniques for ambient intelligence (AIT AMI 2008)* (pp. 1–6). Patras.

Sun, L., Zhang, D., Li, B., Guo, B., Li, S. (2010). Activity Recognition on an Accelerometer Embedded Mobile Phone with Varying Positions and Orientations. UIC 2010, Xian, China, 548–562.

Wang, X., Zhang, D., Gu, T., & Pung, H.K. (2004a). Ontology-based context modeling and reasoning using OWL. In *PerCom CoMoRea workshop* (pp. 18–22). Florida.

Wang, X., et al. (2004b). Semantic space: An infrastructure for smart spaces. *IEEE Pervasive Computing, 3*(3), 32–39.

Wang, X., Tang, L., Gao, H., & Liu, H. (2010). Discovering overlapping groups in social media. In *ICDM* (pp 569–578). Washington: USA.

Wikipedia. http://en.wikipedia.org/wiki/Mobile_social_network. Accessed 15 Nov 2012.

Yu, Z., Yu, Z., Zhou, X., Becker, C., & Nakamura, Y. (2012). Tree-based mining for discovering patterns of human interaction in meetings. *IEEE TKDE, 24*(4), 759–768.

Yu, Z., Zhou, X., Yu, Z., Park, J. H., & Ma, J. (2008). iMuseum: A scalable context-aware intelligent museum system. *Computer Communications (ComCom), 31*, 18,4376–4382.

Zhang, D., Wang, X., Leman, K., & Huang, W. (2003). OSGi-based service infrastructure for context-aware connected homes. In *International conference on smart homes and health telematics* (pp 1–8). Paris.

Zhang, D., Yu, Z., & Chin, C. Y. (2005). Context-aware infrastructure for personalized healthcare. In C. D. Nugent, P. J. McCullagh, E. T. McAdams, & A. Lymberis (Eds.), *Personalised health management systems* (pp. 154–163). The Netherlands: Ios Press.

Zhang, D., Guo, B., Li, B., & Yu, Z. (2010). Extracting social and community intelligence from digital footprints: An emerging research area. In *UIC 2010* (pp. 4–18). Xian: China.

Zhang, D., Guo, B., & Yu, Z. (2011). The emergence of social and community intelligence. *IEEE Computer, 44*(7), 21–28.

Chapter 7
Enhancing Mobile Social Networks with Ambient Intelligence

Kevin Doolin, Nick Taylor, Micheal Crotty, Mark Roddy, Edel Jennings, Ioanna Roussaki, and David McKitterick

Abstract Mobile social computing has exploded into people's lives during the past 10 years, but to become truly pervasive it needs to be much more context-aware and personalizable. The next generation of social media needs to be able to react and adapt to the physical environments in which people live and act. The SOCIETIES project is integrating research undertaken in the field of pervasive computing with social computing to develop the next generation of social media systems. The vehicle for this is the "pervasive community", and this chapter outlines the innovations required to realize this concept. Pervasive communities can restore the symbiosis between our digital and physical worlds.

7.1 Introduction

When considering the future of mobile social computing, it is important to recall that people are mobile creatures, and from the time we developed our first indirect means of communicating (writing) the tools we use to socialize have

K. Doolin (✉) • M. Crotty • M. Roddy • E. Jennings
Telecommunications Software and Systems Group (TSSG), Waterford Institute of Technology, Waterford, Ireland
e-mail: kdoolin@tssg.org; mcrotty@tssg.org; mroddy@tssg.org; ejennings@tssg.org

N. Taylor
School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, UK
e-mail: N.K.Taylor@hw.ac.uk

I. Roussaki
National Technical University of Athens, 157-173 Zographou, Athens, Greece
e-mail: ioanna.roussaki@cn.ntua.gr

D. McKitterick
INTEL, Collinstown Industrial Park, Leixlip, County Kildare, Ireland
e-mail: davidx.mckitterick@intel.com

A. Chin and D. Zhang (eds.), *Mobile Social Networking: An Innovative Approach*, Computational Social Sciences, DOI 10.1007/978-1-4614-8579-7_7,

accommodated our roaming nature. Historically, those interludes during which humanity has surrendered to attending fixed locations in order to socialize indirectly have been brief. The fixed telephone was supplemented by the mobile phone within 50 years of it entering most homes. The early years of social networking, which tied people to their desktop computers, were just another such interlude. The users of those sites already had SMS available to them when they were on the move, and mobile social networking was inevitable with the advent of the smartphone.

Pervasive computing has the potential to totally transform traditional Web-based social networking sites (SNSs), freeing them from the Web-browser and letting them loose in the real world (Pallis et al. 2011; Xu et al. 2011). The most popular of these, such as Facebook, LinkedIn, MySpace, and Twitter, have already embraced smartphone technology, offering mobile versions for their users to access while on the move (Semertzidis and Laso-Ballesteros 2010). Future social networking users will demand the context-aware "everyware" services supporting personalization, which pervasive computing can provide anywhere and at any time.

The SOCIETIES project (SOCIETIES 2012) aims to bridge the gap (Doolin et al. 2012) between the current Web and smart phone application-based social networking services (Zaphiris and Ang 2009) and the next generation of pervasive computing services (Obaidat et al. 2011; Hansmann et al. 2012). In its simplest manifestation, the integration of pervasive and social computing will permit the online "poke" or "nudge" to morph into an image of the poker appearing on a display in the poked person's vicinity, wherever they are. This might still be a computer screen but it could also be a public display, TV, digital photoframe, a haptic or audio notification on a mobile device, etc. More advanced convergence will unify the two concepts completely into "pervasive communities". Mobile devices form a core element of this pervasive community concept. It is envisaged that the mobile devices of a user will be fully integrated into a community of common mobile resources, applications, and services – in essence, the mobile device will provide the user's hook into the pervasive community, will provide their identity, and will provide aspects of their context. As such, the community will always follow the user (based on their current context), and will integrate with fixed and mobile smart spaces as required. This fully ties into the concept of local–social–mobile.

The people in one's social network could be locatable and, if in sight, identifiable, within the limits of the personal privacy policies of the individuals concerned, of course. People with shared interests are likely to belong to a common pervasive community or group within an SNS. They are likely to attend the same events or places of interest and, if they were at such a place concurrently, pervasive communities could alert them to each other's presence and enable them to meet physically. Pervasive communities will enable people to leave messages about exhibits, restaurants, etc. which can be picked up automatically by other members who visit the same place at a later date. Streaming video facilitated by pervasive communities will enable their members to see and hear what their friends are seeing and hearing live with no effort and no delay.

The context-awareness of pervasive communities is not simply limited to location-awareness, however (Lukowicz et al. 2012). The "everyware" (Greenfield 2006) information available to a pervasive community can report a host of measurable characteristics about a member's environment and pass them to other members, even when on the move. As previously stated, it is the mobile device that is the hook that associates the user to their pervasive community. Weather conditions, for instance, helping one to decide where to go and what to wear on a free afternoon, as well as traffic reports, queue lengths at tourist attractions, delay times for planes, trains, and buses, and even dynamic information of other community members, such as their current activity, status, and location, are all types of information that can be made available to a pervasive community member anywhere anytime, either in their raw form, or filtered through the perceptions of what like-minded fellow members believe to be important and/or relevant.

Nor are the actions which a pervasive community member can take limited to data manipulation. Pervasive services will enable a member to affect things in their environment and, if one was minded to, one could allow one's friends to change one's environment using those same services. A friend could change one's TV channel to a particular show that they have suggested you watch, for instance or set one's oven to a temperature just right for baking the perfect cake. Robotic companions could be controlled by friends who are geographically separated. Computer-supported co-operative work (CSCW) could be completely transformed by the unification of pervasive and social computing in pervasive communities.

The rest of this chapter is structured as follows. Section 7.2 elaborates on the vision of the SOCIETIES project introducing the *discover–connect–organize* approach. Section 7.3 defines the necessary concepts and presents the SOCIETIES system architecture. Section 7.4 describes the main innovations of the SOCIETIES project, classifying these into seven categories. Section 7.5 elaborates on the SOCIETIES methodology that is based on the inclusion of three different user communities: an Enterprise, a Student and a Disaster Management Community. Section 7.6 provides some details on the project status, as well as on the future research and exploitation plans. Finally, in Sect. 7.7, conclusions are drawn.

7.2 SOCIETIES Vision

To date, pervasive computing systems have been designed mainly to address the needs of individual users. The sensor nets and services, which form part of the smart spaces of these pervasive systems, while supporting multi-user operations, are targeted at people who are assumed to be only interacting with the smart space and not directly with other users. At the same time, the recent development and rise in popularity of social computing has occurred more or less in isolation from developments in pervasive computing (Parthasarathy et al. 2011).

Experience has taught us that some of the most beneficial and rapidly assimilated uses of computer technology have arisen when that technology has brought people together and allowed them to communicate and collaborate. Computer-supported co-operative work (CSCW), massively multiplayer online role-playing games (MMORPGs), social networking sites (SNSs), even the humble SMS or instant messenger service have seen meteoric rises in popularity because they address a deep-rooted human need to socialize. The integration of pervasive technologies, from the real physical world, with social computing has the potential to generate a new wave of innovations in this novel hybrid area, and to re-factor ICT systems at the point of use so that they become more user-friendly, responsive, relevant, and capable of providing new means to harness data and engage more people, with a view towards tackling societal problems.

The SOCIETIES project aims to address the gap between pervasive and social computing by designing, implementing, and evaluating an open scalable service architecture and platform for pervasive communities. A pervasive community is inherently context-aware, and so can adapt to factors such as the user's location, activity, environment, physiological/psychological status, and current goals. In order to achieve this, a pervasive community has to be self-organizing, self-improving and capable of proactive behavior, which will optimize and personalize the pervasive experience of an individual user or an entire community.

A growing challenge that is faced by both social and pervasive computing is the enormous volume and variety of resources available to their users. As more people participate in social media, and as those media provide richer ways in which people can interact and share content, it has become essential to provide mechanisms which enable users to filter and organize the content they receive, in order to home in on what is relevant to them. Similarly, as pervasive technologies make more, and richer, sensors, networks, and services available to people at any time and in any place, identifying what is relevant, and what is not, becomes a fundamental requirement for usability.

The SOCIETIES platform recognizes, the paramount importance of relevance and the ways in which social and pervasive information and methods can be brought together to provide mutual support in identifying what is relevant to a user at any given time. Examples of factors that can contribute to determining what is relevant at a time, *t*, are:

- *Location* – at current time (t), historically at (t−1, t−2, ..), predicted at (t+1, t+2, ..)
- *Activity* and *status* – recognized by physical action monitoring and digital service usage, or predicted from prior tasks at (t−1, t−2, …)
- *Interactions* – people with whom one is communicating at present (t) or in the recent past (t−1, t−2, …), people with whom one is co-located, people with whom one is sharing services
- *Profile* – interests, background, affiliations, etc.
- *Physiology* – body temperature, perspiration rate, blood sugar level, etc.
- *Psychological/emotional state* – (un)happy, angry, bored, interested, etc.
- *Preferences* – service, interface, group preferences, etc.

Evidence from such a wide variety of sources (some pervasive, some social) can be bound by confidence limits, and combining the degrees of confidence from all sources can permit the calculation of an overall "quality of relevance" (QoR). This QoR can, in turn, be used by pervasive and social computing systems to decide what to filter out and what to alert a user to at a specific time and/or place. Such a decision process also needs to be personalizable, of course, so that each user can focus it to his or her own needs and tastes in a context-dependent fashion.

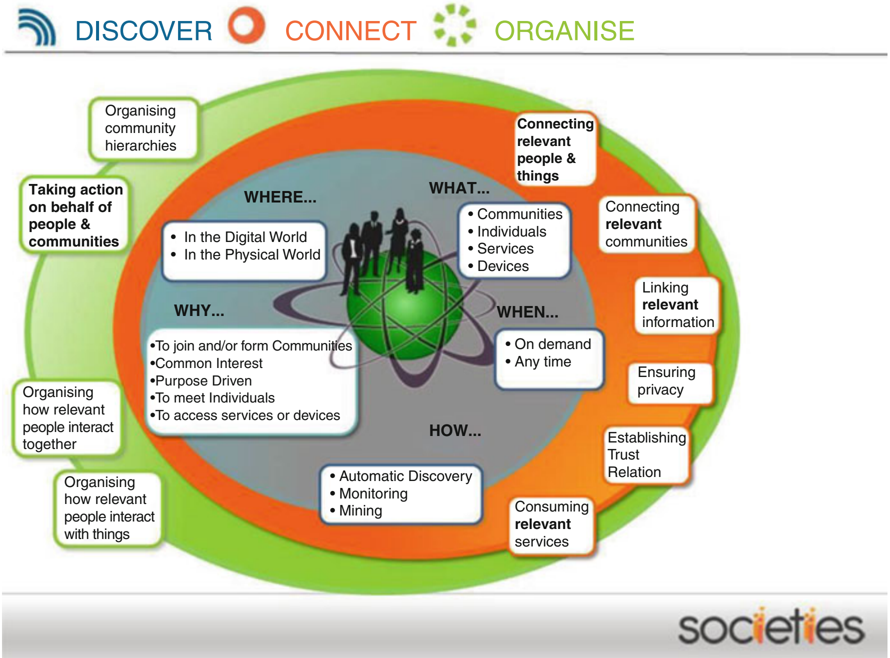


Fig. 7.1 The discover, connect, organize approach of SOCIETIES

The vision of SOCIETIES arises from this requirement to distinguish the relevant from the irrelevant, and can be summarized in three key concepts, each of which contributes to the formation of our pervasive communities. These concepts are: *discover*, *connect* and, *organize*, and the respective approach is illustrated in Fig. 7.1. In short, *SOCIETIES enables the discovery, connecting, and organizing of relevant people, resources, and things, crossing the boundary between the real and virtual worlds*. They are elaborated in turn in the subsequent sections.

7.2.1 Discover

SOCIETIES enables the discovery of trusted entities that reside in both the physical and digital worlds, such as individual people, communities, devices, resources, and services. What is important to note here is that the discovered entities have a distinct relevance either to each other or to the user on whose behalf the discovery process is operating – this concept of relevance is key throughout all innovations realized by the project, and sets SOCIETIES apart from current systems which typically adopt an all or nothing approach. Determining relevance is elaborated on further in the innovation sections that follow.

Existing systems such as LinkedIn or Facebook facilitate the discovery of people based on lightweight personal preferences and associations (primarily through manipulation of friends lists), while others harness lightweight contextual information (for example the use of a manual check-in on a mobile device); however, these systems do not enable the full exploitation of the capabilities of pervasive technologies or sensor networks to discover entities across the virtual – physical worlds of a user. SOCIETIES enables, for example, the discovery of a community of people with a common interest without a dependence on social network information (although it has to be noted that SOCIETIES integrates information from social networks in order to provide a deeper association between entities). Similarly, SOCIETIES can enrich existing social network services through the provision of rich contextual data about their users. The discovery notion in SOCIETIES therefore goes beyond the capabilities of existing technologies. Goal-, performance-, intention-, and commonality-driven discovery can greatly increase the performance of systems where access to relevant but unknown entities is required.

In terms of convenience, SOCIETIES' capability to discover based on relevance will increase the level of convenience provided to the entity seeking connections.

The discovery system is highly personalizable and context-aware. It can provide goal/social connection-, or situation-based discoveries, as well as learning and taking proactive discovery actions on behalf of users. Discovery also takes users' privacy requirements into consideration as SOCIETIES adopts a "privacy-by-design" approach throughout. Through the use of open federated identity management services, SOCIETIES allows the current social network monopoly to be broken, so that users can decide who provides their identity service. This will improve the overall quality of service and promote diversity of business models and opportunities.

Note that the provision of a privacy-aware external interface to allow third-party service developers to access this rich set of relevant entities opens up new possibilities in terms of next-generation mobile and social computing applications.

7.2.2 Connect

Building on the discovered relevant entities, SOCIETIES provides mechanisms for the interconnection of these entities across the physical and digital worlds. This enables communication connections to be established which capitalize on the capabilities of the entities they connect to. From a user's perspective, connections can take many forms – person to person, person to group, person to object, person to service – and only by accommodating all of these can any individual user's digital and physical worlds be seamlessly bridged. With the provision of external interfaces, third-party service providers are able to use SOCIETIES to enter the community-based mobile services market. Additionally, management of community resources, and off-device aggregated analytics, will enhance system and device performance (Lukowicz et al. 2012).

Convenience for the end user is greatly enhanced, since SOCIETIES enables the pervasive notion of having a user's/community's resources interconnected seamlessly, which alleviates the need for complex manual setup of entities.

Of course, the interconnections can be personalized and made context-aware on an individual basis, as will be discussed in the innovations section later.

7.2.3 Organize

The final stage of the DCO concept is of course "organize" (note that although the DCO concepts are presented in a linear manner in this document, in reality they form part of a continuous cycle).

Organizing, from a SOCIETIES perspective, relates to the complete lifecycle of discovered and connected entities. This lifecycle management includes the introduction of new entities to a connected community (based on further discover and connect cycles), and removal of entities that are no longer relevant or which may no longer require to be part of a specific community.

Enhanced intelligence functionality is available to community entities, and can be enriched by third-party services that can leverage SOCIETIES' enabling technologies such as privacy/trust management, security, learning, personalization, and context-awareness. It is critical to note here that once a community is formed, SOCIETIES then enables the deployment of new community services that harness contextual information from the physical and digital world, which can be fully personalized on an individual and community basis, and which protect privacy at all times. Some simple examples include the organization of a dynamic community for a business meeting or conference, or organizing a group of experts when needed in a disaster situation.

SOCIETIES' support for community hierarchies, members, and lifecycles facilitates the formation of temporary and ongoing communities for example. Analysis of an ad hoc temporary community's activities can facilitate the creation of a more permanent/ongoing community (for example a temporary community formed in a museum could become an ongoing community interested in a particular era of history). Additionally, identification of cohesive sub-communities can lead to the formation of useful community hierarchies (for example, a community of students could be formed into sub-communities based on course content, and therefore receive more focused, relevant services based on learning materials).

7.3 SOCIETIES Framework and System Architecture

7.3.1 Core Concepts

The project is formed around three interdependent concepts, illustrated in Fig. 7.2.

Pervasive community – a group of two or more individuals who have agreed to share some, but not necessarily all, of their pervasive resources.

CIS – a pervasive community, once constituted, forms a community interaction space (CIS). Individuals may belong to any number of pervasive communities and thus CISs simultaneously.

CSS – members of a pervasive community interact with a CIS via their own personal cooperating smart space (CSS), which lives on their mobile device(s). They can also interact with other CSSs directly, or without using CSSs at all. People can interact in person.

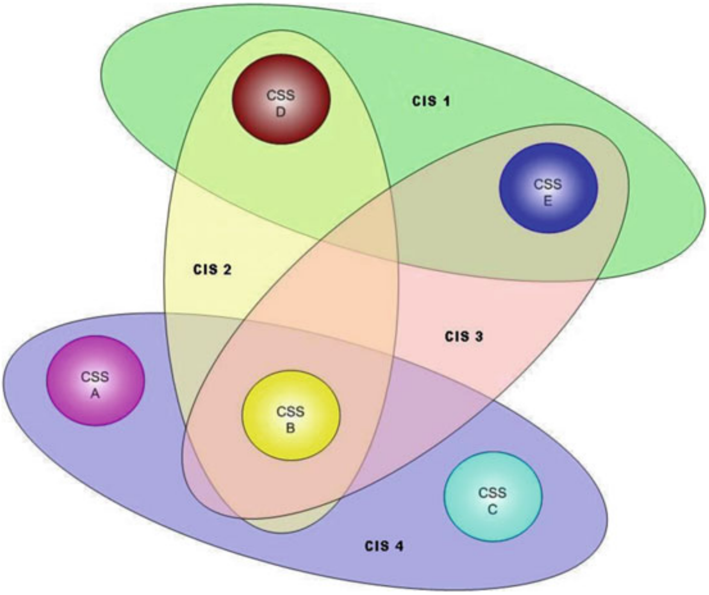


Fig. 7.2 Five individuals using their CSSs to form four pervasive communities

7.3.2 Architecture

This section provides an overview of the functional architecture for the SOCIETIES project. The diagram in Fig. 7.3 provides an overview of the platform "core services" provided. These are described at a conceptually very high level to give an understandable overview of the capabilities of the system.

These are further described under a grouping assigned from the logical deployment they support, e.g. services for an administrative domain, for CISs, for CSSs or for every node.

Multi CSS/CIS grouping – these services support a wider open group of stakeholders in a federated domain model. They offer federated search, identity, and domain administration functions, and store multiple CSS or CISs public information. There is one instance of these services per administrative domain, and other federated domains can request information.

This group includes the following services: the *domain authority* (which provides and manages the CSS and CIS identities in a decentralized manner, allowing authentication between multiple domains), the *CIS directory* (which manages the CIS information in a decentralized repository; it records available CISs within a domain or set of domains, it enables searching for CISs based on specific criteria, and it allows a CIS to be removed from the repository), the *CSS directory* (which provides search facilities for CSSs, based on their identifier or by specifying search criteria, such as public profile attributes and tags), the *CIS recommendation* (which is responsible for handling CIS recommendations, allowing for

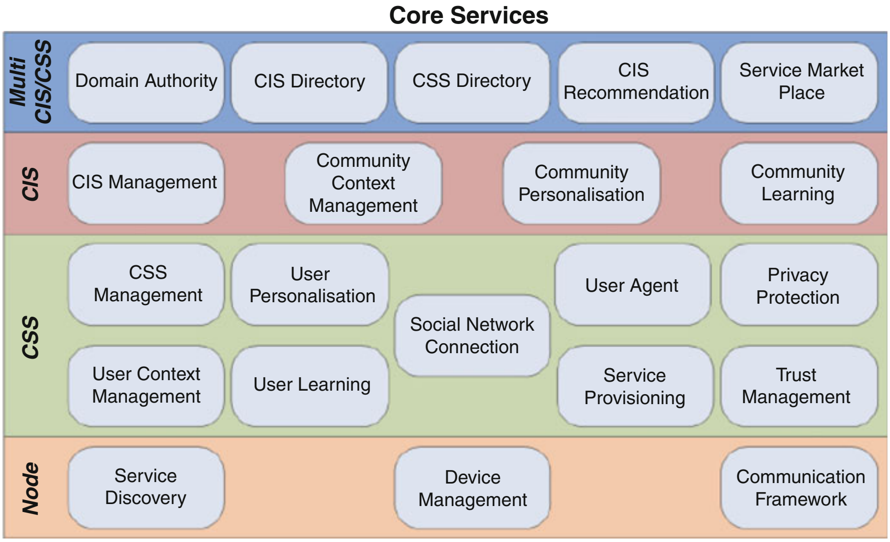


Fig. 7.3 High level architecture of the SOCIETIES platform

recommendations of CISs to users), and the *service market place* (which provides access to a repository of installable third-party (3P) services and optional "core" services, and provides mechanisms for charging for these services).

CIS grouping – these services support a community and their community interaction space (CIS). There is at least one instance of these services per CIS, and an instance of these services can be used by multiple CISs.

The CIS services are: the *CIS management* [which is responsible for handling all aspects of CIS lifecycle management (creation, update, and removal), provides control over CIS membership and includes a community profile manager and a role manager to specify the governance model for the CIS], the *community context management* (which enables access to and maintenance of community context, providing query capabilities, as well as addition/update/removal operations for community context, maintaining the history of context for a CIS, and inferring community context information), the *community learning* (which supports community preferences and community intent learning), and the *community personalization* (which manages the community preferences and community intent, and exposes interfaces for community members to retrieve these preferences and intent models for their own use).

CSS grouping – these services support a participant and their cooperating smart space (CSS). The word 'participant' is used to *refer to a single user or organization*. There is at least one instance of these services per participant, and an instance of these services can be used by multiple participants.

The CSS services are:

- The *CSS management* (which controls which nodes – devices or cloud instances – are part of the CSS, assigns them a common identity and manages resource sharing and configuration policies)

- The *user context management* (which is responsible for acquiring the user context from sensors and other context sources, for modeling and managing the collected data, for maintaining current and historic context in appropriate data repositories, and for the provision of inference techniques enabling the extraction of high level information from raw context data)
- The *user personalization* (which evaluates the user behavioral models, to identify and apply preference changes)
- The *social network connection* [which integrates with existing social networking systems (SNSs), enabling the extraction of public info available in SNSs, as well as access/update of non-public information for the specified user]
- The *privacy protection* (which provides identity management mechanisms, facilities for managing the CSS privacy policies, which specify the terms and conditions the CSS will respect concerning the personal data, also offering Privacy Policy Negotiation facilities)
- The *user learning* (which supports learning of user behavior models exploiting the user's history of actions stored in the system) and the *user agent* (that acts on behalf of a single CSS based on captured behavior models to establish the system's proactive behavior, and capturing any feedback)
- The *trust management* (which is responsible for collecting, maintaining, and managing all information required for assessing the trust relationships and evaluating direct, indirect, and user perceived trust) and *service provisioning* (which supports the setup and lifecycle control of a 3P service or CSS resource, allowing for installation)

Node grouping – the core services in this group are available per node. A CSS node is a *logical node/device/cloud instance running CSS software that coordinates with other CSS nodes to form a participant's CSS*. There is an instance of these services per CSS node.

This grouping includes the following services:

- The *communication framework* [which provides the necessary mechanisms to support intra- and inter-CSS communication through the discovery of CSS nodes (devices)]
- The *device management* (which provides mechanisms for managing attached devices and management of their capabilities)
- The *service discovery* (which provides service discovery and advertisement mechanisms, enabling the discovery of core platform services within a CSS, as well as, the discovery of 3P services shared by other CSSs or CISs)

7.4 SOCIETIES Innovations

In order to realize the SOCIETIES vision and support the three key concepts of discover, connect and organize, for pervasive communities, a number of challenges arise. The primary innovations required to address the challenges will now be detailed.

7.4.1 SNS Interoperability

An important aspect of the pervasive community vision is to capture and facilitate social Web interactions. However we are not producing just another social network. Instead, we need to develop a bridging scheme between the Social Web and pervasive communities.

Interoperability with Social Web initiatives needs to be taken into account, with regard both to standards work on federation and also to non-standardized approaches (e.g., Facebook, Twitter, Google+) given their popularity.

The SOCIETIES Social Connector/Extractor leverages and impacts on Web-based standard APIs and protocols specifically defined for efficient exchange of data among current and upcoming social networking initiatives.

1. *Bridging Social and Pervasive Communities*
   Our aim is to make pervasive communities interact seamlessly with the Social Web (Gallacher et al. 2011). This provides a key bridging or staging point for users not currently using pervasive communities. They can still interact with members of a pervasive community using familiar social networks, while not accessing all the benefits.

   Our first step is to link the profiles and activities of users in our pervasive communities with those in their online social networks (e.g., Facebook, Twitter, Diaspora, etc.). In order to do this, we need a common representation of people, profiles, and activities.

   A key second step is to facilitate exchanges of information between both systems. For example, details such as the name of the pervasive community, the participants' identities, and the activities of individual participants.

Data is exchanged by the following mechanisms:

Pull: SNS information is extracted to inform a participant of what is happening in the Social Web.

Push: pervasive community activity information is updated as user status in the Social Web.

Alignment: automatic and real-time pervasive community profile extraction and updating of corresponding Social Web profiles.

7.4.2 Context-Awareness

In today's systems, no context models have been developed that can support the management of context for a dynamic community in large- scale systems. While research has been carried out regarding community context, there is no support for resolving context conflicts or for inheritance of context information across hierarchical communities. Additionally, while research has been carried out with regard to the formulation of groups of people with common interests, the lifecycle management of these groups does not consider the many ongoing context changes that occur for each user. SOCIETIES is innovating in a number of areas (Doolin et al. 2012) related to this problem space, as described below. These innovations can be related back to the discover, connect, organize paradigm as follows:

Discover: context sensing resources, user context values and available resources based on current context information

Connect: users based on context similarity and connect users with relevant resources based on current and historic context information

Organize: community lifecycles and membership based on context information of individual members and of the entire community

1. *Context modeling*

   The SOCIETIES community context model (CCM) assists in the management of dynamically formed communities of people, crossing the digital and physical worlds, in which users discover relevant entities. Context information is one of the main criteria that is used to identify which users are relevant to which individuals, resources, and communities.

   SOCIETIES CCM allows modeling of context-related information for: communities of individuals, quality of community context, history of community context, social relationships (and social media information in general), bonds and relationships between users, and interactions among users.

   New context modeling approaches (Roussaki et al. 2012) have been developed in order to properly represent community context information. In this respect, hybrid models are used that support a multitude of community context representations, based on the type of the respective information, its quality and source. Thus, both deterministic and stochastic solutions are adopted, which vary depending on the qualitative and quantitative inherent features each context type may demonstrate. Furthermore, the conflicting context information that may be observed within communities is addressed with a flexible model that is capable of supporting context inheritance and refinement mechanisms.

2. *Community context extraction*

   Estimation of community context must be carried out both on demand and continuously. SOCIETIES context estimation can provide useful real-time information for community activities, and can connect people with common interests or other context commonalities/bonds. This forms a key part of the “relevance” aspect of the SOCIETIES discover, connect, organize paradigm. There are a number of models that are currently used to represent and estimate these community context values, such as: aggregation of the values of user context data for community members; stochastic representation of this mainly for discrete or enumerated context value formats, average/median values for the discrete or enumerated context value formats, most probable value, etc. The community context extraction is a difficult task that becomes more challenging as the complexity and dynamicity of context information rises and as the number of the community members increases.

3. *Sources of context updates*

   In order to update context information, SOCIETIES specifies a number of context information sources that can capture context values. Individual context data updates can be triggered by sensors, user actions, changes in social media data, quality of context thresholds, context refinement (inference) processes, community activity, community membership changes, etc.

As an example, a change in an individual user's context information may trigger the update of one or more pieces of community context information across multiple communities of which that person is a member. This process of keeping a community's context constantly updated enhances its usefulness, relevance, and trustworthiness, thus allowing more people to connect to it and hence to each other.

4. *Context similarity*
   Context similarity evaluation (CSE) in SOCIETIES is a challenging research area that requires new algorithms/mechanisms for context comparison. It is necessary to enable the evaluation of context similarity for both quantifiable context information such as location coordinates, weight, or temperature (which require arithmetic methods), as well as for qualitative context information such as user interests, status, or symbolic location (which require new principles in comparing context semantics to estimate similarity).

   The CSE concept supports a number of decision-making components within the SOCIETIES framework for tasks, such as: creation/discover/deletion of communities, identification of community hierarchies, identification of communities that could be merged/split, community membership management (i.e., addition/removal of members), prediction of user intent based on similar context (as common context often indicates similar intentions), and user preference discovery based on similar context (as the same preferences may be applicable not only under a given combination of context information values, but also when similar context conditions are observed).
5. *Location inference and prediction*
   Location is a core part of context information, and the techniques used to infer and predict location in SOCIETIES deserve some attention. Current approaches usually require manual check-in (e.g., Facebook, FourSquare) or semi-automatic check-in (e.g., Google Maps/Latitude). SOCIETIES focuses on automatic estimation of indoor location (through the use of Wi-Fi sniffers for example) with fine granularity (as opposed to the common practice of the manual check-in seen in many of today's mobile services). In distinction to other competitive solutions that support automatic location tracking, our approach does not require any expensive hardware. Furthermore, SOCIETIES provides an outdoor location prediction framework that exploits cellular traces and communication records, and forecasts a user's locations in a given time frame based on periodicity.

7.4.3 Learning

Learning in a social media context tends to be mainly for the benefit of the service provider, for example data mining for revenue generation through targeted advertising, whereas in a pervasive system, learning focuses more on the user by implicitly acquiring information which would otherwise have to be acquired explicitly – and typically only individual users are considered by such systems. SOCIETIES surpasses this current impasse by learning on a community level to underpin many of the project innovations at various points in the platform.

In previous research projects [such as PERSIST (2012) or DAIDALOS (2012)] preference learning mechanisms focused on the individual user. In SOCIETIES, we break new ground by learning about preferences of communities of users. This learning is based on historic behavior and context information collected from all members of a community (bearing in mind the privacy requirements of users) and fused together to create a single behavior and context history for the entire community. Context-dependent community preferences are extracted with these learning techniques, and are associated with a related community. Current and new community members can inherit all or part of the learned/modified community preferences to enhance their own preference set.

In terms of our DCO concept, learning relates as follows:

Discover: individual and community preferences
Connect: learning supports connection based on individual/community preferences
Organize: community-level learning assists individuals in acquiring information and links from other community members.

7.4.4 Privacy

Within SOCIETIES, all design has privacy built in as a fundamental requirement. Our "privacy-by-design" approach means that privacy protection is intimately integrated in the platform, rather than being an afterthought. SOCIETIES privacy protection supports the user in management of personal data (Gonçalves et al. 2012), enforcing negotiations between user preferences and service/community policies, data disclosure with personalized obfuscation, multi-identity management and selection, and privacy assessment.

In relation to DCO, the following can be stated:

Discover: individuals and communities who will comply with a user's privacy preferences.
Connect: privacy policy negotiation during connection with potential for data obfuscation, micro-agreement, and a privacy-aware social firewall
Organize: privacy audit/assessment contributes to reputation, and this enables more informed organizational activities.

1. *Personalised privacy policy negotiation*
   Privacy policy negotiation provides users with the ability to choose the personal data they wish to disclose to other entities. SOCIETIES uses user preferences to automate the privacy policy negotiation process, which in turn benefits the user by lifting the burden of privacy policy configuration from them. This also ensures that the data owner and requestor are bound by a common data disclosure

agreement. Policies also help the user to stay consistent in terms of the data they disclose, as the same preference can map to numerous situations. Privacy policies can be used to allow the user to assess suggested communities (automatically or manually), and privacy policy negotiation within SOCIETIES communities allows tailored data disclosure based on user needs.

2. *Data obfuscation*
   The purpose of data obfuscation is to reduce the personal content of data. The concept is an amalgamation of two "privacy-by-design" principles – data minimization and enforcement. While location and identity have been obfuscated in other research projects, SOCIETIES formalizes this obfuscation, generalizes it and applies it to many data types. When sharing data on behalf of a user, obfuscation takes place before any data are shared with others.
3. *Personalization and learning for privacy*
   SOCIETIES personalization and learning techniques are applied to data disclosure to support the learning of privacy preferences, and to facilitate automation of negotiation processes, identity selection, and obfuscating personal data.
4. *Micro-agreements for business*
   In support of enterprise, SOCIETIES support for micro-agreements allows employees to expose their company privacy policies for the social and pervasive mobile services they wish to consume or offer using micro-policies. Automatic negotiation of micro-agreements makes the connection between new community service providers and the service consumers in a secure and trustworthy manner. Companies can use the content of micro-agreements to monitor quality of service provided by service providers, and can organize their collaboration with other parties accordingly.
5. *Privacy assessment: privacy-aware social firewall*
   This novel area of research within SOCIETIES protects users from unintentional privacy leakages into social channels by monitoring data flows, estimating privacy risks, and passing control of them over to the user. The firewall operates based on privacy assessment which includes the following elements: auditing actual privacy practices by monitoring and logging how private information is actually used by applications and services, detecting privacy breaches and/or detection of potential privacy information leakage, and enforcing privacy by requesting immediate user feedback in case of potential violations. This usage of privacy assessment generates trust-relationship information, and is consequently an enabler of trustworthy entity discovery and connection.

7.4.5 Trust

Interaction between users, communities, and services requires formal trust assessment including robust and authenticated mechanisms. Aspects of trust (e.g., purposeful or referral) and trust relationships are the subject of current research, but these are not used in social media. Trust-related research supports the SOCIETIES DCO paradigm as follows:

Discover: trustworthiness of individuals, communities, services, and entities that you trust in advance.
Connect: to various entities based on individual and community trust assessment.
Organize: trust-based community membership management and trust-based community lifecycle (merging and splitting) based on trust relationships among members of existing communities.

Community-enhanced trust assessment based on feedback and learning
Assessment of the trustworthiness of individuals, communities, and services in SOCIETIES is based not only on the experiences of the user, but also on the experiences of fellow members of a community. Inherent in this process are SOCIETIES' feedback and learning mechanisms. This assessment provides the necessary facilities for trust-based discovery of individuals, communities and services. It supports automatic interconnection of, and information sharing of privacy-sensitive information with, trusted entities. SOCIETIES also supports community lifecycle and hierarchy management, based on the trust relationships among members of the parent community, and trust-based community lifecycle transition from temporary to ongoing, and facilitates the trust assessment process with community feedback. The SOCIETIES trust management and evaluation framework provides the necessary support infrastructure for the maintenance and management of dynamically changing trustworthiness within collaborative domains.

SOCIETIES researches direct trust (which is evaluated based on the interaction history of users and trust ageing) and indirect trust (in which case trust is inferred based on a user's fellow community members). A fusion of these trust types allows assessment of the aggregate trust value as perceived by the user. More specifically, some people employ objective measures to evaluate their level of trust in an entity, while others rely on a more subjective feeling. Direct trust generally outweighs indirect trust in this fusion process. However, the weight of each factor also depends on the confidence level with which it has been estimated.

7.4.6 Community Orchestration

Existing social networks provide a means to organize social network connections into communities (groups, circles, or cliques). However, the management of communities in current social media is largely manual, with some assistance provided in the form of suggestions.

In addition to the "connection" management overhead of the social interactions, Pervasive communities present a completely new challenge due to the physical element (devices, sensors, locations, and services) captured in the pervasive community concept. This adds "real world" impact to a group or community.

In order to maximize the value for users, we need to allow users to organize all aspects of communities. *Orchestration* refers to the ability to manage the intelligent formation, organization, membership, and termination of communities. Relating this back to the value propositions of the SOCIETIES project:

Discover: potential communities and members.
Connect: individuals via community membership/formation.
Organize: individuals into communities, form, merge, and delete communities and sub-communities.

1. *Context state model*
   Individuals and communities generate large amounts of data continuously, and this data needs to be gathered, with due regard to individuals' privacy policies, and processed so that it can be analyzed in a time-ordered manner. Potential new communities, defunct communities, and existing communities which an individual might wish to join/leave are identified via context state models (CSM).

   Each context state model (CSM) is made up of key attributes that describe characteristics relevant to existing and potential communities. Once a CSM has been created or modified, locality-sensitive hashing (LSH) is used to reduce the CSM to a simple rapidly computable value. The CSM data of individuals and communities is compared to that of others and trajectory mining techniques are used to discover similarities.

   CSMs can be used for analysis in near-real-time group dynamics. CSM modeling allows quick computational comparisons, which is vital for the discovery process, to ensure that a user has a 'live' (near-real-time) experience. CSMs also permit user data to be analyzed anonymously in an abstract fashion. A *community life-cycle manager* enforces the individual privacy constraints of users prior to any automatic or semi-automatic orchestration activity. The process is event-driven, and reacts to events as they occur to determine how group dynamics are affected by CSM state changes.
2. *Community nature*
   Community nature refers to the concept of "temporary" and "ongoing" communities. Some communities form and can last for years, while others are formed on demand and are discarded within minutes. For example, a community formed around a family will have significant longevity, but one formed among a group of people waiting at a bus stop is likely to be short-lived. The ability to distinguish between communities that are likely to be short-lived (e.g., location- or purpose-based) and those that are likely to be of indefinite duration (e.g., family-based) can assist in efficient community orchestration and life-cycle management.

   Algorithms that are specific to the ongoing/temporary nature of communities can autonomously/semi-autonomously drive the creation, configuration, and deletion of communities on behalf of the user. This relieves the user of lots of housekeeping tasks associated with the many communities to which they belong during their daily lives.

The community nature, i.e., the concepts of "ongoing" and "temporary", can be exploited by the system to assist an end user. Longevity of association between individuals is clearly a key driver for potential community formation and, in the online world, it may be acceptable for this to be a prerequisite. However, in the physical pervasive world, we need also to support ad hoc community formation for short-term goals which would not be possible with a longevity prerequisite. Discovery algorithms can look for brief, temporary communities at short intervals and longer-lasting and ongoing ones at longer intervals. The data used for this and how to analyze it varies depending on how long a community might be expected to last. Temporal and community nature semantics are attached to communities, and used to inform decisions such as deleting obsolete communities, configuring them, and creating new ones.

7.4.7 User Intent

The actions of people are invariably guided by a set of tasks that they need to complete over the short to medium term, e.g., buying fresh milk from the local shop. Ideally, for a system to fit in with the user's view of the world, it also needs to be aware of what the user is trying to achieve, i.e., their final goal. We refer to this as user intent.

User intent is not formally considered in social media. Instead they focus on the lower level of social interactions, without considering why these interactions take place. Some systems rely on manual entry of the tasks and subtasks, and require the user to continuously update these tasks with progress. In SOCIETIES, we attempt to capture intent by observing the user, and attempting to deduce what their intentions are.

Most user intent learning approaches aim to predict future user intentions based on raw context data. This is very difficult due to the gap between the low semantic level of raw context and the high semantic level of intentions.

User goals have been considered in pervasive systems, but predicting future behaviors has been limited to the level of the individual and not to a community. Relating this to the value proposition of SOCIETIES:

Discover: proactively discovers both individual and community intent.

Connect: once the intent of different users (e.g., a community) is discovered, we are able to effectively connect them by establishing a community with users who share similar intent.

Organize: provides community intent-aware services, and takes actions on behalf of a community.

1. *User intent prediction based on CRFs*
   Rather than directly using raw context data to learn intents, we introduce the concept of a situation to fill in the semantic gap. A recurring pattern in the context information yields a specific "situation" associated with a clique of raw context data.

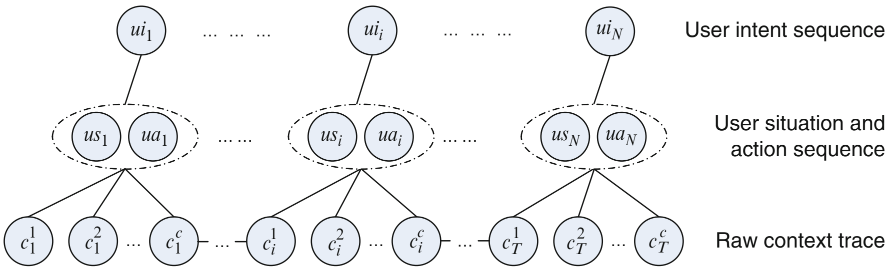


Fig. 7.4 Prediction based on CRFs

Then, based on the situation, as well as the real-time user action, user intent discovery and prediction are performed. Specifically, a 3-layer conditional random field (CRF) is constructed to model the relationship among the raw context data, the situations/actions, and the users' intentions, as shown in Fig. 7.4.

This model is capable of facilitating social interaction by proactively learning/detecting/predicting both individual and community intentions.

2. *Context-aware user intent prediction*
User intent is also predicted in SOCIETIES via a second approach, which also exploits context information but employs methods of statistical analysis. In this approach, the user behavior is modeled based on the user interactions with services. Text compression algorithms are used for identifying sequences of *user actions* and *user tasks*. Each user action is accompanied by a context data snapshot describing the situation of the user. Frequent occurring sequences of user actions, and their associated context snapshots, are grouped into user tasks. The user behavior model is described by stochastic models (variable-order Markov chains) that describe the sequences of user tasks and user actions and the respective transition probabilities.

History logs containing individual user interactions, along with context information are collected in a common context history repository. Alternatively, individual intent models can be provided. Learning algorithms, also used for the discovery of the user intent model of individuals, create a community-wide behavior model. The community model is used in order to enhance the intent prediction for individual users. Figure 7.5 provides an example of this concept.

Case A outlines a scenario where the history logs containing individual user's interactions, along with corresponding context information, are collected in a common context history repository. This common context history repository is used to discover a community intent model.

Case B outlines a scenario where individual intent models, which have been prior-discovered from the individuals' history logs, are provided as a substitute for the common context information, i.e., the community intent model is discovered from the aggregation of the individual intent models.

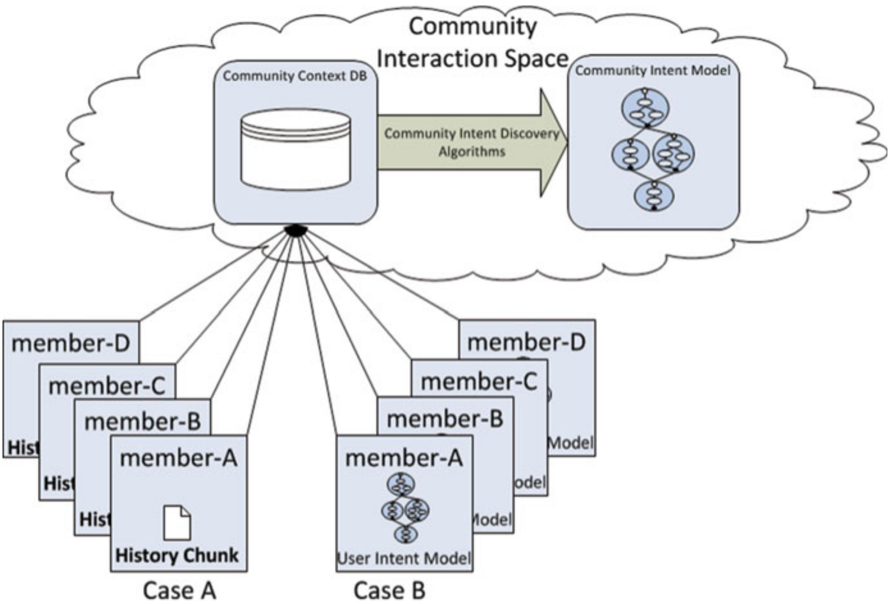


Fig. 7.5 Community intent model learning process

7.5 SOCIETIES Methodology

The SOCIETIES platform is intended to be used by individuals and communities. It is therefore imperative that the design is user-centered and, wherever possible, user-driven. The SOCIETIES methodology (Jennings et al. 2011) is based on the inclusion of three different user communities, who have engaged continually in the design and development process from initial user requirements gathering and early concept creation through to scenario refinements via early prototype evaluations. These are an *enterprise* (Lima et al. 2012) or *business* community, a *student* (Gallacher et al. 2012) or *young technology-wise* community, and a *disaster management* (Floch et al. 2012) or *emergency rapid response* community.

Early user requirements were elicited via ethnographic techniques, involving close observation of each community in their real-world environment, combined with questionnaires and participatory discussions. Requirements and scenarios were further refined through users' engagement with storyboards, and in the case of the students, a Wizard-of-Oz experiment. In the next steps of the project's user evaluation, each user community will evaluate the advanced system prototypes in situ in their real-world environments.

The *student community* is comprised of computer science students, for the most part avid social media users and naturally comfortable with many digital devices. They envision many potential uses for pervasive community systems, such as enhancing management of access to common services, such as libraries, and supporting collaboration both among students themselves, as well as between students and their lecturers. Students collaborated with developers through participatory workshops; using brainstorming, bodystorming, and storyboards to illustrate and investigate scenarios, while some students also opted to participate in a Wizard-of-Oz experiment that examined user acceptance of some proposed features and services of pervasive communities set in a campus environment. The results from the early student trials indicate that in general the students see value in joining pervasive communities. Clear controls over privacy settings and automation of services are a prerequisite for use of such a system according to this student community, but once these are in place, they are open to sharing some personal data and preferences, in order to add value to and leverage value from those communities.

Enterprise users are primarily interested in pervasive community services, which could enhance the workflow of their daily lives, with additional features for more selective relevant communications. These busy executives, in several cases, report being overloaded with communications and information, and the automatic filtering of both, without missing opportunities for useful connections, is attractive to them. Online community group discussions, observation, and storyboard-based workshops have supported the project's engagement with this group, and have led the project to focus on pervasive communities in a conference scenario. Benefits for enterprise attendees could potentially then include services such as personalized real-time information, such as agendas and navigation aides, events tailored to registered attendees' interests, live feedback during presentations, and maximizing networking opportunities. Many enterprise group executives indicate willingness to share some of their preferences with the system, while privacy, with regard to which preferences are shared for services, is clearly also a requirement.

The experienced *disaster management* community, with which we are engaged, is interested in any means of communication that supports the relief effort in a disaster zone. While emergency disaster situations are unpredictable by nature and difficult to envision, the project engaged with this group, composed of professional individuals from several different countries, during intensive assessment mission field training exercises, using mixed methods observation, questionnaires, discussions, and storyboard workshops. Relevant detailed and reliable information is important to disaster missions, and these disaster management professionals recognize great potential value in SOCIETIES providing this via pervasive communities and crowd computing. They are interested in tapping the expertise and knowledge of wider society in a disaster situation, for particular tasks, once that information can be filtered and effectively validated. They see advantages in the system coordinating the broad population who would like to help to assist a disaster mission, for certain tasks, such as checking satellite imagery for infrastructure damage, translation, and transcribing logbooks. Our research findings show it is important for disaster management users to feel in control of the system, and that they would prefer the system to make intelligent suggestions to support humans making decisions, rather than automation for most tasks.

In summary, the 'voice of the user' has been considered and continually integrated into the system design, through user-centered design and participatory methods. The next steps in the project will be to immerse the three user communities into prototypes of pervasive environments that have been derived and built from the scenarios and storyboards presented to them in the previous trials, and to evaluate their responses to these pervasive environments.

7.6 SOCIETIES Status, Future, and Exploitation

At the time of writing, the SOCIETIES project has completed a number of phases, and is now focused on initial implementation and integration of the platform described above.

Requirements and scenarios have been developed and verified with our user groups (SOCIETIES 2011), which encapsulate the critical use-cases of the project. Requirements have been specified in three specific domains – user requirements, technical requirements, and business requirements – each of which is combined to produce the overall requirement specification of this complex system. The requirements and scenarios for the project have been submitted to an initial user evaluation for verification and feedback (SOCIETIES 2011).

The SOCIETIES' system architecture has been specified along a number of interrelated lines of research: the overall system architecture that defines the high level architectural features of the SOCIETIES system, the service architecture that examines infrastructural and end-user service requirements, and finally interoperability architecture that specifies how SOCIETIES will integrate beyond its research remit.

The project core platform and intelligent enabling services have been designed in detail based on the architecture and requirements specified in the early phases of the project, and cover all aspects of the system design from client frameworks, communications management, device management, and service lifecycle management, to intelligent community orchestration, personalization, context management, and privacy & trust.

Development and integration of the SOCIETIES platform is driven by our commitment to carrying out trials with real users by the end of 2012, followed by a further research, development, and trial phase in 2013.

Looking to the future, the SOCIETIES project outputs (SOCIETIES 2011) can be exploited in many ways, from development of a full commercial community management system to integration of various enabling technologies into existing and emerging products, and to development of an open-source community around the results. SOCIETIES' stakeholders, identified in our business models, range from the mass-market to service providers, and onwards towards a number of secondary customer groups such as conference centers, emergency organizations, and universities.

In terms of mass-market adoption, SOCIETIES has involved real user communities throughout all stages of our research, from initial concepts and scenarios to fully developed software. The purpose of this is to identify real consumers' needs and wants, and their level of preparedness to adopt this highly pervasive technology.

Other stakeholders include brokers, who can harness the rich data produced by the SOCIETIES platform and offer it to various service providers. Service consumers, i.e., end users, require solutions to their discovery needs, their device interoperability and connection requirements, and the ability to bridge the gap between the real and virtual worlds. Service providers will have the capability to create new highly intelligent community-based services, as well as enhancing existing offerings with rich relevant information. Social network service providers can increase their SNS capabilities through the integration of SOCIETIES' enhanced discovery results, through the integration of connected communities, and through the provision of context-aware, relevant physical-world data to end users.

As noted previously, privacy is a major concern for many consumers. Consumers need to know that their data will be protected, and which communities have relevant trust levels based on personal requirements.

To support the adoption of SOCIETIES' results, the project is committed to providing external interfaces (APIs) to allow third-party entities to exploit all or parts of the system. The establishment of trust and other required customer relationships will facilitate the creation of a strong user base, as will integration with existing technologies. SOCIETIES is promoting itself to the Future Internet community, through various dissemination channels, in order to stimulate debate and discussion on the results. To increase consumer confidence in the SOCIETIES system, demonstrating a capability to monitor (in a privacy supporting manner) and learn about users in order to take proactive actions on their behalf is a key requirement. Proactive adaptation of a community's characteristics based on context, learned information, individual and group preferences, and membership management will enhance users' confidence in the value of the project results. Finally, the provision of exemplar community-based services, which do not exist in today's market, will further stimulate adoption.

7.7 Summary and Conclusion

In summary, SOCIETIES intends to bridge the gap between new social media and the virtual and real worlds, driving the adoption of pervasive technology through the discovery, connection, and organization of purpose-driven communities of interest. SOCIETIES believes that the individual elements of *discover*, *connect*, and *organize* make for a compelling overall value proposition that provides identified stakeholders with real value in today's competitive marketplace. SOCIETIES merges concepts from pervasive computing and social computing, breaking the physical–digital boundary, in a way that ensures the privacy of the user is safeguarded, letting users control their data, which in turn builds user trust.

Acknowledgments This work was supported in part by the European Commission as part of the SOCIETIES project under the FP7 programme which the authors gratefully acknowledge. The authors also wish to thank our colleagues in the project, without whom this paper would not have been possible (SOCIETIES 2011). Note that apart from funding this project, the European Commission has no responsibility for the content of this chapter.

References

DAIDALOS, (2012). *Project Website*. http://www.ist-daidalos.org. Accessed 16 Nov 2012.

Doolin, K., Roussaki, I., Roddy, M., Kalatzis, N., Papadopoulou, E., Taylor, N. K., Liampotis, N., McKitterick, D., Jennings, E., & Kosmides, P. (2012). SOCIETIES: Where pervasive meets social. *Future Internet Assembly Book, 1*, 30–41.

Floch, J., Angermann, M., Jennings, E., & Roddy, M. (2012). *Exploring cooperating smart spaces for efficient collaboration in disaster management*. Vancouver: ISCRAM.

Gallacher, S., Papadopoulou, E., Taylor, N.K., Blackmun, F.R., Williams, M.H., Roussaki, I., Kalatzis, N., Liampotis, N., & Zhang, D. (2011). Personalisation in a system combining pervasiveness and social networking. In *20th International Conference on Computer Communications and Networks (ICCCN 2011)*. Maui.

Gallacher, S., Papadopoulou, E., Taylor, N.K., Blackmun, F.R., & Williams, M.H. (2012). Intelligent systems that combine pervasive computing and social networking. In *9th IEEE International Conference on Ubiquitous Intelligence and Computing (UIC 2012)*. Fukuoka.

Gonçalves, J., Gomes, D., & Aguiar, R. (2012). Low-latency privacy-enabled context distribution architecture. In *IEEE Communication Software Services and Multimedia Applications Symposium (IEEE ICC 2012)*. Ottawa.

Greenfield, A. (2006). *Everyware: The dawning age of ubiquitous computing*. Berkeley: New Riders.

Hansmann, U., Merk, L., Nicklous, M., & Stober, T. (2012). *Pervasive computing: The mobile world* (2nd ed.). New York: Springer.

Jennings, E., Madden, E., Roddy, M., & Roussaki, I. (2011). Participative user research and evaluation methodologies for pervasive communities. In *Irish Human Computer Interaction Conference 2011 (iHCI 2011), Integrated Practice Inclusive Design (CIT)*. Cork.

Lima, C., Gomes, D., & Aguiar, R. (2012). Pervasive CSCW for smart spaces communities. In *IEEE International Conference on Pervasive Computing and Communications Workshops 2012 (PERCOM Workshops 2012)*. Lugano.

Lukowicz, P., Pentland, S., & Ferscha, A. (2012). From context awareness to socially aware computing. *IEEE Pervasive Computing, 11*(1), 32–41.

Obaidat, M. S., Denko, M., & Woungang, I. (2011). *Pervasive computing and networking* (1st ed.). Hoboken: Wiley.

Pallis, G., Zeinalipour-Yazti, D., & Dikaiakos, M. D. (2011). New directions in web data management. *Studies in Computational Intelligence, 331*, 213–234. doi:10.1007/978-3-642-17551-0_8. Online Social Networks: Status and Trends.

Parthasarathy, S., Ruan, Y., & Satuluri, V. (2011). Community discovery in social networks: Applications, methods and emerging trends. In *Social network data analytics* (pp. 79–113). Springer US. doi: 10.1007/978-1-4419-8462-3_4.

PERSIST. (2012). *Project Website*. http://www.ict-persist.eu. Accessed 16 Nov 2012.

Roussaki, I., Kalatzis, N., Liampotis, N., Kosmides, P., Anagnostou, M., Doolin, K., Jennings, E., Bouloudis, Y., & Xynogalas, S. (2012). Context-awareness in wireless and mobile computing revisited to embrace social networking. *IEEE Communications Magazine, 50*(6), 74–81.

Semertzidis, T., & Laso-Ballesteros, I.D.P. (2010). *Social networks: Current trends and research challenges, coordinated by the next MEDIA coordination action*.

SOCIETIES Deliverable D2.1 *Specification of initial scenarios and user requirements*. http://www.ict-societies.eu/files/2011/11/SOCIETIES_D21.pdf. Accessed 16 Nov 2012.

SOCIETIES Deliverable D2.2 *Scenario description, use cases and technical requirements specification*. http://www.ict-societies.eu/files/2011/11/SOCIETIES_D22.pdf. Accessed 16 Nov 2012.

SOCIETIES Deliverable D8.1 *Paper trial evaluation report*. http://www.ict-societies.eu/files/2011/11/D8.1_public.pdf. Accessed 16 Nov 2012.

SOCIETIES deliverables. http://www.ict-societies.eu/project-deliverables/. Accessed 16 Nov 2012.

SOCIETIES Integrated Project. *Full title: Self Orchestrating CommunIty ambiEnT IntelligEnce Spaces*, http://www.ict-societies.eu. Accessed 16 Nov 2012.

SOCIETIES project partners. http://www.ict-societies.eu/partners/. Accessed 16 Nov 2012.

Xu, B., Chin, A., Wang, H., Chang, L., Zhang, K., Yin, F., & Zhang, L. (2011). Physical proximity and online user behaviour in an indoor mobile social networking application. In *Internet of Things (iThings/CPSCom), 2011 International Conference on and 4th International Conference on Cyber, Physical and Social Computing* (pp. 273–282), doi: 10.1109/iThings/CPSCom.2011.74.

Zaphiris, P., & Ang, C. S. (2009). *Social computing and virtual communities*. Hoboken: Chapman and Hall/CRC/IGI Global.

Chapter 8
Data Analysis on Location-Based Social Networks

Huiji Gao and Huan Liu

Abstract The rapid growth of location-based social networks (LBSNs) has greatly enriched people's urban experience through social media, and attracted increasing number of users in recent years. Typical location-based social networking sites allow users to "check in" at a physical place and share the location with their online friends, and therefore bridge the gap between the real world and online social networks. The availability of large amounts of geographical and social data on LBSNs provides an unprecedented opportunity to study human mobile behavior through data analysis in a spatial–temporal–social context, enabling a variety of location-based services, from mobile marketing to disaster relief. In this chapter, we first introduce the background and framework of location-based mobile social networking. We next discuss the distinct properties, data analysis and research issues of location-based social networks, and present two illustrative examples to show the application of data mining to real-world location-based social networks.

8.1 Introduction

The wide use of mobile devices and location-based services in the world has generated a new concept of online social media, namely location-based social networks (LBSNs). Location-based social networking sites use GPS, Web 2.0 technology and mobile devices to allow people to share their locations (usually referred to as "check-in"), find out local points of interest and discounts, leave comments on specific places, connect with their friends, and find other friends who are nearby. A recent survey from the Pew Internet and American Life Project reports that over the past year, smartphone ownership among American adults has risen from 35 %

H. Gao (✉) • H. Liu
Computer Science and Engineering, Arizona State University, Phoenix, USA
e-mail: Huiji.Gao@asu.edu; Huan.Liu@asu.edu

A. Chin and D. Zhang (eds.), *Mobile Social Networking: An Innovative Approach*, Computational Social Sciences, DOI 10.1007/978-1-4614-8579-7_8,

in 2011 to 46 % in 2012. Almost three-quarters (74 %) of smartphone owners use their phone to get real-time location-based information such as getting directions or recommendations. Meanwhile, 18 % of smartphone owners use geo-social services, such as Foursquare,[1] Gowalla,[2] and Facebook Places,[3] to "check in" to certain locations and share them with their friends, this percentage having risen from 12 % in 2011 (Zickuhr 2012). It is anticipated that more than 82 million users will subscribe to location-based social networking services by 2013 (ABI Research 2008), and location-based marketing will be a $1.8 billion business worldwide by 2015 (ABI Research 2010). Such rapid growth of location-based social networks has led to the availability of a large amount of user data, which consists of both the geographical trajectories and the social friendships of users, providing both opportunities and challenges for researchers to investigate users' mobile behavior in spatial, temporal, and social aspects.

Typical online location-based social networking sites provide location-based services that allow users to "check in" at physical places, and automatically include the location into their posts. "Check-in" is an online activity that posts a user's current geographical location to tell his friends when and where he is through social media. Compared with many other online activities (following, grouping, voting, tagging, etc.) that interact with the virtual world, "check-in" reflects a user's geographical action in the real world, residing where the online world and real world intersect. In this scenario, "check-in" not only adds a spatial dimension to the online social networks, but also plays an important role in bridging the gap between the real world and the virtual world. Thus, the study of check-ins on location-based social networks provides an ideal environment to analyze users' real world behavior through virtual media, and could potentially improve a variety of location-based services such as mobile marketing (Barnes and Scornavacca 2004; Bauer et al. 2005; Scharl et al. 2005), disaster relief (Goodchild and Glennon 2010; Gao et al. 2011a, b), and traffic forecasting (Ben-Akiva et al. 1998; Dia 2001).

The first commercial location-based social networking service available in the United States is Dodgeball,[4] launched in 2000. It allows users to "check in" by broadcasting their current locations through short messages to their friends who are within a ten-block radius; users can also send "shouts" to organize a meeting among friends at a specific place. After being acquired by Google in 2005, the original Dodgeball was replaced with Google Latitude in 2009, while the founder of Dodgeball launched a new location-based social networking service, "Foursquare", in the same year. Foursquare utilizes a game mechanism in which users can compete for virtual positions, such as mayor of a city, based on their check-in activities. It reached 20 million users by April 2012 (Kessler 2012), becoming one of the most successful location-based social networking sites in the United States. Facebook

[1] http://foursquare.com

[2] http://en.wikipedia.org/wiki/Gowalla

[3] http://www.facebook.com/about/location

[4] http://en.wikipedia.org/wiki/Dodgeball

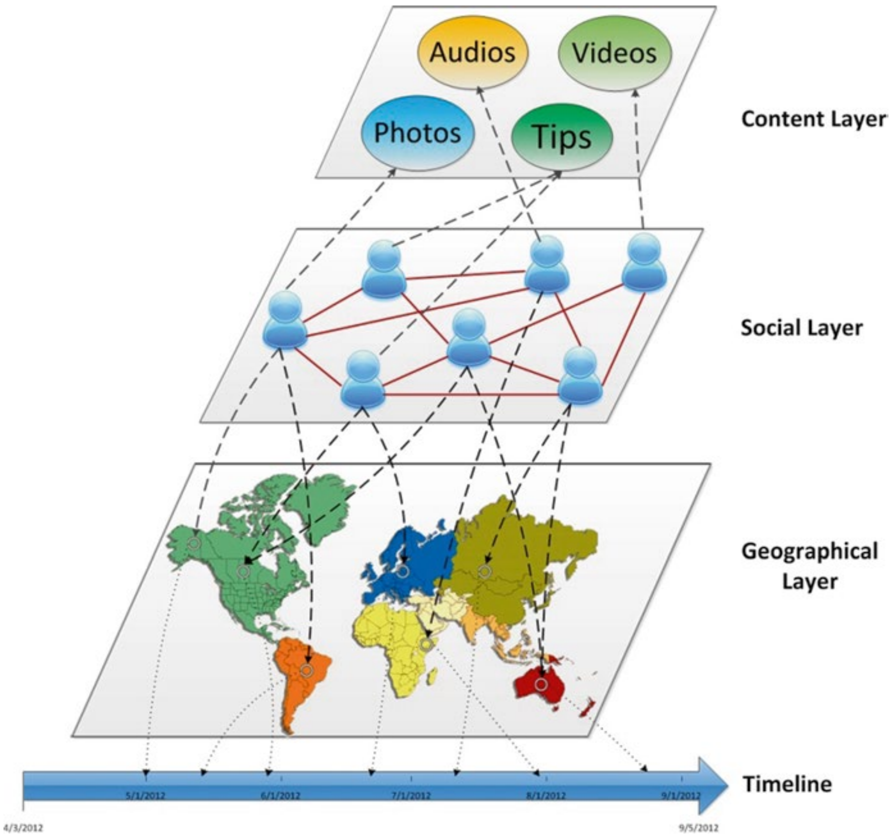


Fig. 8.1 The information layout of location-based social networks

also launched its location-based service, namely Facebook Places, in 2010 with its check-in function, and acquired another popular LBSN, Gowalla,[5] at the end of 2011. All these location-based social networking sites share a "3 + 1" framework, i.e., three layers and one timeline, as shown in Fig. 8.1.

The geographical layer contains the historical check-ins of users, while the social layer contains social friendship information, and the content layer consists of user feedbacks or tips about different places. All these three layers share one timeline, indicating the temporal information of the user "check-in" behavior. Previous research has investigated the social and content layers with traditional online social network data (Hu and Liu 2012), and analyzed the geographical and content layers with mobile phone data (Chen and Kotz 2000). Compared to them, location-based social network data has an additional geographical layer which is not available in traditional online social networks, and an explicit social layer which is not available

[5] http://www.pcmag.com/article2/0,2817,2401433,00.asp

from mobile phone data (usually social friendship information from mobile phone data is derived through smartphone proximity network). The unique geographical property and the social network information presents new challenges for data analysis on location-based social network data, since traditional approaches on social network or mobile phone data may fail due to the lack of pertinence. Furthermore, the "3 + 1" data structure defines six different types of networks, i.e., location–location network, user–user network, content–content network (e.g., word–word network), user-location network, user-content network, and location-content network. Each one can be mined together with the temporal information provided by the timeline, indicating more opportunities for data analysis on LBSNs. Therefore, data analysis techniques specifically designed for LBSNs can efficiently deal with these distinct properties, and help understand user behavior for research and business purposes.

The rest of this chapter is organized as follows. We first introduce the distinct properties of location-based social network data in Sect. 8.2, then discuss the data analysis and research issues in Sect. 8.3, followed by two real-world examples of applying data mining to location-based social networks in Sect. 8.4, and finally provide some conclusions with suggestions for future work in Sect. 8.5.

8.2 Distinct Properties of Location-Based Social Network Data

Location-based social networks provide data consisting of both geographical information and social networks. Compared to traditional online social network data and mobile phone data, location-based social network data have distinct properties in several aspects.

8.2.1 Geographical Property

One of the most significant differences between LBSNs and traditional online social networks is the geographical property, which is considered as the unique facet of location-based social networks. Users on LBSNs are able to check in at a physical place, and let their friends be aware of this check-in. The check-in location indicates the current geographical status of a user in the real world, and generates the local social networks of the user based on this location. In this scenario, the geographical check-in locations bridge the gap between the real world and online social networks (Cranshaw et al. 2010; Gao et al. 2012a), which in turn reflect the user's behavior more closely to the real world compared with other online social networks, and provide an unprecedented opportunity to study a user's real-world behavior through social media. Researchers have studied the distinctions between online and offline social networks (Cranshaw et al. 2010), differences between location-based social networks and content-based social networks (Scellato et al. 2010), and relationship between geographical distance and friendship (Scellato et al. 2011b; Cho et al. 2011), etc. These analyses exploit many fundamental user mobile patterns, and motivate us to make use of geographical properties for the development of better location-based services.

1. *Large-Scale Mobile Data*
   The increasing use of mobile devices and popular location-based mobile social networking sites has led to the massive availability of mobile data. Compared with the traditional cell phone data, which is usually collected through telecommunication carriers with limited number of users (Zheng et al. 2009), location-based social networking services utilize Web 2.0 technology combined with GPS on mobile devices, generating a large amount of geographical and social information from millions of users (Chang and Sun 2011; Scellato et al. 2011b). For example, Google Latitude reported ten million active users in 2011,[6] Yelp had approximately 71 million unique visitors monthly on average in the first quarter of 2012,[7] and Foursquare reached 20 million users and two billion check-ins by April 2012 (Kessler 2012). Researchers can easily obtain these data through public APIs provided by location-based social networking sites, enabling the large-scale data analysis of user behavior in a spatial, temporal, and social context (Cheng et al. 2011; Scellato et al. 2011b; Gao et al. 2012a).
2. *Accurate Description of Geolocations*
   Location-based mobile social networking sites provide more accurate location descriptions than traditional geo-tagged data. For example, in location-based social networks, it is easy to distinguish two adjacent restaurants on a street, two nearby stores in a fashion square, or a pharmacy located upstairs of a bar. This is because the traditional geo-tagged data only provide the longitude and latitude of a location, while location-based social networking sites such as Foursquare and Facebook places could provide additional textual descriptions for popular venues, e.g., categories, comments, and tips, therefore promoting a variety of location-based applications from location recommendation (Ye et al. 2011a) to urban computing (Cranshaw et al. 2012) by endowing the physical places with semantic meaning.
3. *Data Sparseness*
   In traditional cell phone data, a user's geographical location is automatically recorded by the telecommunication tower, while on location-based social networks, the check-in process is user-driven (Noulas et al. 2011a), i.e., the user decides whether to check in at a specific place or not due to certain privacy concerns. For example, a user may usually check in at Starbucks in New York, but with the latest check-in at SeaWorld in San Diego, or check in continuously at the same restaurant many times. Some users even have more than 1-year gaps between consecutive check-ins. Such check-in behavior leads to the significant sparseness of geographical data in location-based social networks, which greatly increases the difficulty of data analysis, especially in investigating human mobility patterns.

[6] http://techcrunch.com/2011/02/01/google-latitude-check-in

[7] http://www.yelp-press.com/phoenix.zhtml?c=250809&p=irol-press

4. *Explicit Social Friendship*
   The social networks on location-based social networking sites consist of social friendship information explicitly defined by users (a user can explicitly add another user as a friend), while in traditional cell phone data, the social network is usually collected through user study (Li et al. 2008; Eagle et al. 2009), or derived from communication network or Bluetooth network (Wang et al. 2011). This property enables more accurate and efficient data analysis and evaluation on location-based social networks, especially for applications such as friend recommendation and location privacy control (Kelley et al. 2008).

8.3 Data Analysis and Research Issues of Location-Based Mobile Social Networks

The heterogeneous data in location-based social networks contain spatial–temporal–social context and present new challenges and opportunities for data analysis. One can ask many interesting questions that can potentially be answered by analyzing LBSN data. For example, are there any relationships between user attitudes and mobile patterns on LBSNs? How does geographical distance affect online social friendship, and vice versa? Why do people use location-based social networking services? Under what circumstances would users not like to share their locations due to privacy issues? Can location prediction help mobile marketing? Can location-recommender systems improve urban experience? How can one best control location privacy to maximize her social networking experience? In this section, we introduce a variety of data analysis techniques and current research on location-based social networks, and show how answers to these challenging questions can be obtained via novel data analysis to improve location-based services.

8.3.1 Social Friendship and Geographical Distance

Traditional social networking analysis mainly studies network structure and properties, which does not consider the geographical distance between nodes. In 2001, Cairncross (2001) proposed the term "the death of distance", claiming that geographical distance begins to play a less important role due to the communication revolution and the rapid development of the Internet, which therefore could lead our world to a "global village". Later, Gastner and Newman (2006) studied the spatial structure networks. They demonstrated that there is a strong correlation between geographical attributes and network properties, indicating the significance of considering the spatial properties of networks for future applications. Other researchers studied geographical distance in the Internet, and argued that the IT revolution does not transfer us into a borderless society, as physical proximity still plays an important role in the Internet era (Goldenberg and Levy 2009; Mok et al. 2010). All these studies are based on traditional networks such as e-mail networks, cell phone contact networks, road networks, and the Internet.

One of the first attempts to investigate how social connection is affected by geographical distance in online social networks was proposed by Liben-Nowell et al. (2005). The authors studied users' social networks and their hometown information obtained from LiveJournal. Their simulation model shows that one-third of friendships are independent of geography. With the wide use of mobile devices, such as Apple iPhones and Google Android phones, and the increasing attention on mobile social networking, location-based social networks focused on the small local social network derived from a user's geographical location become more and more popular. Dodgeball was the first commercial location-based social network service available in the United States, launched in 2000. Humphreys (2007) studied user behavior on Dodgeball, and found that LBSNs do change people's attitude toward locations and their experience of urban life.

The increasing popularity of location-based social networking sites makes it possible to obtain data consisting of the geographical distance between users and their social networks in large-scale, which in turn enables a vast research opportunity for large-scale data analysis on geo-social properties in LBSNs. Scellato et al. (2010) proposed two geo-social metrics, embedding the geographical distance into social structure, to measure the node locality and geographical clustering coefficient. Two findings are presented in this work: (1) users who live close have a higher probability to create friendship links than those who live at a distance, and (2) users in the same social cluster show short geographical distances. Furthermore, the authors compared location-based social networks (Brightkite and Foursquare) with content-sharing-based social networks (LiveJournal and Twitter), discovering the difference of network properties between these two kinds of social networks. They found that people within a social cluster on the LBSNs tend to have smaller geographical distance than those online social networks focusing on content producing and sharing.

Researchers have also investigated how geographical distance influences social networks, and how social networks influence human movement on LBSNs. Scellato et al. (2011b) presented a comprehensive study on three location-based social networking sites, i.e., Brightkite, Foursquare, and Gowalla. They observed strong heterogeneity across users with different geographic scales of interaction across social ties, with the probability of a social tie between two users as a function of the geographical distance between them. Cho et al. (2011) studied Gowalla, Brightkite, and cell phone data, reporting that long-distance travel is more influenced by social friendship, while short-range human movement is not influenced by social networks. More recently, Kulshrestha et al. (2012) investigated the Twitter social network, and concluded that offline geography still matters in online social networks, while one-third of the users would like to have their social links in other countries, which is consistent with the previous findings presented in Liben-Nowell et al. (2005) and Scellato et al. (2010). Brown et al. (2012) extended the research on LBSNs to social community, and discovered that the rise of social groups is affected by both social and spatial factors. They reported that social communities on location-based social networks seem to be more relevant to the spatial factor. This is also consistent with previous findings (Scellato et al. 2010) about the differences between location-based social networks and content-sharing-based social networks.

8.3.2 User Activity and Mobile Pattern Analysis

Sociologists have studied the characteristics of user behavior on location-based social networks, motivated by the potential power of these characteristics for future research and applications. Among the current research, there are two major characteristics that sociologists mostly discussed, i.e., user activity and mobile patterns.

1. *User Activity*
   User activity indicates how frequently a user creates and consumes online content in LBSNs. Researchers attempt to classify users into various groups, representing different levels of user activity. This is motivated by tailoring location-based services to different user types to benefit the majority of users. One of the first large-scale analyses of user activity on a real-world commercial location-based social network was presented in Li and Chen (2009). The authors analyzed user profiles on Brightkite, and observed that the majority of users are male users who are professionals and willing to participate in social media. They also found that users with higher network degree tend to be more mobile and active. The authors further clustered users based on their attributes such as total number of updates, uniquely visited places, etc., and obtained five user groups according to user activity, named as inactive, normal, active, mobile, and trial users. They reported that the majority of users on Brightkite are trial users, while only 6 % of users are clustered as active users. Noulas et al. (2011b) used a spectral clustering algorithm to group users based on their check-in category distribution on Foursquare, aiming at identifying user communities to help develop new applications such as recommender systems.

   Vasconcelos et al. (2012) considered different type of features for user clustering on Foursquare. They focused on the tips, dones, and to-dos of venues, and utilized three related attributes to cluster users, i.e., the number of tipped venues, the total number of dones and to-dos, and the percentage of tips with links. They obtained four groups, with three groups based on user activity level, and one group representing spam users. It is reported that around 86 % of users tend to tip a larger number of venues and get more dones and to-dos in return, forming the largest group on Foursquare. Furthermore, the authors showed that observing a large number of links pointed to unrelated content in tips can be a good predictor for detecting spam users.
2. *Mobile Patterns*
   Cheng et al. (2011) explored millions of check-ins on Facebook, and observed various spatial, temporal, and social patterns. For example, human movement follows a “Lévy Flight” (Rhee et al. 2011), in which people tend to move to nearby places and occasionally to distant places. The authors observed that user

mobility is influenced by social status, geographical, and economic factors. Furthermore, the user check-in behavior presents strong daily/weekly patterns and periodic property, indicating the potential to improve location-based applications. In Noulas et al. (2011a), the authors observed similar geo-temporal patterns of check-ins on weekdays and weekends. They reported that around 20 % of consecutive check-ins in Foursquare happen within 1 km of one another, 60 % between 1 and 10 km, and 20 % over 10 km. Li and Chen (2009) studied users' mobility characteristics on Brightkite. They clustered users based on their mobility patterns derived from user updates and movement paths, and obtained four user groups, namely home users, home–vacation users, home–work users, and other users which present different mobility patterns from previous groups.

8.3.3 Location Prediction

Location prediction is a traditional task in mobile computing. It has been studied over a long period. Researchers analyze human mobility patterns to improve location prediction services, and therefore exploit their potential power on various applications such as mobile marketing (Barnes and Scornavacca 2004; Barwise and Strong 2002), traffic planning (Ben-Akiva et al. 1998; Dia 2001), and even disaster relief (Gao et al. 2011a; Goodchild and Glennon 2010; Gao et al. 2012a; Wang and Huang 2010). Current research on location prediction in LBSNs mainly focuses on two tasks: (1) predicting a user's home location, and (2) predicting a user's location at any time. The former task considers the static home location of a user, while the latter considers more about a user's moving trajectories, with his location in movement.

Before we delve into different location prediction methods, we first discuss two commonly used evaluation metrics in the location prediction task. The first metric is *prediction accuracy*, i.e., the fraction of correctly predicted locations over the total number of predicted locations in the testing set, which has been widely used in current work (Gao et al. 2012a; Cho et al. 2011; Backstrom et al. 2010). Sometimes its variants have also been used for additional evaluation. For example, the top-k accuracy is utilized in Cheng et al. (2010). It returns the top k candidates as the predictions for a location, and treats a prediction as correct as long as the ground truth location is among the top k returned locations. Here, k is usually selected as 2, 3, 5, and 10. The second metric is *expected distance error* (Cho et al. 2011), as shown below, which computes the average geographical distance between the real location and the estimated location, over all predicted locations.

$$ErrD = \frac{1}{|L|} \sum_{l P L} d\left(l_{act}, l_{est}\right) \quad (8.1)$$

where L is the unknown locations in the testing set, l_{act} is the actual location, and l_{est} is the estimated location. d(x,y) is a function that computes the geographical distance between two locations x and y.

The motivation of home location prediction arises from the sparseness of available user home locations on popular social networks such as Twitter and Facebook. Based on the statistics from Cheng et al. (2010), only 26 % of Twitter users list their locations as granularly as a city name, and less than 0.42 % of all tweets use the geo-tagging function to indicate their locations. On the other hand, the availability of user home location leads to a user-centric social network. It provides an opportunity to study social networks from a user's ego view, and in turn benefits applications such as targeting advertisement regions, and summarizing the local news for nearby users. Therefore, obtaining the user home location is critical to studying human mobility on location-based social networks.

Current work in home location prediction on LBSNs uses two kinds of resources, i.e., content information and social network information. The content-based approaches (Cheng et al. 2010; Hecht et al. 2011) studied the location information implicated in a user's tweet content, and proposed a location prediction framework based on the correlation between specific terms in tweets and their corresponding locations.

Backstrom et al. (2010) utilized social network information on Facebook to predict the user's home location. They predicted a Facebook user's home address based on the provided home addresses of his friends. One observation was leveraged so that the probability of a link being present between two nodes is a function of their geographical distance. By maximizing the likelihood of observations on friendship and non-friendship of a user, the unknown home location could be computed according to friends' addresses. All these methods predict the location at country, state, or city level, while the spatial resolution is low.

To predict a user's location at any time, usually referred to as *next location prediction*, various approaches have been proposed in the last decade. Without the social network information being available, these methods mainly consider the spatial trajectories (Monreale et al. 2009; Spaccapietra et al. 2008), temporal patterns (Thanh and Phuong 2007), or spatial–temporal patterns (Scellato et al. 2011a; Gao et al. 2012c) for location prediction. With the availability of social information on LBSNs, Gao et al. (2012a) proposed the first work of modeling social information for next location prediction on LBSNs with a social-historical model. Later, Noulas et al. (2012) further investigated the next location prediction problem and proposed a set of features regarding various facets of user behavior for prediction. Researchers have made a great effort to investigate the role of social friendship in explaining a user's mobile patterns. On the other hand, leveraging social networking information for location prediction becomes a new challenge, since how to embed the social property into geographical patterns is still an open issue on location-based social networks (Gao et al. 2012b).

Current work on LBSNs has proposed various approaches to combing social network information with traditional spatial–temporal patterns. Chang and Sun (2010) utilized logistic regression model to combine a set of features extracted from Facebook data. The features include a user's previous check-ins, user's friends' check-ins, demographic data, distance of place to user's usual location, etc. Their results demonstrated that the number of previous check-ins by the user is a strong predictor, while previous check-ins made by friends and the age of the user are also good features for prediction.

Linear combination has been mostly used for integrating social friendship with spatial–temporal patterns (Cho et al. 2011; Gao et al. 2012a). Cho et al. (2011) considered the user check-in probability as a linear combination of social effect and non-social effect. The social effect assumes the check-in of a user to be close to the check-ins of his friends, both in space and in time; while the non-social effect captures the periodical patterns, which considers the user's personal movement following a 2-D Gaussian distribution, with the two Gaussian centers focusing on home and work. Gao et al. (2012a) proposed a social–historical model integrating the social ties and historical ties of a user for location prediction. Both ties generate the probability of next location based on the observation of previous check-in sequence. The historical ties consider the user's own check-in sequence, and the social ties consider the check-in sequences of the user's friends. Based on the observation that word sequence and location trajectory share a set of common properties, a language model is then introduced for generating the next location probability.

All of the current work reports very limited improvement by utilizing social network information in LBSNs. The model that considers social networks slightly improves those that do not consider social networks. However, this does not lead to the conclusion that social network has no contributions to a user's mobility. The best way to integrate the social network and leverage it for location prediction is still under study.

8.3.4 Recommender Systems

Recommender systems are designed to recommend items to users in various situations such as online shopping, dating, and social events. Since the exploration of city and neighborhood provides us with more choices of life experience than before, recommendation is indispensable to help users filter uninteresting items, and therefore reduce their time in decision-making. Furthermore, recommender systems could also benefit virtual marketing, since the appropriate recommendations could attract users with specific interests. Recommender systems on location-based social networks only started just a few years ago, and three items are mainly recommended in current work, which are locations, tags, and friends.

1. *Location Recommendation*
   Location recommendation aims to recommend a set of locations to a user based on the user's interests. The major difference between location prediction and location recommendation is that location prediction usually predicts the next location as an existing location that the user has been before, while location recommendation would recommend a new location that the user has never been before. From a research standpoint, location prediction on LBSNs considers more how to utilize the social information, while current research in location

recommendation on LBSNs mainly focuses on the geo-spatial and temporal influence, and the social network information is usually utilized through traditional collaborative filtering (Berjani and Strufe 2011; Zhou et al. 2012), which considers the location as an item such as that on Epinions (Tang et al. 2012a, b). For evaluation, performance@N (Ye et al. 2011c) is usually adopted to evaluate the location recommendation performance. The performance@N metric consists of precision@N and recall@N. It consider all the locations that should be recommended as uncovered locations, and the set of correctly recommended locations as recovered locations. The precision@N evaluates the ratio of recovered locations to the N recommended locations, and the recall@N calculates the ratio of recovered locations to uncovered locations.

Ye et al. (2010) first introduced location recommendation on location-based social networks. In this paper, the major focus is location recommendation efficiency. The essential content contains: (1) only friendship information was used for collaborative filtering, and (2) instead of calculating the user similarity based on historical behavior (e.g., check-in history), the authors captured the correlations between geographical distance and user similarity, and leveraged them for user similarity calculation. This work is later extended in Ye et al. (2011c), which considers both spatial influence and social friendships for location recommendation. Three factors are investigated and combined to recommend locations. The first factor represents influence from similar users, the second factor indicates influence from friends, and the third factor captures geographical influence, under the hypothesis that people tend to visit close places more often than distant places. A spatial constraint is generated to capture the geographical influence by exploiting the relationship between a user visiting two places and the geographical distance between these two places. These three factors are then represented by three probabilities, and linearly combined together with corresponding weights. The results demonstrated that the most influential factor actually comes from the similar users, while friendship and geographical distance together have around 30 % influences.

2. *Tag Recommendation*
Tag recommendation is motivated to enrich the semantic meaning of places and to facilitate the development of recommender systems such as “Point of Interest” retrieval services. Temporal patterns have been usually considered for tag recommendation on location-based social networks. In Ye et al. (2011a), the authors proposed “temporal bands” to capture the temporal patterns of each place, and suggested their potential ability for tag recommendation. For example, a bar may be visited frequently at 11:00 p.m. to 1:00 a.m., while a restaurant may have more visits around 12:00 p.m. and 6:00 p.m. Therefore, tags associated with the bar or restaurant present different visiting distributions over time, i.e., temporal bands. By considering the visiting probability at different hours of a day and different days of a week, one can compare such visiting distributions between candidate tags and target places; the recommender system could then recommend a set of tags that mostly fit the temporal band of that place. In this work, the authors only proposed the idea of temporal bands, but did not apply it to real-world datasets for tag recommendation.

In Ye et al. (2011b), the temporal information has been formally utilized for tag recommendation and place annotation. In this work, the authors considered tag recommendation as a classification problem. Two sets of features, named explicit patterns and implicit patterns, are firstly defined to generate the feature space for each place, then a SVM classifier is learned for each tag, based on the observed feature vectors that are associated and not associated with the tag. The explicit patterns include features that can be explicitly observed in the data, e.g., total number of check-ins, total number of unique visitors, etc. The implicit patterns generate the relatedness between two places based on their common visiting users and common temporal patterns, while the latter factor is similar to Ye et al. (2011a). These two factors are linearly combined together, which generates a ranking list of places based on their relatedness to the target place. A place with high relatedness is referred to as a semantic neighbor, and the corresponding relatedness indicates the probability of the target place to be labeled with a given semantic tag from this neighbor. The final implicit patterns are the probabilities for each possible tag on the target place. The hypothesis of this method is that two places checked in by the same user around the same time should have strong relatedness, and therefore share more common tags. The experiment showed that most people follow the same temporal patterns in visiting places, while the explicit and implicit features both need to be considered for tag recommendation.

3. *Friend Recommendation*
Friend recommendation analyzes the similar patterns between a target user and other users, and then recommends users with the most similar patterns to the target user. Here, the similar patterns may represent the common interests, shopping habits, traveling trajectories, etc. Friend recommendation on location-based social network mostly uses supervised learning in terms of link prediction. A set of features is firstly extracted from the historical data for each pair of users, and then a classifier is trained based on the extracted features and finally used to predict the link between two users. The social network information is used as ground truth to evaluate their proposed approaches, and ROC curves (Scellato et al. 2011c; Sadilek et al. 2012a) are usually used as evaluation metrics.

Current work on friend recommendation differs in how to choose the feature space and classifier. Chang and Sun (2011) used logistic regression to predict the link between two users who have co-locations. Feature extraction was based on the tuples of (place x, actor1, actor2), indicating that actor1 and actor2 have checked-in into place x at least once. Three features are extracted: the total number of check-ins at place x, and numbers of check-ins of actor1 and actor2 respectively. Cranshaw et al. (2010) extracted 67 features from the data on Locaccino (Sadeh et al. 2009) for each co-location observation between two users. Their features include intensity and duration, location diversity, mobility regularity, structure properties, etc., with respect to co-location properties and user attributes. Three classifiers are selected for predicting the link, while the results show that AdaBoost has the best classification performance. They also reported that there is a positive correlation between the location diversity and the number of social ties a user has in the social network, and that considering the number of co-locations between two users is not sufficient for friend recommendation.

Sadilet et al. (2012a) adopted a similar scenario, while in addition considering the content features from tweets. Scellato et al. (2011c) exploited the place features such as common check-ins, social features like common friends, and global features such as distance between homes, then adopted various classifiers in WEKA for link prediction on Gowalla. Their results demonstrated that the purely social-based features contribute least to the prediction performance, while space features and global features lead to better performance, indicating the importance of location-based activities on location-based social networking analysis.

8.3.5 Location Privacy

Location sharing is an indispensable function of location-based social networking services. Users share their locations by checking in on location-based social networking sites to let their friends know where they are and when. The location awareness can then form location-based social networks and enhance the user's social connections. For example, a user may want to hang out with his friend after learning he is nearby through his check-in status. On the other hand, while location sharing significantly enhances user experience in social networks, it also leads to privacy and security concerns. In recent years, location privacy on location-based social networks has attracted more and more attention from both academia and industry. Previous work (Lederer et al. 2003; Consolvo et al. 2005; Gundecha et al. 2011; Tsai et al. 2009) has found that privacy is a critical concern for user considering adopting location-sharing services. When using location-sharing services, some users would like to share their location with friends for social purposes, while other users may believe that sharing personal location discloses one's personal preferences and movement track, which may cause potential physical security risks. Therefore, it is inevitable to consider privacy control when designing location-sharing applications.

Researchers are interested in understanding users' preference regarding location privacy in location-based social networks, such as why people are using location-sharing services and under what circumstances they do not want to share locations, therefore improving the design of new location-sharing applications. Humphreys (2007) analyzed user behavior on Dodgeball by conducting interviews with 21 Dodgeball users, and discovered that location-based social services do influence the way people experience urban public places and their social relations. Lindqvist et al. (2011) explored how and why people use Foursquare through interviews and surveys of Foursquare users, and reported five major factors that explain the reasons: i.e., badges and fun, social connection, place discovery, keeping track of places, and competition with themselves. Furthermore, the authors also found that the majority of users had few privacy concerns, and users choose not to check in at specific locations mainly because the places are embarrassing, non-interesting, or sensitive.

Mobile applications have also been developed to help manage privacy on LBSNs. Toch et al. developed a location sharing application "Locaccino",[8] focusing on privacy

[8] http://locaccino.org

control based on the Facebook social network (Toch et al. 2010b; Sadeh et al. 2009). A Locaccino user can request the location of his Facebook friends. It allows a user to set detailed location-sharing privacy preferences, such as when and where his location can be visible to a set of pre-specified users. Toch et al. (2010a) utilized the data collected from Locaccino to investigate the location factors that influence users' location-sharing preferences. They deployed Locaccino to a set of participants, and conducted surveys on them. Their analysis showed that locations with higher location entropy (Cranshaw et al. 2010) (a measure that is utilized to evaluate the user diversity of a location: higher location entropy indicates the location has been visited by a diverse set of unique users) are more comfortable for users to share, while highly mobile users receive more requests from their friends for location sharing. Kelley et al. (2008) introduced a machine learning approach to control the sharing policy. They proposed a Gaussian Mixture based method to classify the privacy control policies of users, with evaluation on Locaccino data from 43 users and 124 pre-defined privacy policies. The prediction accuracy is chosen as the evaluation metric.

8.3.6 Related Efforts

Aside from the topics discussed in the previous sections, even more efforts have been made in mining location-based social networks. In event detection, Sakaki et al. (2010) constructed an earthquake reporting system in Japan to report earthquakes using an event detection algorithm. They considered each user who makes tweets about a target event to be a sensor of the event, and proposed a spatial–temporal model to track the event center and trajectory. De Longueville et al. (2009) utilized twitter data to analyze the spatial, temporal, and social dynamics and URL property of events related to the Marseille forest fire, aiming to investigate the potential power of leveraging Twitter for emergency planning and disaster relief.

In geographical topic analysis, researchers utilize generative models, which are combined with spatial–temporal regularities to explore the space–time structures of topical content (Pozdnoukhov and Kaiser 2011), or devised with embedded content, user preference, and geographical locations to model tweet density (Hong et al. 2012), or generated as a combination of geographical clustering and topic model to discover and compare geographical topics (Yin et al. 2011). However, among all these works, social network information is not utilized, and the evaluation of the geo-topic model is also controversial to a certain extent.

In urban computing, Cranshaw et al. (2012) developed an online system, Livehoods,[9] to explore the social dynamics of the city and reveal the different characterized regions. The authors used a spectral clustering approach to cluster the check-in locations from 18 million check-ins into different areas, with each one

[9] http://livehoods.org

representing the character of lifestyle in that area. Sadilek et al. (2012b) modeled the spread of disease through Twitter data. They proposed a detection framework to identify the sick individual based on tweet content, and showed that there is a strong correlation between a person's number of infected friends and his probability of getting sick, where the probability increases exponentially as the number of infected friends grows.

8.4 Illustrative Examples of Mining Location-Based Social Network Data

In this section, we present two examples to illustrate how to mine real-world LBSN data to improve location-based services. The first example investigates a user's social–historical ties in check-in behavior for location prediction, and the second example leverages the social network information on LBSNs to address the "cold-start" check-in problem.

8.4.1 Exploring Social–Historical Ties on Location-Based Social Networks

On location-based social networking sites, a user's check-in behavior can be analyzed as an integration of his social ties and historical ties, while both ties have varying tie strengths, as illustrated in Fig. 8.2 with the tie strengths represented by line width (Gao et al. 2010a).

1. *Discovering the Properties of Social–Historical Ties*
   The historical ties of a user's check-in behavior have two properties in LBSNs. Firstly, a user's check-in history approximately follows a power-law distribution, i.e., a user goes to a few places many times and to many places a few times. Figure 8.3a shows the distribution of check-in frequency (in log scale) on a real-world dataset[10] collected from Foursquare, with detailed dataset statistics shown in Table 8.1. The figure suggests that the check-in history follows a power-law distribution, and the corresponding exponent is approximately 1.42. The check-in distribution of an individual also shows the power-law property, as shown in Fig. 8.3b. Secondly, historical ties have a short-term effect. As illustrated in Fig. 8.2, a user arrives at the airport and then takes a shuttle to the hotel. After his dinner, he sips a cup of coffee. The historical ties of the previous check-ins at the airport, shuttle stop, hotel, and restaurant have different strengths with respect to the latest check-in at the coffee shop. Furthermore, historical tie strength decreases over time.

[10] The dataset used in this example is available at: http://www.public.asu.edu/~hgao16/dataset/SHTiesData.zip

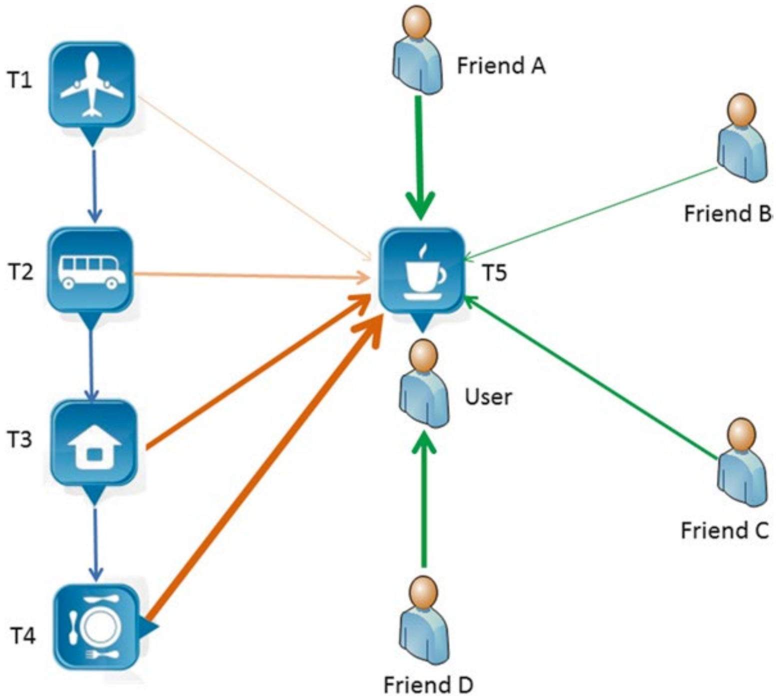


Fig. 8.2 An example: how social and historical ties may affect a user's check-ins at time T_5

To discover the properties of social ties, we compare the check-in similarity between users with friendship and those without. For each user, let $\mathbf{f} \in R^m$ be his check-in vector with the k-th element $\mathbf{f}(k)$ being the number of check-ins at location $l_k \in \pounds$, where $m = |\pounds|$ is the vocabulary size. The cosine similarity of two users u_i and u_j is defined as:

$$sim\left(u_i, u_j\right) = \frac{\mathbf{f}_i \mathbf{f}_j}{\left|\mathbf{f}_i\right|_2 \times \left|\mathbf{f}_j\right|_2}, \tag{8.2}$$

where $|\bullet|_2$ is the 2-norm of a vector.

We define the check-in similarity between u_i and a group G of other users as the average similarity between user u_i and the users in group G,

$$S_G\left(u_i\right) = \frac{\sum_{u_j \in G} sim\left(u_i, u_j\right)}{|G|}. \tag{8.3}$$

For each u_i, we calculate two similarities; i.e., $\mathbf{S_F}(u_i)$ is the average similarity of u_i and his friendship network; $\mathbf{S_R}(u_i)$ is the average similarity of u_i and a group of randomly chosen users, who are not in the friendship network of u_i. The number of randomly chosen users is the same as the amount of u_i's friends.

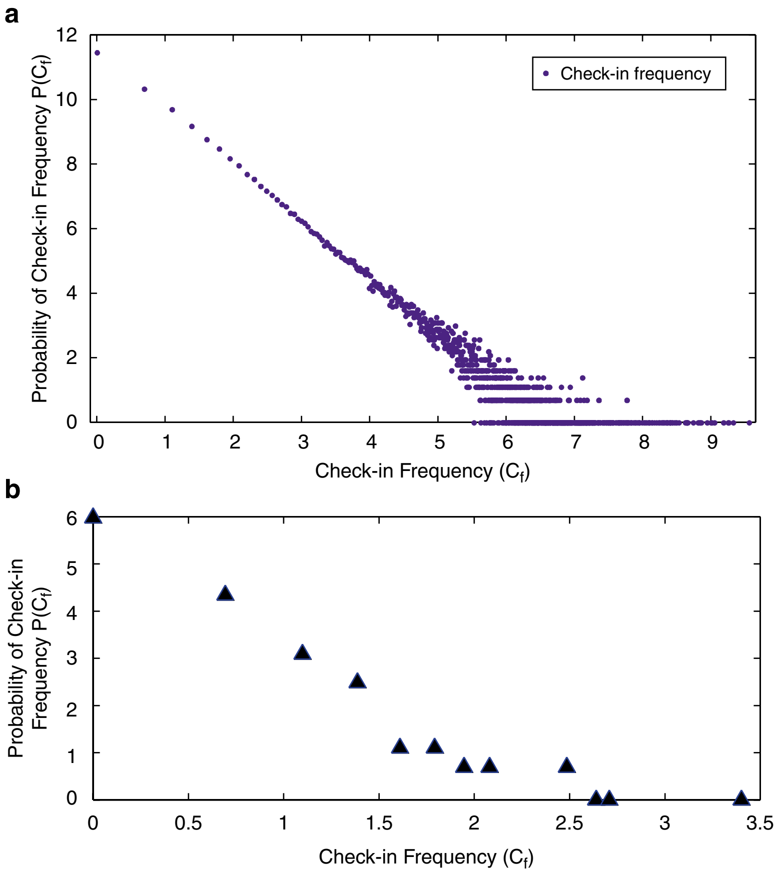


Fig. 8.3 The power-law distribution of check-ins. (**a**) Power-law distribution of check-ins in whole dataset. (**b**) Power-law distribution of check-ins in whole dataset

Table 8.1 Statistical information of Foursquare dataset

Duration	Mar. 8, 2010–Jan. 21, 2011
Number of users	18,107
Number of check-ins	2,073,740
Number of unique locations	43,063
Number of links	115,574

We conduct a two-sample *t*-test on the vectors $\mathbf{S_F}$ and $\mathbf{S_R}$. The null hypothesis is H_0: $\mathbf{S_F} \leq \mathbf{S_R}$, i.e., users with friendship share fewer common check-ins than those without, and the alternative hypothesis is H_1: $\mathbf{S_F} > \mathbf{S_R}$. In our experiment, the null hypothesis is rejected at significant level $\alpha = 0.001$ with *p*-value of 2.6*e*-6, i.e., users with friendship have higher check-in similarity than those without.

Table 8.2 Corresponding features between language and LBSN modeling

Language modeling		LBSN modeling	
Corpus		Check-in collection	
Document		Individual check-ins	
Document structure	Paragraph	Check-in structure	Monthly check-in sequence
	Sentence		Weekly check-in sequence
	Phrase		Daily check-in sequence
	Word		Check-in location

2. *Modeling Social–Historical Ties for Location Prediction*
To capture the two properties of historical ties, i.e., power-law distribution and short-term effect, a language model is utilized to model the check-in behavior. There are many features shared between language processing and LBSN mining. First, the text data and check-in data have similar structures, as shown in Table 8.2. For example, a document in language processing can correspond to an individual check-in sequence in LBSNs, while a word in the sentence corresponds to a check-in location. Second, the power-law distribution and short-term effect observed in LBSNs have also been found in natural language processing, where the word distribution is closely approximated by power-law (Zipf 1932), and the current word is more relevant to its adjacent words than distant ones. Therefore, to model the historical ties of a user, we introduce the hierarchical Pitman–Yor (HPY) language model (Teh 2006a, b) to the location-based social networks, which is a state-of-the-art language model that generates a power-law distribution of word tokens (Goldwater et al. 2006) while considering the short-term effect. We define the historical model (HM) as below,

$$P_H^i\left(c_t = l\right) = P_{HPY}^i\left(c_t = l \mid \Omega_i, \Theta\right), \tag{8.4}$$

where $P_{HPY}{}^i(c_t = l|\Omega_i, \Theta)$ is the probability of user u_i's check-in c_t at location l generated by the HPY with u_i's observed check-in history Ω_i, and Θ is the parameter set for the HPY language model. More technical details can be found in Gao et al. (2012a).

To model the social ties of check-in behavior, we define the social model (SM) as below,

$$P_S^i\left(c_t = l\right) = \sum_{u_j \in F(u_i)} sim\left(u_i, u_j\right) P_{HPY}^i\left(c_t = l \mid \Omega_j, \Theta\right), \tag{8.5}$$

where F(u_i) is the set of u_i's friends. $P_{HPY}{}^i(c_t = l|\Omega_j, \Theta)$ is the probability of u_i's next check-in c_t at location l computed by HPY with u_j's check-in history Ω_j as training data. Note that only the check-ins before the prediction time are included in the training data.

Finally, a social–historical model (SHM) is proposed to explore a user's check-in behavior, integrating both historical and social effects,

$$P_{SH}^i\left(c_t = l\right) = \eta P_H^i\left(c_t = l\right) + \left(1 - \eta\right) P_S^i\left(c_t = l\right). \tag{8.6}$$

where η controls the weight from historical ties and social ties.

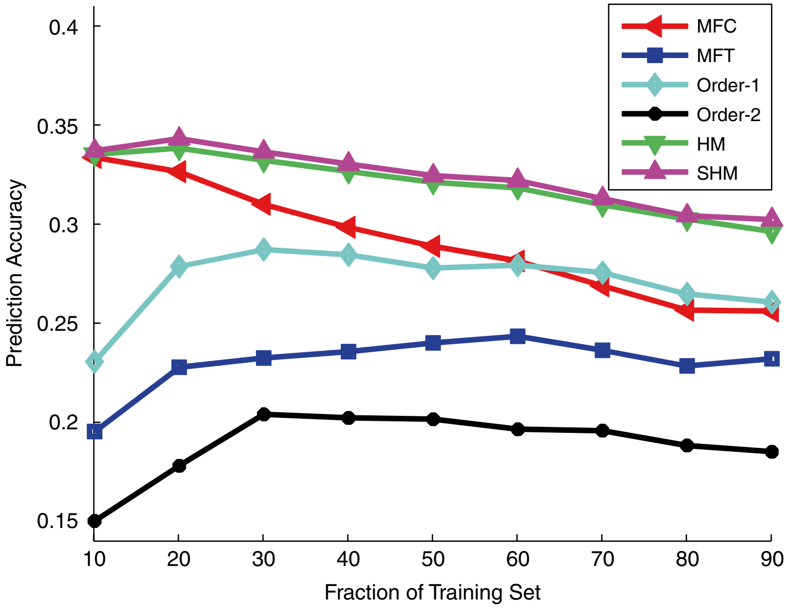


Fig. 8.4 The performance comparison of prediction models

The experimental results of location prediction on a real-world LBSN dataset are plotted in Fig. 8.4, with the performance comparison of the proposed model (HM and SHM) and four baseline models (Gao et al. 2012a). The results demonstrate that the proposed approach properly captures a user's check-in behavior by considering social–historical ties, and outperforms the current state-of-the-art prediction models.

8.4.2 gSCorr: Modeling Geo-Social Correlations for New Check-ins on Location-Based Social Networks

On location-based social networking sites, users explore various POIs and check in at places that interest them. The power-law property of users' check-in behavior in Fig. 8.3 indicates that users do visit new places, resulting in the "cold-start" check-in problem (Gao et al. 2012b). Predicting the "cold-start" check-in locations (i.e., predicting a user's next location where he has never been before) exacerbates the already difficult problem of location prediction, as there is no historical information on the user for the new place; hence, traditional prediction models relying on the observation of historical check-ins would fail to predict the "cold-start" check-ins. In this scenario, social network information could be utilized to help address the "cold-start" problem, since social theories (e.g., social correlation (Anagnostopoulos et al. 2008)) suggest that the movement of humans is usually affected by their social networks.

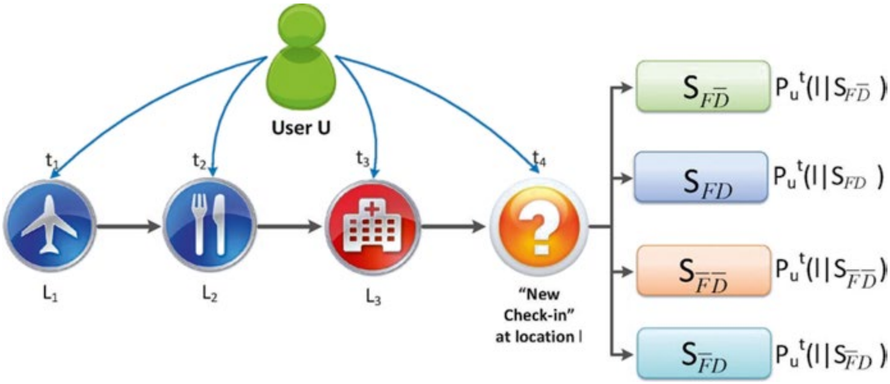


Fig. 8.5 Geo-social correlations of new check-in behavior

Table 8.3 Geo-social correlations

	F	$\bar{F}$
$\bar{D}$	$S_{F\bar{D}}$: local friends	$S_{\bar{F}\bar{D}}$: local non-friends
D	S_{FD}: distant friends	$S_{\bar{F}D}$: distant non-friends

Figure 8.5 illustrates a user's "new check-in" behavior in different social correlation aspects. User u goes to the airport at t_1, and then the restaurant at t_2 followed by the hospital at t_3. When u performs a "new check-in" at t_4, i.e., the check-in location does not belong to $\{L_1, L_2, L_3\}$, it may be correlated to those users that are from u's different geo-social circles $S_{F\bar{D}}$, S_{FD}, $S_{\bar{F}\bar{D}}$, and $S_{\bar{F}D}$, as defined in Table 8.3. Investigating these four circles enables us to study a user's check-in behavior in four corresponding aspects: local social correlation, distant social correlation, confounding, and unknown effect.

1. *Modeling Geo-Social Correlations*
   To model the geo-social correlations of "new check-in" behavior, we consider the probability of a user u checking-in at a new location l at time t as $P_u^t(l)$. We define this probability as a combination of the four geo-social correlations,

$$P_u^t(l) = \Phi_1 P_u^t(l|S_{\bar{F}\bar{D}}) + \Phi_1 P_u^t(l|S_{F\bar{D}})$$

$$+\Phi_3 P_u^t(l|S_{FD}) + \Phi_4 P_u^t(l|S_{\bar{F}D}) \tag{8.7}$$

   where Φ_1, Φ_2, Φ_3, and Φ_4 are four distributions that govern the strength of different geo-social correlations, $P_u^t(l|S_x)$ indicates the probability of user u checking-in at location l that is correlated to u's geo-social circle S_x.

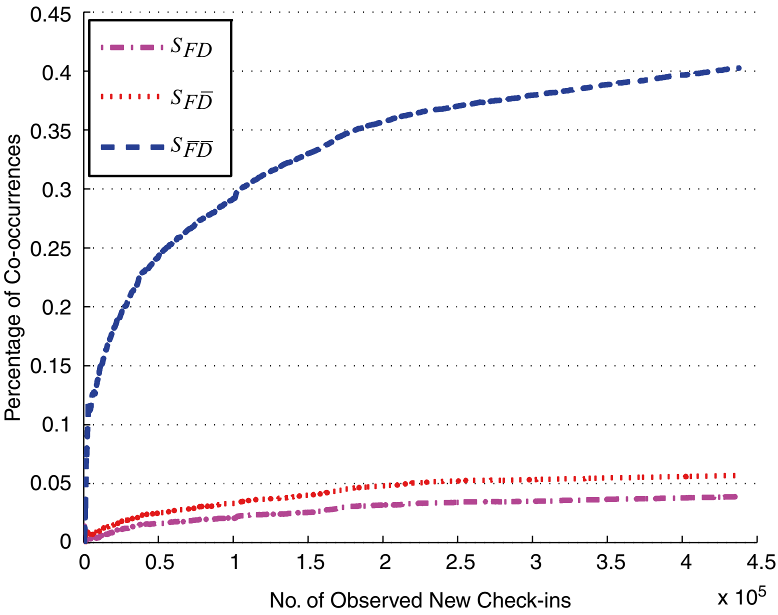


Fig. 8.6 Observed social correlations on new check-ins

Table 8.4 Statistical information of Foursquare dataset

Duration	Jan. 1, 2011–July 31, 2011
Number of users	11,326
Number of check-ins	1,385,223
Number of unique locations	182,968
Number of links	47,164

The modeling of Φ_1, Φ_2, Φ_3 and Φ_4 is based on the observation of "new check-in" distribution in Fig. 8.6, with the corresponding dataset[11] collected from Foursquare shown in Table 8.4. From Fig. 8.6, it is observed that Φ_1 is a real-valued and differentiable increasing function, and Φ_2 and Φ_3 are fairly constant. The percentage of "new check-ins" from $S_{\bar{F}D}$ is not presented, since it can be deduced from the other three. Therefore,

$$\begin{aligned}
\Phi_1 &= f\left(\mathbf{w}^T \mathbf{f}_u^t + b\right) \\
\Phi_2 &= \left(1-\Phi_1\right)\varnothing_1 \\
\Phi_3 &= \left(1-\Phi_1\right)\left(1-\varnothing_1\right)\varnothing_2 \\
\Phi_4 &= \left(1-\Phi_1\right)\left(1-\varnothing_1\right)\left(1-\varnothing_2\right) \\
0 \le \Phi_1 &\le 1, 0 \le \varnothing_1 \le 1, 0 \le \Phi_1 \le 1
\end{aligned} \tag{8.8}$$

[11] The dataset used in this example is available at: http://www.public.asu.edu/~hgao16/dataset/gScorrData.zip

Table 8.5 Check-in and social features

Features	Description
N^{c}	Number of check-ins in *u*'s history
N^{nc}	Number of new check-ins in *u*'s history
$N_{F\bar{D}}$	Number of friends in $S_{F\bar{D}}$
$N^{c}_{F\bar{D}}$	Number of check-ins from $S_{F\bar{D}}$
$N^{uc}_{F\bar{D}}$	Number of unique check-ins from $S_{F\bar{D}}$
$N^{vc}_{F\bar{D}}$	Number of visited check-ins from $S_{F\bar{D}}$
$N^{nvc}_{F\bar{D}}$	Number of visited unique check-ins from $S_{F\bar{D}}$
N_{FD}	Number of friends in S_{FD}
${N_{FD}}^{c}$	Number of check-ins from S_{FD}
${N_{FD}}^{uc}$	Number of unique check-ins from S_{FD}
${N_{FD}}^{vc}$	Number of visited check-ins from S_{FD}
${N_{FD}}^{nvc}$	Number of visited unique check-ins from S_{FD}
$N_{\bar{F}\bar{D}}$	Number of users in $S_{\bar{F}\bar{D}}$
$N^{c}_{\bar{F}\bar{D}}$	Number of check-ins from $S_{\bar{F}\bar{D}}$
$N^{uc}_{\bar{F}\bar{D}}$	Number of unique check-ins from $S_{\bar{F}\bar{D}}$
$N^{vc}_{\bar{F}\bar{D}}$	Number of visited check-ins from $S_{\bar{F}\bar{D}}$
$N^{nvc}_{\bar{F}\bar{D}}$	Number of visited unique check-ins from $S_{\bar{F}\bar{D}}$

where $\mathbf{f}_u^t$ is a check-in feature vector of a single user u at time t, $\mathbf{w}$ is a vector of the weights of $\mathbf{f}_u^t$, and b controls the bias. In this work, we define a user's check-in and social features $\mathbf{f}_u^t$ in Table 8.5. Φ_1 and Φ_2 are two constants.

To capture the geo-social correlation probabilities $P_u^t(l|S_x)$, three geo-social correlation measures are proposed considering the factors of location frequency, user frequency and user similarity, as described below,

- *Sim-Location Frequency (S.Lf)*

$$P_u^t\left(l \mid S_x\right) = \frac{\sum_{v \in S_x} s\left(u,v\right) N_v^t\left(l\right)}{\sum_{v \in S_x} s\left(u,v\right) N_v^t} \tag{8.9}$$

where $s(u,v)$ represents the user similarity between user u and user v. $N_v^t(l)$ represents the number of check-ins at location l by user v before time t, and N_v^t the total number of locations visited by user v that user u has not visited before time t .

- *Sim-User Frequency (S.Uf)*

$$P_u^t\left(l \mid S_x\right) = \frac{\sum_{v \in S_x} \delta_v^t\left(l\right) s\left(u,v\right)}{\sum_{v \in S_x} s\left(u,v\right)} \tag{8.10}$$

where $\delta_v^t(l)$ equals to 1 if user v has checked in at l before t, and 0 otherwise.

- *Sim-Location Frequency & User Frequency (S.Lf.Uf)*

$$P_u^t\left(l \mid S_x\right) = \frac{\sum_{v \in S_x} s\left(u,v\right) N_v^t\left(l\right)}{\sum_{v \in S_x} s\left(u,v\right) N_v^t} \frac{\sum_{v \in S_x} \delta_v^t\left(l\right)}{N_{S_x}} \tag{8.11}$$

Table 8.6 Evaluation metrics

	Single measure	Various measures
Equal strength	EsSm	EsVm
Random strength	RsSm	RsVm
Various strength	VsSm	gSCorr

We adopt *S.Lf.Uf*, *S.Lf*, and *S.Uf* to compute $P_u^t\left(l \mid S_{F\bar{D}}\right)$, $P_u^t(l|S_{FD})$ and $P_u^t\left(l \mid S_{\bar{F}\bar{D}}\right)$ respectively, based on our observation of their good performance on corresponding geo-social circles. To reduce time complexity, we consider $P_u^t\left(l \mid S_{\bar{F}D}\right)$ as a probability of random jump to a location in current location vocabulary that u has not checked in before.

2. *Evaluating gSCorr*
To evaluate gSCorr, we consider the effect of both geo-social correlation strength and measures in capturing the user's "new check-in" behavior. Therefore, we set up five baselines to compare the location prediction performance with gSCorr, as shown in Table 8.6. Each baseline adopts a different combination of correlation strength and measures, where "Es", "Rs", "Vs", "Sm", "Vm" represent "equal strength" (set all geo-social correlation strengths as 1), "random strength" (randomly assign the geo-social correlation strengths), "various strength" (the same as gScorr), "single measure" (use *S.Lf.Uf* to measure the correlation probabilities for all the geo-social circles) and "various measures" (the same as gScorr) respectively. Note that gSCorr is a various strength and various metrics approach. Following the evaluation metrics of recommendation system, we use top-k accuracy as evaluation metric and set k=*1*, *2*, *3* in the experiment. For each random strength approach (RsSm and RsVm), we run *30* times and report the average accuracy.

Table 8.7 shows the detailed prediction accuracy of each method for further comparison, with the best performance highlight as italics. We summarize the essential observations below:

- The geo-social correlations from different geo-social circles contribute variously to a user's check-in behavior. Both *VsSm* and *gSCorr* perform better than their equal strength versions (i.e., *EsSm* and *EsVm*) respectively, indicating that the geo-social correlations are not equally weighted.
- The randomly assigned strength approaches (*RsSm* and *RsVm*) perform the worst compared to the other approaches, where the performance of *VsSm* has a 10.50 % relative improvement over *RsSm*, and *gSCorr* has a 26.11 % relative improvement over *RsVm*, indicating that social correlation strengths do affect check-in behavior.
- The single metric approaches (*EsSm*, *RsSm*, *VsSm*) always perform worse than the various metrics approaches (*EsVm*, *RsVm*, *gSCorr*), which suggests that for different social circles, there are different suitable correlation metrics.

gSCorr performs the best among all the approaches. To demonstrate the significance of its improvement over other baseline methods, we launch a random guess approach to predict the "new check-ins". The prediction accuracy of the random guess is always below 0.005 % for top-1 prediction, and below 0.01 % for top-2 and

Table 8.7 Location prediction with various geo-social correlation strengths and measures

Methods	Top-1(%)	Top-2(%)	Top-3(%)
EsVm	17.88	24.06	27.86
EsSm	16.20	21.92	25.43
VsSm	16.49	22.28	25.92
RsSm	14.93	20.30	23.70
RsVm	15.23	20.85	24.50
gSCorr	*19.21*	*25.19*	*28.69*

top-3 prediction, indicating that gSCorr significantly improves the baseline methods, suggesting the advantage of gSCorr as considering different geo-social correlation strength and metrics for each geo-social circle.

8.5 Conclusions and Future Work

Location-based social networks carry user-driven geographical information, and bridge the gap between real world and online social media. Typical location-based social networking sites contain a triple-layer data structure including geographical, social, and content information, providing an unprecedented opportunity for studying mobile user behavior from a spatial, temporal, and social standpoint. In this chapter, we discuss the distinct properties of location-based social network data and their challenges, and elaborate current work for data analysis and research issues on location-based social networks.

This chapter has only discussed some essential issues. There are a number of interesting directions for further exploration.

- How do we better utilize social network information on LBSNs?
  Current work (Gao et al. 2012a; Cho et al. 2011; Ye et al. 2011c) on LBSNs reports very limited contributions from social networks. In their approaches for location prediction and recommender systems, models with social network information perform slightly better than those without social information. This leads to the question "is social network information really useful in explaining human mobile behavior?". The answer is probably still "yes", but the consequent problem is how to appropriately and efficiently make use of social information in LBSNs. For example, social information could be helpful on certain specific problems, such as the "cold-start" problem (Huang et al. 2004).
- How do we handle the check-in sparseness of LBSNs?
  The sparseness of user-driven check-ins in geographical sequence in LBSNs presents challenges to application of traditional approaches that cannot handle data sparseness. For example, in Cho et al. (2011), the authors evaluate their location prediction approaches on two location-based social network datasets and one cell phone dataset, reporting significantly higher accuracy on cell phone

data compared with LBSN data. The sparseness of LBSNs data can be one of the reasons that explain this phenomenon. Finding an efficient way to handle this sparse data is very challenging.

- How do we efficiently make use of user-generated content on LBSNs? User-generated content such as comments and tips for locations reflects the interest of the user within a spatial–temporal context. Current work mostly focuses on geographical patterns and social contexts; very few attempts have been made to make use of the user-generated content for understanding human behavior in LBSNs. Traditional text analysis approaches in social media could be leveraged for mining such content. For example, semantic knowledge that are used to enrich short texts (Hu et al. 2009, 2011) can be utilized to analyze the tips on LBSNs. Furthermore, an interesting research direction would consider the spatial–temporal, social, and content information together for improving location-based services. Investigating such information could help design new applications more closely to a user's daily life, and therefore improve the urban experience of citizen life.

Acknowledgments This work is supported, in part, by ONR (N000141010091). The authors would like to acknowledge all of the researchers in Arizona State University's Data Mining and Machine Learning Laboratory. The views expressed in this chapter are solely attributed to the authors, and do not represent the opinions or policies of any of the funding agencies.

References

ABI Research. (2008). *Location-based mobile social networking: Hype or reality?* http://www.abiresearch.com/research/product/1002345-location-based-mobile-social-networking-hy/ Accessed 17 Nov 2012.

ABI Research. (2010). *Location-based marketing*. http://www.abiresearch.com/research/product/1005770-location-based-marketing/ Accessed 17 Nov 2012.

Anagnostopoulos, A., Kumar, R., & Mahdian, M. (2008). Influence and correlation in social networks. Las Vegas: In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 7–15).

Backstrom, L., Sun, E., & Marlow, C. (2010). Find me if you can: Improving geographical prediction with social and spatial proximity. North Carolina: In *Proceedings of the 19th International Conference on World Wide Web* (pp. 61–70).

Barnes, S., & Scornavacca, E. (2004). Mobile marketing: The role of permission and acceptance. *International Journal of Mobile Communications, 2*(2), 128–139.

Barwise, P., & Strong, C. (2002). Permission-based mobile advertising. *Journal of Interactive Marketing, 16*(1), 14–24.

Bauer, H., Barnes, S., Reichardt, T., & Neumann, M. (2005). Driving consumer acceptance of mobile marketing: A theoretical framework and empirical study. *Journal of Electronic Commerce Research, 6*(3), 181–192.

Ben-Akiva, M., Bierlaire, M., Koutsopoulos, H., & Mishalani, R. (1998). Dynamit: A simulation-based system for traffic prediction. In *DACCORS Short Term Forecasting Workshop*. The Netherlands: Citeseer.

Berjani, B., & Strufe, T. (2011). A recommendation system for spots in location-based online social networks. Salzburg: In *Proceedings of the 4th Workshop on Social Network Systems* (pp. 1–6).

Brown, C., Nicosia, V., Scellato, S., Noulas, A., & Mascolo, C. (2012). Where online friends meet: Social communities in location-based networks. Dublin: In *Sixth International AAAI Conference on Weblogs and Social Media*.

Cairncross, F. (2001). *The death of distance: How the communications revolution is changing our lives*. Boston: Harvard Business Press.

Chang, J., & Sun, E. (2011). Location 3: How users share and respond to location-based data on social networking sites. Barcelona: In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media.*

Chen, G., & Kotz, D. (2000). *A survey of context-aware mobile computing research* (Technical Report No. TR2000-381). Hanover: Department of Computer Science, Dartmouth College.

Cheng, Z., Caverlee, J., & Lee, K. (2010). You are where you tweet: A content-based approach to geolocating Twitter users. Toronto: In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management* (pp. 759–768).

Cheng, Z., Caverlee, J., Lee, K., & Sui, D. (2011). Exploring millions of footprints in location sharing services. Barcelona: In *Proceedings of the Fifth International Conference on Weblogs and Social Media.*

Cho, E., Myers, S., & Leskovec, J. (2011). Friendship and mobility: User movement in location-based social networks. San Diego: In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1082–1090).

Consolvo, S., Smith, I., Matthews, T., LaMarca, A., Tabert, J., & Powledge, P. (2005). Location disclosure to social relations: Why, when, & what people want to share. Portland: In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 81–90).

Cranshaw, J., Toch, E., Hong, J., Kittur, A., & Sadeh, N. (2010). Bridging the gap between physical location and online social networks. Copenhagen: In *Proceedings of the 12th ACM International Conference on Ubiquitous Computing* (pp. 119–128).

Cranshaw, J., Schwartz, R., Hong, J., & Sadeh, N. (2012). The Livehoods project: Utilizing social media to understand the dynamics of a city. Dublin: In *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media: Vol. 12.*

De Longueville, B., Smith, R., & Luraschi, G. (2009). OMG, from here, I can see the flames!: A use case of mining location based social networks to acquire spatio-temporal data on forest fires. In *Proceedings of the 2009 International Workshop on Location Based Social Networks* (pp. 73–80).

Dia, H. (2001). An object-oriented neural network approach to short-term traffic forecasting. *European Journal of Operational Research, 131*(2), 253–261.

Eagle, N., Pentland, A., & Lazer, D. (2009). Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences, 106*(36), 15274–15278.

Gao, H., Barbier, G., & Goolsby, R. (2011a). Harnessing the crowdsourcing power of social media for disaster relief. *IEEE Intelligent Systems, 26*(3), 10–14.

Gao, H., Wang, X., Barbier, G., & Liu, H. (2011b). Promoting coordination for disaster relief – from crowdsourcing to coordination. *Social Computing, Behavioral-Cultural Modeling and Prediction* (pp. 197–204).

Gao, H., Tang, J., & Liu, H. (2012a). Exploring social–historical ties on location-based social networks. Dublin: In *Proceedings of the Sixth International Conference on Weblogs and Social Media.*

Gao, H., Tang, J., & Liu, H. (2012b). gSCorr: Modeling geo-social correlations for new check-ins on location-based social networks. Hawaii: In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management.*

Gao, H., Tang, J., & Liu, H. (2012c). Mobile location prediction in spatio-temporal context. In *Proceedings of the Mobile Data Challenge by Nokia Workshop in conjunction with International Conference on Pervasive Computing*. Newcastle.

Gastner, M., & Newman, M. (2006). The spatial structure of networks. *The European Physical Journal B-Condensed Matter and Complex Systems, 49*(2), 247–252.

Goldenberg, J., & Levy, M. (2009). *Distance is not dead: Social interaction and geographical distance in the internet era*. Arxiv preprint arXiv:0906.3202.

Goldwater, S., Griffiths, T., & Johnson, M. (2006). Interpolating between types and tokens by estimating power-law generators. *Advances in Neural Information Processing Systems, 18*, 459.

Goodchild, M., & Glennon, J. (2010). Crowdsourcing geographic information for disaster response: A research frontier. *International Journal of Digital Earth, 3*(3), 231–241.

Gundecha, P., Barbier, G., & Liu, H. (2011). Exploiting vulnerability to secure user privacy on a social networking site. San Diego: In *Proceedings of the 17th ACM SIGKDD Conference* (pp. 511–519).

Hecht, B., Hong, L., Suh, B., & Chi, E. (2011). Tweets from Justin Bieber's heart: The dynamics of the location field in user profiles. Vancouver: In *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems* (pp. 237–246).

Hong, L., Ahmed, A., Gurumurthy, S., Smola, A., & Tsioutsiouliklis, K. (2012). Discovering geographical topics in the Twitter stream. Beijing: In *Proceeding of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.*

Hu, X., & Liu, H. (2012). Text analytics in social media. In C. C. Aggawal & C. Zhai (Eds.), *Mining text data* (pp. 385–414). New York: Springer.

Hu, X., Sun, N., Zhang, C., & Chua, T. (2009). Exploiting internal and external semantics for the clustering of short texts using world knowledge. Hong Kong: In *Proceeding of the 18th ACM Conference on Information and Knowledge Management* (pp. 919–928).

Hu, X., Tang, L., & Liu, H. (2011). Enhancing accessibility of microblogging messages using semantic knowledge. Glasgow: In *Proceedings of the 20th ACM international conference on Information and knowledge management* (pp. 2465–2468).

Huang, Z., Chen, H., & Zeng, D. (2004). Applying associative retrieval techniques to alleviate the sparsity problem in collaborative filtering. *ACM Transactions on Information Systems (TOIS), 22*(1), 116–142.

Humphreys, L. (2007). Mobile social networks and social practice: A case study of Dodgeball. *Journal of Computer-Mediated Communication, 13*(1), 341–360.

Kelley, P., Hankes Drielsma, P., Sadeh, N., & Cranor, L. (2008). User-controllable learning of security and privacy policies. Alexandria: In *Proceedings of the 1st ACM Workshop on Workshop on AISec* (pp. 11–18).

Kessler, S. (2012). *Foursquare tops 20 million users*. http://mashable.com/2012/04/16/foursquare-20-million/. Accessed 16 Nov 2012.

Kulshrestha, J., Kooti, F., Nikravesh, A., & Gummadi, K. (2012). Geographic dissection of the Twitter network. Dublin: In *AAAI International Conference on Weblogs and Social Media.*

Lederer, S., Mankoff, J., & Dey, A. (2003). Who wants to know what when? Privacy preference determinants in ubiquitous computing. Florida: In *CHI'03 Extended Abstracts on Human Factors in Computing Systems* (pp. 724–725).

Li, N., & Chen, G. (2009). Analysis of a location-based social network. Vancouver: In *International Conference on Computational Science and Engineering*: *Vol. 4*. (pp. 263–270).

Li, Q., Zheng, Y., Xie, X., Chen, Y., Liu, W., & Ma, W. (2008). Mining user similarity based on location history. Irvine: In *Proceedings of the 16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems: Vol. 34*.

Liben-Nowell, D., Novak, J., Kumar, R., Raghavan, P., & Tomkins, A. (2005). Geographic routing in social networks. *Proceedings of the National Academy of Sciences, 102*(33), 11623–11628.

Lindqvist, J., Cranshaw, J., Wiese, J., Hong, J., & Zimmerman, J. (2011). I'm the mayor of my house: Examining why people use Foursquare – a social-driven location sharing application. In *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems* (pp. 2409–2418).

Mok, D., Wellman, B., & Carrasco, J. (2010). Does distance matter in the age of the Internet? *Urban Studies, 47*(13), 2747.

Monreale, A., Pinelli, F., Trasarti, R., & Giannotti, F. (2009). Wherenext: A location predictor on trajectory pattern mining. Paris: In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 637–646).

Noulas, A., Scellato, S., Mascolo, C., & Pontil, M. (2011a). An empirical study of geographic user activity patterns in Foursquare. Barcelona: In *Proceedings of the 5th International AAAI Conference on Weblogs and Social Media*

Noulas, A., Scellato, S., Mascolo, C., & Pontil, M. (2011b). Exploiting semantic annotations for clustering geographic areas and users in location-based social networks. Barcelona: In *Proceedings of SMW11.*

Noulas, A., Scellato, S., Lathia, N., & Mascolo, C. (2012). Mining user mobility features for next place prediction in location-based services. In *Proceedings of the 12th IEEE international conference on data mining* (pp. 1038–1043). Brussels.

Pozdnoukhov, A., & Kaiser, C. (2011). Space-time dynamics of topics in streaming text. Chicago: In *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Location-Based Social Networks*: *Vol. 8*.

Rhee, I., Shin, M., Hong, S., Lee, K., Kim, S., & Chong, S. (2011). On the Levy-walk nature of human mobility. *IEEE/ACM Transactions on Networking (TON), 19*(3), 630–643.

Sadeh, N., Hong, J., Cranor, L., Fette, I., Kelley, P., Prabaker, M., & Rao, J. (2009). Understanding and capturing people's privacy policies in a mobile social networking application. *Personal and Ubiquitous Computing, 13*(6), 401–412.

Sadilek, A., Kautz, H., & Bigham, J. (2012a). Finding your friends and following them to where you are. Seattle: In *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining* (pp. 723–732).

Sadilek, A., Kautz, H., & Silenzio, V. (2012b). Dublin: Modeling spread of disease from social interactions. In *Proceedings of Sixth AAAI International Conference on Weblogs and Social Media (ICWSM)*.

Sakaki, T., Okazaki, M., & Matsuo, Y. (2010). Earthquake shakes Twitter users: Real-time event detection by social sensors. North Carolina: In *Proceedings of the 19th international conference on World wide web* (pp. 851–860).

Scellato, S., Mascolo, C., Musolesi, M., & Latora, V. (2010). Distance matters: Geo-social metrics for online social networks. Berkeley: In *Proceedings of the 3rd Conference on Online Social Networks* (pp. 8–8). USENIX Association.

Scellato, S., Musolesi, M., Mascolo, C., Latora, V., & Campbell, A. (2011a). Nextplace: A spatio-temporal prediction framework for pervasive systems. San Francisco: *Pervasive Computing*, 152–169.

Scellato, S., Noulas, A., Lambiotte, R., & Mascolo, C. (2011b). Socio-spatial properties of online location-based social networks. Barcelona: In *Proceedings of the 5th International AAAI Conference on Weblogs and Social Media*, 329–336.

Scellato, S., Noulas, A., & Mascolo, C. (2011c). Exploiting place features in link prediction on location-based social networks. San Diego: In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1046–1054).

Scharl, A., Dickinger, A., & Murphy, J. (2005). Diffusion and success factors of mobile marketing. *Electronic Commerce Research and Applications, 4*(2), 159–173.

Spaccapietra, S., Parent, C., Damiani, M., De Macedo, J., Porto, F., & Vangenot, C. (2008). A conceptual view on trajectories. *Data and Knowledge Engineering, 65*(1), 126–146.

Tang, J., Gao, H., & Liu, H. (2012a). mTrust: Discerning multi-faceted trust in a connected world. Seattle: In *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining* (pp. 93–102).

Tang, J., Gao, H., Liu, H., & Sarma, A. (2012b). eTrust: Understanding trust evolution in an online world. Beijing: In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 253–261).

Teh, Y. (2006a). A Bayesian interpretation of interpolated Kneser–Ney (Technical Report TRA2/06). Singapore: National University of Singapore.

Teh, Y. (2006b). A hierarchical Bayesian language model based on Pitman–Yor processes. Sydney: In *ACL* (pp. 985–992). Association for Computational Linguistics.

Thanh, N., & Phuong, T. (2007). A Gaussian mixture model for mobile location prediction. Hanoi: In *2007 IEEE International Conference on Research, Innovation and Vision for the Future* (pp. 152–157).

Toch, E., Cranshaw, J., Drielsma, P., Tsai, J., Kelley, P., Springfield, J., Cranor, L., Hong, J., & Sadeh, N. (2010a). Empirical models of privacy in location sharing. Copenhagen: In *Proceedings of the 12th ACM International Conference on Ubiquitous Computing* (pp. 129–138).

Toch, E., Cranshaw, J., Hankes-Drielsma, P., Springfield, J., Kelley, P., Cranor, L., Hong, J., Sadeh, N. (2010b). Locaccino: A privacy-centric location-sharing application. Copenhagen: In *Proceedings of the 12th ACM International Conference Adjunct Papers on Ubiquitous Computing* (pp. 381–382).

Tsai, J., Kelley, P., Drielsma, P., Cranor, L., Hong, J., & Sadeh, N. (2009). Who's viewed you? The impact of feedback in a mobile location-sharing application. Boston: In *Proceedings of the 27th International Conference on Human Factors in Computing Systems* (pp. 2003–2012).

Vasconcelos, M., Ricci, S., Almeida, J., Benevenuto, F., & Almeida, V., (2012). Seattle: Tips, dones and to-dos: Uncovering user profiles in Foursquare. In *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining* (pp. 653–662).

Wang, F., & Huang, Q. (2010). The importance of spatial–temporal issues for case-based reasoning in disaster management. Beijing: In *2010 18th International Conference on Geoinformatics* (pp. 1–5). IEEE.

Wang, D., Pedreschi, D., Song, C., Giannotti, F., & Barab'asi, A. (2011). Human mobility, social ties, and link prediction. San Diego: In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1100–1108).

Ye, M., Yin, P., & Lee, W. (2010). Location recommendation for location-based social networks. San Jose: In *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems* (pp. 458–461).

Ye, M., Janowicz, K., M¨ulligann, C., & Lee, W. (2011a). What you are is when you are: The temporal dimension of feature types in location-based social networks. Chicago: In *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* (pp. 102–111).

Ye, M., Shou, D., Lee, W., Yin, P., & Janowicz, K. (2011b) On the semantic annotation of places in location-based social networks. San Diego: In Proceedings of the 17th ACM SIGKDD *International Conference on Knowledge Discovery and Data Mining* (pp. 520–528).

Ye, M., Yin, P., Lee, W., & Lee, D. (2011c). Exploiting geographical influence for collaborative point-of-interest recommendation. Beijing: In *Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 325–334).

Yin, Z., Cao, L., Han, J., Zhai, C., & Huang, T. (2011). Geographical topic discovery and comparison. Hyderabad: In *Proceedings of the 20th International Conference on World Wide Web* (pp. 247–256).

Zheng, Y., Zhang, L., Xie, X., & Ma, W. (2009). Mining interesting locations and travel sequences from GPS trajectories. Madrid: In *WWW* (pp. 791–800).

Zhou, D., Wang, B., Rahimi, S., & Wang, X. (2012). A study of recommending locations on location-based social network by collaborative filtering. *Advances in Artificial Intelligence*, 255–266.

Zickuhr, K. (2012). Three-quarters of smartphone owners use location-based services. *Pew Internet & American Life Project*.

Zipf, G. (1932). *Selective studies and the principle of relative frequency in language*. Cambridge: Harvard University Press.

Chapter 9
Towards Trustworthy Mobile Social Networking

Zheng Yan, Valtteri Niemi, Yu Chen, Peng Zhang, and Raimo Kantola

Abstract Self-organized networks based on mobile devices, e.g., mobile ad hoc networks (MANET), are becoming a practical platform for mobile social networking. People, either familiar or strangers, communicate with each other via such a network for instant social activities. How to help mobile users to build up trust in mobile social networking is becoming an important and interesting issue. Trust concerns not only security, but also privacy, as well as quality of social networking experiences. It relates to many properties that are essential for establishing a trust relationship in ephemeral and dynamically changed mobile social environments. This book chapter reviews the literature with regard to how to build up trust in mobile social networking. We explore whether mobile social networking is demanded, considering many existing popular Internet social networking services.

Z. Yan (✉)
The State Key Laboratory of ISN, Xidian University, Xi'an, China

Department of ComNet, Aalto University, Espoo, Finland
e-mail: zyan@xidian.edu.cn; zheng.yan@aalto.fi

V. Niemi
Department of Mathematics and Statistics, University of Turku, Turku, Finland

Y. Chen
Human Computer Interaction Group, EPFL, Lausanne, Switzerland
e-mail: yu.chen@epfl.ch

P. Zhang
Xi'an University of Post and Telecommunications, Xi'an, China
e-mail: pengzhangzhang@gmail.com; pzhang@xupt.edu.cn

R. Kantola
Department of ComNet, Aalto University, Espoo, Finland
e-mail: raimo.kantola@aalto.fi

A. Chin and D. Zhang (eds.), *Mobile Social Networking: An Innovative Approach*,
Computational Social Sciences, DOI 10.1007/978-1-4614-8579-7_9,

Based on a need assessment survey, we propose a trust management framework that supports context-aware trust/reputation generation, trustworthy content recommendations, secure communications, unwanted traffic control, user privacy recommendation and preservation, and other trust and privacy enhancement technologies. Simulations, prototype implementation, and trial experiments further prove the effectiveness of proposed solutions.

9.1 Introduction

With the rapid growth of mobile computing and social networking technologies, the social network has extended its popularity from Internet to the mobile domain. Personal mobile devices (e.g., smart phones) can be self-organized and communicate with each other for social activities by forming a multi-hop radio network and maintaining connectivity in a decentralized manner. We refer to such a kind of social networking based on mobile devices that supports instant and pervasive social activities as mobile social networking (MSN). Nowadays, the mobile ad hoc network (MANET) has become a practical platform for mobile social networking and computing, playing as a valuable extension and complement of traditional Internet social networks. For example, a user could query people in the vicinity using his/her mobile device about which shop is on sale, which movie is recommended to see, or which mobile application should be installed for tagging the locations of photos. The user neighbors can respond to these queries by providing their recommendations via MSN. The users can also chat with people nearby via MSN with regard to sharing a taxi ride, or sharing the cost of a series of movie tickets. Moreover, they can seek services or aids from nearby strangers through MSN. This kind of social networking is very valuable for mobile users, especially when fixed networks (e.g., Internet) or mobile networks are temporarily unavailable or costly to access.

Several research groups in academia have focused on social activities based on mobile ad hoc networks. The Stanford MobiSocial Group has developed Junction, a mobile ad hoc and multiparty platform for MANET applications (Junction 2012). Micro-blog (2012), developed by SyNRG in Duke University, helps users to post micro-blogs tagged by locations. AdSocial (Stuedi et al. 2008), introduced by the ETHz Systems Group, provides a pervasive social communication platform. Floating content concept has been analyzed based on a theoretical framework to study the fundamental quantities of an ephemeral content-sharing service in opportunistic networking, such as node encounter rate, mean contact times as a function of location, and achievable transmission rates and transmission ranges (Hyytia et al. 2011; Ott et al. 2011). In a proposed floating content system, content is only shared within an anchor zone in a best-effort manner, i.e., copies are kept available within that zone while they are deleted outside the anchor zone.

In industry, quite a number of companies, such as Microsoft, Nokia, and Intel, have conducted research in the area of MSN. For example, Microsoft Research Asia developed the EZSetup system in order to enable a mobile user to find services provided by his/her neighbors (EZSetup 2012). The Nokia Instant Community (NIC) developed by Nokia Research Center provides an instant social networking platform to allow people in the vicinity to communicate, get to know, and share with each other (Nokia Instant Community 2010a; Nokia Instant Community 2010b; Ahtiainen et al. 2009). Similarly, the Intel Berkeley Lab ran a project named Familiar Stranger based on mobile devices, to extend our feelings and relationships with strangers that we regularly observe but do not interact with in public places (Paulos and Goodman 2012).

However, trust, security, and privacy aspects in mobile social networking have not been seriously considered in existing projects. Traditional centralized social networking systems (e.g., Facebook) have not taken user privacy into account. They cannot satisfy instant social networking demands, especially when users do not have Internet connection, but have location proximity with each other. Issues on trust management for security assurance and privacy enhancement need serious research, in order to deploy a practical mobile social networking system that can be easily accepted by mobile users. A number of unsolved crucial issues with regard to trust, security, and privacy should be overcome to move towards trustworthy mobile social networking.

First, key management for secure data access control in MSN should adapt to trust levels of users and social context changes. Due to the dynamic changes of MSN topology and user trust level, the encryption key used for securing social communications needs to be frequently changed, and the decryption key should be distributed to each of the eligible users. This introduces a heavy traffic and processing load, which may cause a serious performance bottleneck. How to automatically control MSN data access in a secure and efficient way is a challenge. Most existing work in MSN didn't provide a solution to control social communication data access and effectively support user revocation (Junction 2012; MicroBlog 2012; Stuedi et al. 2008; EZSetup 2012; Nokia Instant Community 2010a; Nokia Instant Community 2010b; Ahtiainen et al. 2009; Paulos and Goodman 2002). Past key management solutions didn't consider applying the trust level of user and context attributes as control conditions for MSN data access in order to adapt context changes. Thus, they are not effective in the practice of MSN (Bethencourt et al. 2007; Goyal et al. 2006; Muller et al. 2008; Sahai and Waters 2005; Wang et al. 2010).

Second, a practical privacy enhanced trust management system with context-awareness and easy user acceptance is demanded in MSN. During mobile social networking, how much should the users trust each other in order to make a decision? Moreover, users generally want to preserve personal privacy and avoid malicious tracking. Although there are a number of research activities concerning MSN in academia and industry, issues with regard to usable and autonomic trust management for adaptive security assurance and privacy enhancement still need serious research. Herein, we define trust as the confidence, belief, and expectation regarding the reliability, integrity, ability, or character of an entity. Reputation is a measure derived from direct or indirect knowledge/experience on earlier interactions of entities, and is used to assess the level of trust put into an entity.

Third, controlling unwanted traffic is another crucial issue in MSN. People communicate in different contexts for different purposes, creating various content information flows. Examples are a mobile application installation link, a URL of a service, a Web page, and a textual message typed by a user. However, at the time when mobile users expect useful and valuable contents via MSN, they may also receive unwanted or unexpected contents, for example, malware; a virus or a Trojan; spammed emails, www pages, spammed voice-VoIP, spammed instant messaging, SMS, Web contents; and denial of service (DoS) and distributed denial of service (DDoS) attacks. The unwanted contents could penetrate user devices, consume user time, occupy user device memory, and irritate the user. How much should a user trust different contents received over the MSN? What contents are unwanted by the users, and thus should be controlled? Controlling unwanted contents becomes a crucial issue in MSN. Herein, unwanted content is defined as the content that is not expected by its destination or the content consumer.

Finally, a context-aware recommender system for user privacy is expected in MSN. For accessing services in MSN, a mobile user may be requested to share personal information and data (e.g., user profile and location information) with other node service providers. Privacy becomes a crucial issue in MSN, because it is the ability of an entity to seclude itself or information about itself and thereby reveal itself selectively. However, it is normally difficult for a user to justify whether it is safe and proper to disclose personal data to others in different contexts. In addition, privacy is a subjective issue. Different users treat personal privacy differently, even in the same situation. To solve all the above problems, there is a demand to provide a context-aware recommender system for user privacy that could help the user make a decision on personal data sharing in MSN.

This book chapter explores whether mobile social networking is demanded, considering many existing popular Internet social networking services. Based on a need assessment survey, we propose a hybrid trust management framework – AwareTrust – which supports context-aware trust/reputation generation, trustworthy recommendations, secure communications, unwanted traffic control, and user privacy recommendation. Meanwhile, we also propose a number of other trust and privacy technologies that can be adopted in MSN in order to enhance its trustworthiness. We have carried out a series of simulations, implemented a number of prototype systems, and conducted user/trial experiments to prove the effectiveness of the proposed solutions.

The rest of the chapter is organized as follows. Section 9.2 reviews literature background and related work. We introduce the result of a need assessment survey to investigate how users consider and expect to cope with a reputation system for mobile social networking in Sect. 9.3. Next, we describe the design of a trust management framework – AwareTrust – for MSN in Sect. 9.4. Section 9.5 reports the main properties for trust that are supported by the framework. Simulation results, prototype implementation and the findings and implications of user/trial experiments are then presented in Sect. 9.6. In Sect. 9.7, we review a number of interesting issues related to trust, security, and privacy of MSN. Finally, we conclude by discussing the contributions of this chapter and suggesting future work towards trustworthy mobile social networking.

9.2 The State-of-the-Art

9.2.1 Trust and Reputation Mechanism

Trust and reputation mechanisms have been widely studied in various fields of distributed systems, such as ad hoc networks, peer-to-peer (P2P) systems, grid computing, pervasive computing, and e-commerce (Yan 2010). Many mechanisms have been developed for supporting trusted communications and collaborations among computing nodes (Sun et al. 2006a; Theodorakopoulos and Baras 2006; Yan and Holtmanns 2008). Examples are FuzzyTrust system (Song et al. 2005), the eBay user feedback system (Resnick and Zeckhauser 2002), PeerTrust model (Xiong and Liu 2004), an objective trust management framework (OTMF) for MANET (Li et al. 2008), and Credence – a robust and decentralized system for evaluating the reputation of files in a P2P system (Walsh and Sirer 2005). Some work evaluates trust based on social relationships (Trifunovic et al. 2010). In the above research, trust can be modeled, calculated, and thus expressed using a value. However, none of the above studies consider how to evaluate trust and reputation based on social networking behaviors and experiences, especially in the context of mobile social networking. None of them support user privacy. Some factors influencing trust in MSN were never considered in the previous work. Moreover, only a small amount of the work in the literature develops a reputation system driven by the concern of users (Yan and Niemi 2009). A reputation rating system based on past behavior of evaluators was proposed in Kujimura and Nishihara (2003). Trust in the evaluator indexes its impact on the rating system. The trust value is dynamically adjusted based on past estimation performance. In our research, the user reputation is adjusted based on his/her past social performance. It is evaluated by each individual node and a trusted server (TS) based on ephemeral and historical experiences respectively.

9.2.2 Trust and Reputation System Architecture

Jøsang et al. classified trust/reputation system architecture into two main types: centralized and distributed (Jøsang et al. 2007). The system architecture determines how ratings and reputation scores are communicated between participants in a reputation system. In the literature, distributed trust evaluations have been studied in MANET, but the solutions seldom support node privacy (Sun et al. 2006a; Theodorakopoulos and Baras 2006; Raya et al. 2008). This could cause potential attacks such as bad-mouthing attacks or unfair rating attacks targeting a specific node (Sun et al. 2008). Most existing systems maintain a statistical representation of reputation by borrowing tools from the realms of game theory (Confidant (Buchegger and Boudec 2002) and Core (Michiardi 2002)), and peer-to-peer networks (Xiong and Liu 2004) and Bayesian analytics (Buchegger and Boudec 2003). These systems try to counter selfish routing misbehavior of nodes by enforcing nodes to cooperate with each other and counter any arbitrary misbehavior of nodes. However, little attention has been paid to the content and social communication reputation issue in MANET with node privacy as a main concern. On the other hand, practical reputation systems generally apply a centralized server to collect feedback for reputation generation (e.g., eBay (Resnick and Zeckhauser 2002), Yahoo auctions (Resnick et al. 2000), and Internet-based systems such as Keynote (Blaze et al. 1996)). However, many existing systems (e.g., Amazon, eBay) lack consideration of the credibility of user vote. This greatly influences the quality of produced reputation. The usage of pseudonyms and the ease of changing them additionally complicates the picture by allowing participants to effectively erase their prior history. We plan to adopt a hybrid trust management system architecture in MSN to overcome the weakness of the above two kinds of architecture, where reputation is evaluated in a distributed way, but with the support of a centralized trusted server.

Recently, a number of reputation systems have been proposed in the context of digital contents and social networking. For example, Thomas Adler and Alfaro proposed a content-driven reputation system for Wikipedia authors solely on the basis of content evolution; but not on user-to-user comments or ratings (Adler and Alfaro 2007). The concept of data centric trust in volatile environments, such as ad hoc networks, was introduced in (Raya et al. 2008) to evaluate the node trust through the data reported by it. Gupta et al. proposed a partially distributed reputation system for P2P systems by introducing a reputation computation agent (RCA) (Gupta et al. 2003). But this system does not consider the challenges caused by privacy enhancement and context-awareness. In addition, the RCA is applied only for calculating peer's reputation based on its contribution to the system. This is not suitable for our research scenario. In our research, we plan to apply the trusted server to update the node user reputation on the basis of historical social behaviors.

9.2.3 Problems of Trust and Reputation Management

In most reputation systems in the ad hoc networks, the reputation of a node is shared globally in the network. The purpose is to make the reputation of a node known to all other nodes and decrease the detection time. Thus, maintaining and disseminating indirect reputation information incur overhead at both the individual node and the network. OCEAN (Bansal and Baker 2003) discounts second-hand reputation exchange, and only utilizes local reputation based on direct observations in order to achieve a reasonable performance. We consider both local-aware and global-aware reputations by aggregating local experiences and global experiences together. By deploying a trusted server, the overhead of reputation maintenance and dissemination is eliminated among MSN nodes.

Inconsistent reputation problem (i.e., different nodes may have different reputation values for the same node) often occurs in the ad hoc networks due to subjective reasons and/or different local experiences. This makes it hard to distinguish correct reputation ratings from reputation voting messages. LARS (locally aware reputation system) (Hu and Burmester 2006) was proposed to deal with selfish behaviors and malicious behaviors (e.g., packet dropping and unfair rating). In LARS, the reputation of a node is derived from direct observation, and exchange of second-hand reputation information is disallowed. We plan to apply a trusted server to unify node reputation based on local experiences reported by nodes. This reputation information is issued to the node by the server. Serving as the initial value of reputation, it is further evolved based on experiences newly collected at the individual node. In addition, the above process is iterated. Thereby, we avoid the inconsistent reputation problem and eliminate user reputation inaccuracy caused by multi-hop reputation dissemination. Reputation generation is based on first-hand experiences and direct votes, whether at the TS or the node.

Nowadays, reputation systems may face the problem of unfair ratings by artificially inflating or deflating reputations (Resnick and Zeckhauser 2002; Resnick et al. 2000; Corritore et al. 2003; Yang et al. 2002). They are vulnerable to a number of potential attacks, such as Sybil attack, on–off attack, independent/collaborative bad-mouthing attack, and conflict behavior attack, ballot-stuffing attack, newcomer attack (Douceur 2002; Liu et al. 2008a). The usage of pseudonyms introduces new challenges, since it makes hard to trace malicious behaviors. It also influences the accuracy of reputation. Sun et al. proposed a number of schemes to overcome some of the above attacks, but they did not consider the additional challenges caused by privacy preservation (Sun et al. 2006a, b) and context-awareness. We aim to overcome the above traditional attacks and challenges in our designed trust management system for MSN.

9.2.4 Projects of Mobile Social Networking

There are a number of projects related to our work. Herein, we give a brief review.

A number of research projects have social networking functions, such as Microblog (MicroBlog 2012) and AdSocial (Studi et al. 2008). However, trust and reputation in social networking are not considered in these projects.

Micro-blog by Duke University supports query, chatting, and recommendation in ad hoc networks. Posts about different contents float on the map of the application user interface. Whenever a user travels, posts about the location are floated to the user. If there is not enough information, users can query in ad hoc networks, and replies could be added to the current location. It can be applied in many application areas such as tourism, advertisement, emergent alerts, and so on.

AdSocial is a social network based on ad hoc networks. Important functions include presence detector, buddy search and chatting, VoIP calls, video calls, and ad hoc games. It supports real-time communication with friends. Example use cases are chatting in the bar or on the train. Besides that, AdSocial supports multiple message exchange methods, e.g., via Wi-Fi and Bluetooth.

9.2.5 Recommendation Services

Advogato (http://www.advogato.org/) is an online service to provide a platform for free software developers to advocate software and promote research. One significant impact of Advogato is the trust metrics behind the service, which is the basis of many research projects (http://www.advogato.org/trust-metric.html). The Advogato trust metric stimulates users to contribute quality-assured software, and protects against attacks. However, Advogato only applies a centralized architecture for reputation generation, while AwareTrust adopts a hybrid reputation system architecture.

Netflix prize (http://www.netflixprize.com/) is a competition to encourage the design of best movie recommendation algorithms. However, privacy has been a concern for Netflix algorithms since 2007 (http://en.wikipedia.org/wiki/Netflix_Prize). AwareTrust enhances privacy through frequent change of pseudonyms in mobile social networking, but at the same time keeps tracking users' reputation information by applying a centralized trusted server that can map the user's pseudonyms to his/her real identity.

MovieLens (http://www.movielens.org/login) is a movie sharing and recommendation website developed by GroupLens Research (http://www.grouplens.org/) at the University of Minnesota. While MovieLens focuses on movie (content) reputation, AwareTrust can provide node user reputation and content reputation based on social behaviors in the context of MSN content recommendation.

9.2.6 User Interface Trust

A lot of work has been conducted with regard to user interface design in order to improve user's trust, mainly for web sites and in the context of e-commerce. Still, previous activity has left room for further studies on the effects of trust information on social networking and, in particular, on how to provide trust information for mobile users. We plan to use a reputation indicator to indicate each user's local reputation during mobile social networking, and provide detailed information about local reputation generation and global reputation. These are interface design elements that provide the cue of trust information in MSN. In particular, a user's local reputation could affect the credibility of a user's voting on other users. But few previous researches have investigated the visualization of the effect of reputation on mobile users in the context of social networking.

9.2.7 Unwanted Content Control Based on Trust Management

A number of anti-spam solutions have been proposed based on trust/reputation management. Examples are MailTrust (Zhang 2010) and IRGroupRep (Zhang et al. 2009a) based on user feedback, SpimRank (Bi et al. 2008) for instant messaging according to user history tracks. Tang et al. proposed a solution based on spam sender behavior analysis (Tang et al. 2008). A few studies have focused on web spam control, such as TrustRank, Topical TrustRank, PageRank, etc. (Wu et al. 2006; Zhang et al. 2009b; Becchetti et al. 2006). Kolan and Dantu proposed a spam filter based on presence (location, mood, and time), trust, reputation, voice-specific trust and reputation analysis (Kolan and Dantu 2007). This is a specific unwanted traffic control mechanism used to filter voice spam via VoIP. Based on our knowledge, few existing studies have tried to solve the issue of unwanted content control in social networking, particularly in mobile social networking, which, however, is a crucial issue for the success of MSN services.

9.2.8 Recommendation for Privacy Preservation

Recommender systems generally apply an information filtering technique that attempts to recommend information items (e.g., films, books, web pages, music, etc.) that are likely to be of interest to users (Resnick and Varian 1997). Typically, a recommender system compares a user profile to some reference characteristics, and seeks to predict the 'rating' that a user would give to an item they had not yet experienced (Hancock et al. 2007). These characteristics may be from an information item (a content-based approach) or the user's social environment (a collaborative filtering approach) (Su and Khoshgoftaar 2009). In O'Donovan and Smyth (2006), the authors introduced the use of trust as both weighting and filtering in recommendations. The recommendation partners should have similar tastes or preferences and they should be trustworthy, with a history of making reliable recommendations. This trust information can be incorporated into the recommendation process. But to our knowledge, most characteristics used for recommendations are not based on users' private data-sharing behavior which, however, is an important clue to users' preferences on privacy. In AwareTrust, we consider both recommender's trust and service provider's trust in the generation of recommendations. Little existing work provides recommendations on user data privacy preservation, especially in mobile social services with context-awareness (Resnick and Varian 1997; Luo et al. 2009; Yu et al. 2006; Yap 2007; Wang 2010; Zhang and Yu 2007; Chuong et al. 2009; Liiv et al. 2009; Liu et al. 2008b; Xiao 2010; Seetharam and Ramakrishnan 2008; Berkovsky 2010). Although some recommender systems concern privacy preservation, they seldom support mobile social services (Polat and Du 2005; Luo et al. 2009; Li et al. 2009; Bilge and Polat 2010; Ahn and Amatriain 2010; Tada et al. 2010; Kikuchi 2009; Katzenbeisser and Petkovic 2008).

Privacy preservation has been considered in a number of existing recommender systems (Polat and Du 2005; Luo et al. 2009; Li et al. 2009; Bilge and Polat 2010; Ahn and Amatriain 2010; Tada et al. 2010; Kikuchi 2009). Current research focuses on preserving user data provided for generating recommendations without leaking the private information of users to service providers, not on recommendations for the purpose of user data privacy, which is the focus of AwareTrust.

Meanwhile, AwareTrust adopts a trusted server (TS) to support the frequent change of node pseudonyms, in order to avoid potential privacy tracking.

We can find a number of solutions for context-aware or personalized recommendations (Yu et al. 2006; Yap 2007; Wang 2010; Zhang and Yu 2007; Chuong et al. 2009; Liiv et al. 2009; Liu et al. 2008b; Xiao 2010; Seetharam and Ramakrishnan 2008; Berkovsky 2010). However, none of them provide recommendations on user data privacy preservation. Most of them, e.g., CoMeR (Yu et al. 2006) and SCAMREF (Zhang and Yu 2007), apply client–server architecture, while AwareTrust supports both centralized and distributed recommendation generation at the trusted server and in MSN nodes. AwareTrust uses a context ID to identify a context that is described with context parameters and their values. It supports context-awareness by recommending a private data-sharing strategy on different types of data in different contexts.

In the literature, distributed trust and reputation evaluations have been studied in MANET (Sun et al. 2006a; Theodorakopoulos and Baras 2006; Raya et al. 2008). Few solutions support node privacy. The existing recommendation-based trust/reputation mechanisms aim to support secure node collaboration for the purpose of routing and networking. Infrequently, they consider recommendation issues in applications and services using MANET as a social networking and computing platform (Zouridaki et al. 2006).

9.2.9 Trust and Privacy in Mobile Social Networking

In current literature, the most widely investigated privacy concern is user personal data and location (Krishna et al. 2010; Chen and Rahman 2008). Common solutions include data encryption and key distribution (Chen and Rahman 2008), introducing a server to issue anonymous identities (Beach et al. 2009). In order to validate the effectiveness of their solutions, researchers also evaluate the robustness against several attacks, such as eavesdropping, spoofing, replay, and wormhole attacks (Beach et al. 2009). Several MSN applications have addressed the privacy concern by allowing users to decide privacy policy and configure privacy settings. Sadeh et al. (2009) discussed related issues in the case of PeopleFinder, an application that lets users share their location with others, with which users can refine privacy policies over time. Miluzzo et al. (2008) investigated data privacy issues in CenceMe, a mobile social networking application using sensors. The solution to the privacy problem is to allow users to manually configure which sensors to use on their mobile phones. This means users have to disable certain sensors in order to achieve privacy goals. AwareTrust advances the art by integrating and adopting the advantages of various mechanisms, including trust evaluation, reputation generation, privacy recommendation, unwanted content control, encryption, MSN communication data access control, and identity issuing from the trusted server, as well as end-user setting, etc. Furthermore, our model is more sophisticated by integrating privacy into the trust mechanism in MSN.

9.3 Need Assessment

We conducted a need assessment survey in order to explore the potential usefulness and significance of developing a trust management system for MSN, and user preference with regard to reputation visualization. We applied a five-point Likert scale in the survey.

9.3.1 Design

The survey contains two parts. The first part evaluates the potential significance of developing trust/reputation management for MSN based on three MSN scenarios:

Scenario 1: sharing the cost of ‘buy three pay for two’ goods in a shopping mall: Right now you are at a shopping center, and a product you want is on sale with a condition ‘buy three pay for two’. However, you only need one. You want to ask people nearby, whom you don’t know, via your mobile phone whether they would like to share the discount with you.

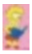
Scenario 2: sharing the price of a packet of five movie tickets in front of a movie theater. After shopping, you want to watch ‘Avatar’ in a movie theater. The ticket price is 13.8 €. However, if you buy a packet of five tickets, it will be 8.6 € for each. You want to share the ticket packet with people nearby whom you don’t know. You discuss whether they would like to share the discount with you via your mobile phone.

Scenario 3: sharing a taxi ride after movie: After the movie, a lot of people are leaving the theatre. You want to watch a figure skating competition quite far away. You would like to take a taxi and think about sharing a ride. You discuss with people nearby via your mobile phone whether they would like to share the ride with you.

The participants were asked to express their opinions on the usefulness of a reputation system in the above MANET chatting scenarios.

The second part attempts to study user preference on reputation visualization. We proposed four visualization methods, as illustrated in Table 9.1: UI1 – reputation is indicated based on the font size of an input chatting message; UI2 – reputation is indicated by the number of stars; UI3 – reputation is indicated through a growing process of a cartoon character; UI4 – reputation is indicated through a role in a community, which can be customized by a user. We asked the participants to mark their preferences. Note that UI2 is a traditional reputation visualization method applied by Amazon and eBay.

Table 9.1 Reputation visualization methods

Reputation visualization methods
UI1) Based on the font size of the input text of a chatter: the bigger size of the font, the more reputable.
UI2) Based on the number of stars, the more the higher reputation, e.g .
UI3) Through a growing process of a cartoon character, the more mature the higher reputation, e.g. .
UI4) Through a role in a community, a user can personally select or define the roles and their represented reputation levels, e.g., some characters from the Simpsons (http://www.thesimpsons.com/).
 represents few reputable histories with trouble records;
 represents some reputable histories with trouble records;
 represents some reputable histories without trouble records;
 represents high reputation with some trouble records;
 represents high reputation without trouble records.

9.3.2 Participants and Results

The survey was distributed through a mailing list. We conducted it in Finland and China, and collected the survey response via email. A small gift was awarded to each participant. We got a total of 107 valid responses; among them, 83 were university students, 68 male (63.6 %) and 39 female (36.4 %). Most participants were between 21 and 28 years old. All of them had Internet chatting experiences, 84.1 % had mobile Internet chatting experiences, and 18.7 % of them had experience on MANET-based mobile social chatting.

The survey result and its implication are summarized as below:

1. The average rating scales regarding the potential usefulness of a reputation system in three mobile social chatting scenarios were 3.74 (SD = 1.08), 3.90 (SD = 0.98) and 4.00 (SD = 1.00) respectively. All of them are over 3.5. This implies that a reputation system for MSN (e.g., pervasive social chatting) is thought to be useful in some scenarios. Therefore, this is a significant contribution to the design and development of such a system.
2. Most participants preferred the traditional reputation visualization style UI2 (with an average value 4.07), but they were also interested in the new styles UI3 and UI4 proposed by us (UI3-3.32 and UI4-3.21). However, the font size based

reputation indication was not preferred (UI1-2.50). Some participants commented that UI4 design is very interesting, and expect an implementation for optional selection. In addition, other schemes of UI2 design (e.g., number of crowns or diamonds) were also preferred by the participants. Thus, personalized reputation visualization is suggested in practice.

9.4 AwareTrust: A Trust Management Framework for MSN

We designed AwareTrust, a trust management framework for mobile social networking. This framework can provide context-aware trust management for different MSN scenarios, and support key management for secure data access control, unwanted content control, and user privacy preservation, etc.

9.4.1 System Design Requirements

A number of constraints and requirements are raised in the practical deployment of such a framework. We summarize them as basic requirements of AwareTrust design.

First, the framework is expected to support user privacy. Node real identity should be hidden in MSN. Node pseudonyms could be frequently changed in order to enhance privacy and avoid potential tracking; while the node real identity can be registered at a trusted party.

Second, the nodes can connect to the trusted server periodically (e.g., once per day), although real-time connection is not always available.

Third, a small amount of transferring data is preferred among MSN node-to-node communications in order to achieve effective power consumption (Ahtiainen et al. 2009). However, there is no such strict constraint with regard to message length for the node and TS communications (Ahtiainen et al. 2009).

Fourth, AwareTrust aims at overcoming potential system attacks. In particular, AwareTrust can fight against attacks on AwareTrust, such as bad-mouthing attack, collaborative bad-mouthing attack, on–off attack, conflict behavior attack, etc.

Fifth, the system provides a mechanism to allow each node to control/filter traffic sourced from a distrusted node. It has the capability of identifying a malicious node and filtering messages from it (based on reputation evaluation).

Sixth, the information or message communicated between the TS and AwareTrust nodes should be protected to avoid eavesdropping; the MANET communications should be protected at a certain level in order to achieve expected efficiency.

Finally, AwareTrust should support various mobile social services, and provide trust and reputation information according to the context of social behaviors. That is, AwareTrust supports scalability, flexibility, and context-awareness.

9.4.2 *Reasons to Introduce a Centralized Trusted Server*

Introducing such a hybrid trust management system structure has a number of merits. First, this design can support privacy preservation by frequently changing the pseudonyms of nodes and avoid the inconsistent reputation problem. It can support accurate node reputation/trust evaluation based on the registered unique node ID, even though the node pseudonyms could be changed. Second, this design provides an economic approach to the collection of useful data from mobile social networking, which can further support other promising services, e.g., location-based content recommendation services. Thereby, the design can potentially support new business models. Finally and more importantly, this design is flexible in order to provide the reputation information whether the TS is available or not. AwareTrust can generate reputation in either a distributed or centralized manner or both.

In AwareTrust, we assume that TS is trustworthy enough to preserve the private data of nodes. The potential weakness is that each party should trust TS, thus the TS could be the target of attackers. Notably, AwareTrust can also work in a distributed way based on the node pseudonyms when TS is not available.

9.5 AwareTrust System Structure

We can find both distributed and centralized reputation architecture in the literature. Both have advantages and disadvantages (Jøsang et al. 2007). We attempt to utilize their advantages in AwareTrust. Figure 9.1 illustrates AwareTrust's structure. It is a hybrid reputation management infrastructure. At each node device, a *user behavior observer* records node communications (e.g., chatting) to generate real user behavior based trust cue (Yan et al. 2012a). An MSN UI (i.e., a set of MSN applications, e.g., TWIN, Ghost Talk, Facebook/LinkedIn friends) provides a user interface for the node user to do social networking. *Communication reporter* and *voter* report the communication records and local reputation information to the TS. Meanwhile, the user can also vote other entities through it to the TS. *A reputation evaluator* evaluates the entity's reputation, and provides the results to the user via a *reputation visualization* UI. A *reputation extractor* receives the reputation tokens (containing the node reputation value) issued by the TS. Trust DataSet stores all data related to the above functional blocks in the node in a secure manner. In addition, a *node profile manager* is used to maintain node user's personal information. It can communicate with the TS to register the node into the AwareTrust system and update the node pseudonym and reputation token.

At the *trusted server*, a *reputation generator/predictor* calculates the node reputation values; meanwhile it identifies malicious nodes. A *reputation distributor* distributes the reputation tokens (containing the node's reputation value) to each node periodically or by request. A *node ID manager* handles node registration and issues new node pseudonyms (by request or periodically). An *information receiver*

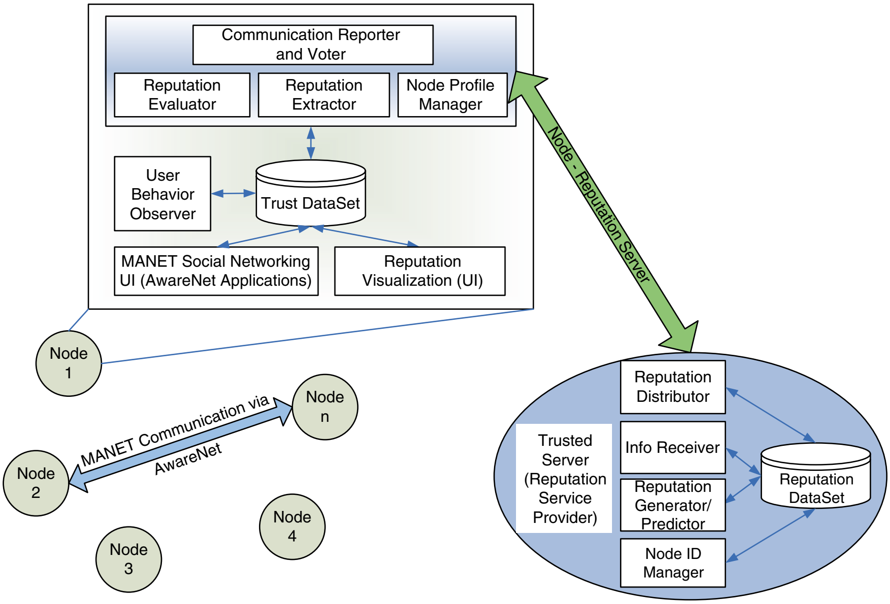


Fig. 9.1 AwareTrust system structure

collects the records reported by the nodes and saves them into *reputation DataSet*, which also saves the reputation token of each node and its real ID and pseudonyms.

9.5.1 AwareTrust Model

The way to calculate trust is often called the trust model (Yang et al. 2002). Based on the above system design, we propose a hybrid trust model – AwareTrust model to generate node user reputation, as shown in Fig. 9.2. This information is evolved at both node sides based on ephemeral local experience and TS side according to collected historical MSN information.

In AwareTrust, TS generates/predicts node reputation according to the context model and issue a reputation token to the node. The reputation token contains node reputation values, context IDs, pseudonym, and token expired time. Based on the reputation token attached to each node and mobile social networking experience, a node generates the reputation of other nodes instructed by the current context model. Furthermore, the node updates other nodes' reputation values based on their social networking behaviors and performance (according to the context model). Particularly, the node reports mobile social networking experiences and opinions to TS (e.g., report past chatting records, or vote other users who chatted with him/her). After collecting additional mobile social networking records, the TS updates node

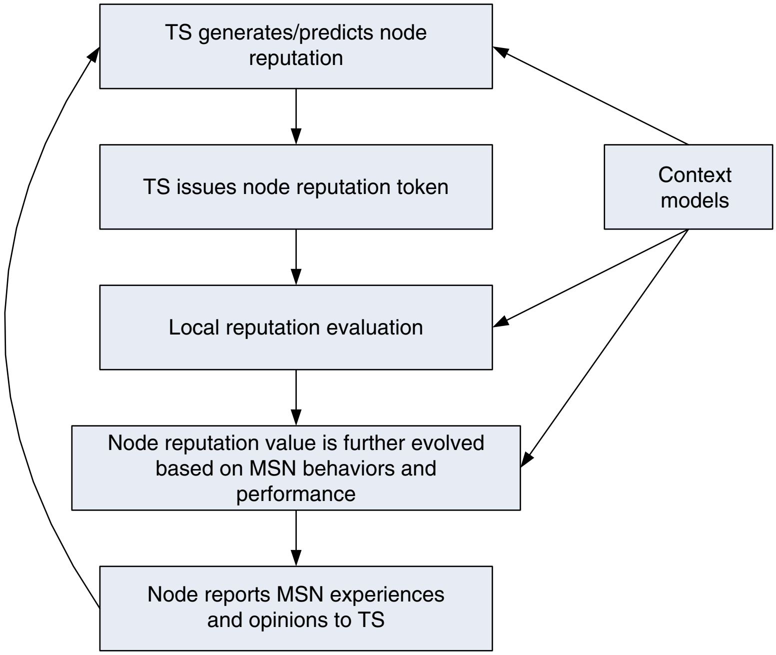


Fig. 9.2 AwareTrust model

reputation according to his/her performance in different contexts. The evaluation of node reputation is iterative at both the node and TS, based on newly accumulated experiences and information.

In order to preserve privacy, AwareTrust applies the TS to issue and manage the pseudonyms of nodes. The TS can identify the real identities of nodes. It could frequently or periodically issue a new pseudonym to the node. Thus, the local experiences are accumulated only based on valid pseudonyms. For example, node A would assume node B as a different node once node B uses a new pseudonym, even though they had interaction with each other before. Historical evaluation on the node reputation can only be conducted at the TS by considering all social networking behaviors in different contexts related to a concrete node that possibly has multiple pseudonyms. Notably, a new reputation token is always issued when the node pseudonym is changed. However, the node may request the TS to update the reputation value by issuing a new reputation token without changing its pseudonym.

The pseudonym can be the node ID expressed by its MAC address and/or network layer address. This ID could be frequently changed in order to enhance the node privacy. The system supporting the node pseudonym is selected from a preprepared list stored at both the trusted server and each node. Thus, the server knows 'who is who' after getting the reports from the nodes, even though the node could change the pseudonym without contacting the server.

Particularly, the TS also collects direct votes on node users and the statistical data of node interaction, for the purpose of evaluating the reputation of the node. The node reputation token is issued by the TS with non-repudiation. Due to the usage of pseudonyms, only the TS can evaluate the node reputation in an accurate way based on historical mobile social networking behaviors. Thus, introducing the TS can greatly improve the precision of reputation evaluation (Yan and Chen 2010). Optionally, the TS is also responsible for issuing the blacklist of malicious nodes and the favorite list of honest nodes to the AwareTrust nodes, if needed in practice. For example, the blacklist can be used to filter or control messages from disreputable nodes.

The context model is the way to describe the situations that need AwareTrust's assistance and specify the way to adopt a suitable reputation evaluation algorithm. Applying the context models, AwareTrust can support various applications that need AwareTrust's assistance, thus providing sound system flexibility and scalability.

The example of the context model for mobile social chatting is:

```
<?xml version="1.0" encoding="UTF-8"?>
<contextIdentity='AwareTrust_chatting'>
  <contextDescription>
      <name> chatting service in MSN </name>
  </contextDescription >
  <operationRequest>
      <operation>
          <name> On-Chat reputation evaluation at node </name>
           <host> node side </host>
           <algorithm> AwareTrust_algo1 </algorithm>
      </operation>
      <operation>
           <name> Reputation generation at TS </name>
           <host> TS side </host>
           <algorithm> AwareTrust_algo2 </algorithm>
      </operation>
      <operation>
           <name> Personal reputation prediction </name>
           <host> TS side </host>
           <algorithm> AwareTrust_algo3 </algorithm>
      </operation>
    </operationRequest>
```

Based on the context model, the AwareTrust system knows which algorithm should be applied in which context based on the context ID and operation name. We have prototyped two demo systems supported by AwareTrust; one is an ad hoc content recommendation service ("AdContRep") (Yan and Chen 2010) and the other is an ad hoc chatting service ("AdChatRep") (Yan and Chen 2011). Additional MANET social services can be supported by AwareTrust by adding new context models into the system.

It is important to note that the node reputation values generated in different contexts can be further combined to get a single value that represents the general node reputation. In addition, the reputation generated in other contexts can play as an important reference for decision-making in a new context if the reputation value linked to the underlying context is not available. AwareTrust supports multiple contexts. It can evaluate node reputation for multiple instances or events in different contexts. The algorithms used in different contexts could be different in *reputation evaluator*. The reputation evaluator has the capability of evaluating and maintaining the reputation values in various contexts with the help of Trust DataSet.

9.6 Properties for Trust

AwareTrust supports a number of key system properties for achieving trust in mobile social networking. This section introduces a number of key properties for trust, such as trust/reputation generation, trustworthy content recommendations, secure data communications, unwanted traffic control, user privacy recommendation and preservation, and other trust and privacy enhancement technologies.

9.6.1 Trust/Reputation Generation for Mobile Social Chatting

To generate the local reputation of a node $R(i \to j)$ (i.e., node j's local reputation evaluated by node i), we consider to aggregate three trust impact aspects (Yan and Chen 2011). The first part is the sum of previous local reputations (or personalized reputations if any) and the general reputation, which serves as the initial reputation of current chatting. The second part is the reputation generated on the basis of current chatting experience. Since there could be multiple on-chat votes during chatting, we integrate them together by averaging the product of on-chat votes and the depth of chatting at the vote, which impacts the preciseness of the opinion of a user. Furthermore, this part is weighted by $is(i,j)$ (which is *the interest similarity between node i and j*) and $cl(i)$ (which is *the crucial level of chatting topic from the view of node i*), which are other two on-chat factors influencing trust, as we explored in a user survey (Yan and Chen 2011). The third part is generated based on the on-chat votes on node j provided by other nodes than i, which is certified by an opinion deviation factor *od* as described in Formula (9.2). The opinion deviation factor indicates the opinion deviation of two nodes on a target node. Applying this factor makes it easy to figure out the nodes that hold different opinions from the reputation evaluating node. Thus, applying it in reputation generation can avoid the negative influence of a bad-mouthing attack. We apply Formula (9.1) by considering the first four on-chat factors explored in the survey that are available at each node (Yan and Chen 2011).

$$R(i \to j) = f\begin{pmatrix} \alpha\left(R'(i \to j) + R(j)\right) + \beta\sum_{l=1}^{L}\left\{cv(i \to j)_l * dc(i,j)_l\right\} \\ *is(i,j) * cl(i) + \gamma\sum_{k \neq i}\sum_{l=1}^{L'}\left\{cv(k \to j)_l\right\} * od(i \leftrightarrow k, j) \end{pmatrix} \tag{9.1}$$

$$od(i \leftrightarrow k, j) = 1 - 2 * \left| f\left\{\begin{array}{l} \sum_{l=1}^{L} cv(i \to j)_l * dc(i,j)_l \\ -\sum_{l=1}^{L'} cv(k \to j)_l * dc(k,j)_l \end{array}\right\} - \frac{1}{2} \right| \tag{9.2}$$

where $R(i \to j)$ is the reputation of node j locally evaluated by node i. $R'(i \to j)$ denotes the personalized reputation predicted and issued by TS, or the reputation previously evaluated by node i on j. $R(j)$ is the general reputation of node j issued by TS. $cv(i \to j)_l$ is the lth on-chat voting by node i on the message of node j. $is(i,j)$ is the number of common communities shared by node i and j, which indicates their common interests. $cl(i)$ is the crucial level of chatting topic. $dc(i,j)_l$ is the depth of chatting between nodes i and j at the time of the lth voting of node j. It is the minimum number of messages input by node i and j at the time of the voting. For example, if node j has input four messages and node i has input six messages at the time of the lth voting of node j on i during their chatting, $dc(i,j)_l = 4$. $f(x) = \frac{1}{1+e^{-x}}$ is the Sigmoid function used to normalize an arbitrary value into (0, 1). L denotes the total number of on-chat votes by node i on j. L' denotes the total number of on-chat votes by node k on j. α, β, γ are parameters to indicate the weights of different contributions. Note that $\alpha + \beta + \gamma = 1$. $od(i \leftrightarrow k, j)$ is the opinion deviation factor that indicates the difference of opinions between nodes k and j on the chatting messages input by node i in the same chat. Note that Formula (9.1) is suitable to be applied if AwareTrust users would like to use their own opinions to measure the trustworthiness of other people in the mobile social chatting. As shown in our simulation evaluation, Formula (9.1) is effective to overcome a collaborative bad-mouthing attack.

Based on afterwards voting and local reputation, we generate two types of node reputation at TS: personalized reputation and general reputation. We apply weighted aggregation using local reputation $R(i \to j)$ as credibility to overcome an unfair rating attack. Meanwhile, we also consider time influence on the afterwards voting in order to overcome on–off and conflict behavior attacks (Sun et al. 2008).

Formula (9.3) is applied to generate the personalized reputation $\overline{R(i \to j)}$ of node j evaluated by node i by considering the fifth factor – afterwards voting and time decaying. The afterwards voting is the vote on a party after chatting based on interaction experiences.

$$\overline{R(i \to j)} = \frac{1}{O}\sum_{m} R(i \to j)^{t_m} * V_i^{j(t_m)} e^{-\frac{|t-t_m|^2}{\tau}} \tag{9.3}$$

where $O = \sum_{m} R(i \rightarrow j)^{t_m} * e^{-\frac{|t-t_m|^2}{\tau}}$; $V_i^{j(t_m)}$ is the afterwards voting of node i on node j at time t_m; t is the calculation time of node reputation; parameter τ is used to control time decaying. $R(i \rightarrow j)^{t_m}$ is the local reputation of node j reported by node i at time t_m, with afterwards voting $V_i^{j(t_m)}$ attached. If $V_i^{j(t_m)}$ is not provided by the node, we set $V_i^{j(t_m)} = 0.5$. Note that $V_i^{j(t_m)} \in [0,1]$ and $R(i \rightarrow j)^{t_m} \in [0,1]$.

To get the general reputation of node j, denoted as $R(j)$, we aggregate the evaluation of all nodes $\overline{R(i \rightarrow j)}$ based on Formula (9.4). The aggregation is weighted by the general reputation $R(i)$ as credibility. Meanwhile, we also consider the influence of the number of reputation contributors on the general reputation generation, since the more contributors, the more convincing the generation result is.

$$R(j) = \frac{f(K)}{W} \sum_{i=1}^{K} R(i) * \overline{R(i \rightarrow j)} \quad (i \neq j) \tag{9.4}$$

where, K is the total number of nodes who have direct experiences with node j. $W = \sum_{i=1}^{K} R(i)$ is the total sum of the general reputation values of those nodes. $f(K) = \left\{1 - \exp\left(\frac{-K^2}{2(\sigma+\varepsilon)^2}\right)\right\}$ is the Rayleigh cumulative distribution function to model the impact of K on node reputation, $\varepsilon = -K/K'$, is a factor to indicate the sociability of node j. Parameter K' is the total number of users registered in the system.

9.6.2 Trustworthy Content Recommendations

Figure 9.3 illustrates the process for generating trustworthy recommendations in MSN (Yan and Chen 2010). As shown in Fig. 9.3, the query node processes query response about content recommendations. The trust evaluator calculates the content reputation based on the votes provided by recommender nodes, the number of collected recommendations (i.e., content popularity) and the recommender node trust values certified by the TS. The node trust issued by TS is further evolved based on recommendation performance at the query node. The evaluated content reputation will be displayed to help the node user make a selection decision. The query node will report its mobile social activities to the TS. Thus the TS can generate content reputation by aggregating all content recommendations, and precisely evaluate node trust based on node real IDs. The TS can issue the trust token that contains the node trust and update the content reputation periodically or by request. Herein, the trust value of a recommender node partially serves as the credibility of its recommendation. This trust value contains two parts: one is provided by the TS by assessing historical node recommendation behaviors. The other is the query node

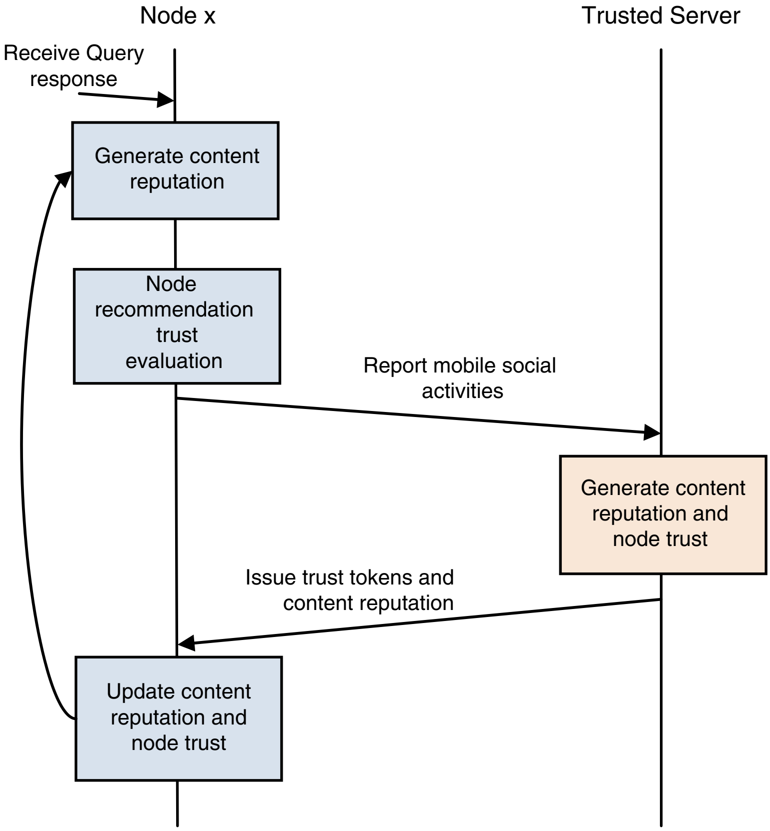


Fig. 9.3 Process for generating trustworthy recommendations in MSN

evaluation on the recommender node based on locally accumulated experiences. The local evaluation is evolved on the basis of the trust value issued by the TS. Notably, credibility provides a reason to trust. Recommendation credibility is essential for generating a reliable reputation value.

In order to preserve privacy, AwareTrust applies the TS to issue and manage the pseudonyms of nodes. The TS can identify the real identities of nodes. It could frequently or periodically issue a new pseudonym to the node. Thus, the local experiences are accumulated only based on valid pseudonyms. For example, node A would assume node B as a different node once node B uses a new pseudonym, even though they had had interaction with each other before. Historical evaluation on the node recommendation trust can only be conducted at the TS by considering all recommendation behaviors related to a concrete node which has possibly applied multiple pseudonyms. Notably, a new trust token is always issued when the pseudonym is changed. However, the node may only request an update to the trust value by issuing a new trust token without changing its pseudonym.

Particularly, the TS also collects direct votes on different contents, and the statistical data concerning node interaction for the purpose of evaluating the node trust and content reputation. The node trust token (containing the node trust value) is issued by the TS with non-repudiation. Due to the usage of pseudonyms, only the TS can evaluate the node trust in an accurate way based on historical recommendation behaviors. Thus, introducing the TS can greatly improve the precision of trust/reputation evaluation.

Meanwhile, the TS can generate the reputation of various contents based on the feedback and recommendation records reported by the query nodes. A content recommendation service based on content reputation can be provided. The evaluation of node trust and content reputation is iterative at both the node and TS, based on newly accumulated experiences and information. The algorithms used for trust/reputation evaluation and recommendation generation at a query node and TS are described in Yan and Chen (2010).

9.6.3 Secure Data Communications

AwareTrust uses either a local trust (LT) evaluated at nodes or a general trust (GT) generated at TS or both to control access of data in the MSN communications. Any MSN node can select other nodes with at least a minimum level of local and/or general trust for secure communications. The nodes with a lower trust level cannot access the data sent from the source node.

In the case that TS is available and the node would like to control its data access only based on the general trust level, the general trust level controlled access keys (i.e., encryption public keys and personalized decryption keys) are generated and issued by the trusted server.

In the case that the server is not available, each user generates the encryption key and corresponding personalized secret keys based on the local trust level for decryption. The user issues the secret keys to those users that satisfy the decryption conditions. Then, it broadcasts the encrypted messages to nearby nodes. In case that the local trust levels of some users have a big change, the node will re-generate suitable keys for encrypting and decrypting later communication data sent from him/her and then re-send new secret keys to eligible users.

In the case that the trusted server is available and the user would like to control communication data access with both trust levels (the general trust level and the local trust level), based on periodically issued keys from the server, the user further controls the access to their data by generating new public/secret keys based on the server-issued keys and local trust level.

We have proposed a concrete solution to achieve above secure data communications based on trust with attribute-based encryption schemes (Muller et al. 2008).

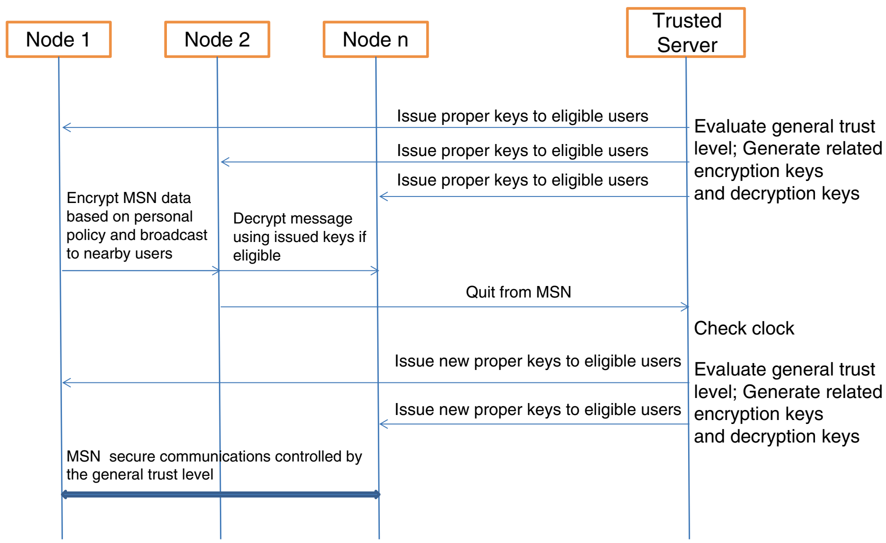


Fig. 9.4 Secure MSN communications controlled by general trust

9.6.3.1 Secure MSN Communication Procedures

We illustrate the procedures of flexible secure mobile social networking based on two-dimensional trust levels in three cases.

In the case that the data access is only controlled by the general trust level, the general trust level controlled access keys (i.e., encryption public keys and personalized decryption keys) are generated and issued by the trusted server (e.g., periodically). The procedure is shown in Fig. 9.4.

Issue keys: TS evaluates the general trust level of nodes; it generates encryption public keys and personalized decryption secret keys of attributes, and then issues the proper keys (both public encryption keys and personalized decryption keys) to eligible nodes.

Secure MSN communications: the node encrypts its message based on personal access policies using corresponding public keys and broadcast to nearby nodes; other nodes check the encryption policy, and use their personalized decryption keys to decrypt the messages if eligible to access.

Quit from MSN: a node may quit the MSN at any time by sending a request to TS; the TS confirms this by not sending any new keys to this node after its current keys are expired.

Re-issue keys: TS checks the clock frequently; if the current keys' valid period has expired, it will re-evaluate the general trust level, then re-generate encryption public keys and personalized decryption secret keys, set their valid period, and re-issue the new keys to the eligible nodes.

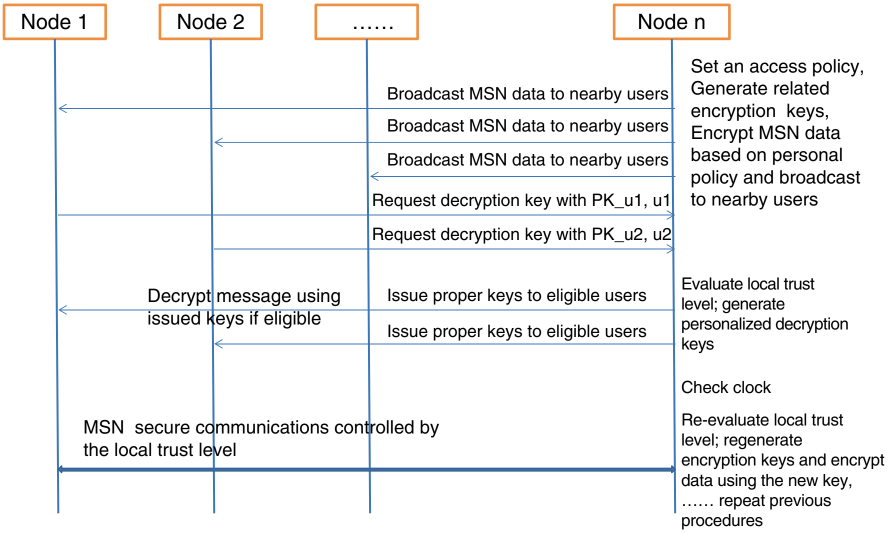


Fig. 9.5 Secure MSN communications controlled by local trust

The MSN nodes continue secure communications with the new keys when their valid period starts.

In the case that the server is not available, each node generates the encryption key and corresponding personalized secret keys based on the local trust level for decryption. It issues the secret keys to those nodes that satisfy the decryption conditions. Then, it broadcasts the encrypted messages to nearby nodes. In the case that the local trust levels of some node change, the node will re-generate suitable keys for encrypting and decrypting later communication data and then re-send new secret keys to eligible nodes. The procedure is shown in Fig. 9.5.

Protect MSN data: a node u sets an access policy, generates related encryption keys, and encrypts its data in MSN based on the personal policy and broadcast to nearby users. After detecting the broadcast message from this node, other nodes check the local trust value of the node and decide whether to communicate with it. If the decision is positive, the other node sends a key request to the node u with identity. The node u evaluates the local trust level of requesting nodes; it generates personalized decryption secret keys of local trust, and then issues the proper keys (personalized decryption keys) to eligible nodes.

Secure MSN communications: the eligible node u' decrypts the message of the node u based on personal access policies using corresponding secret keys.

Re-generate keys: The node u checks the clock frequently. If it is time to re-evaluate the local trust, the node re-evaluates the local trust of other users. If the results have a big difference, the node u will re-generate encryption keys based on a new access policy. It will adopt the new keys to protect later MSN communication data.

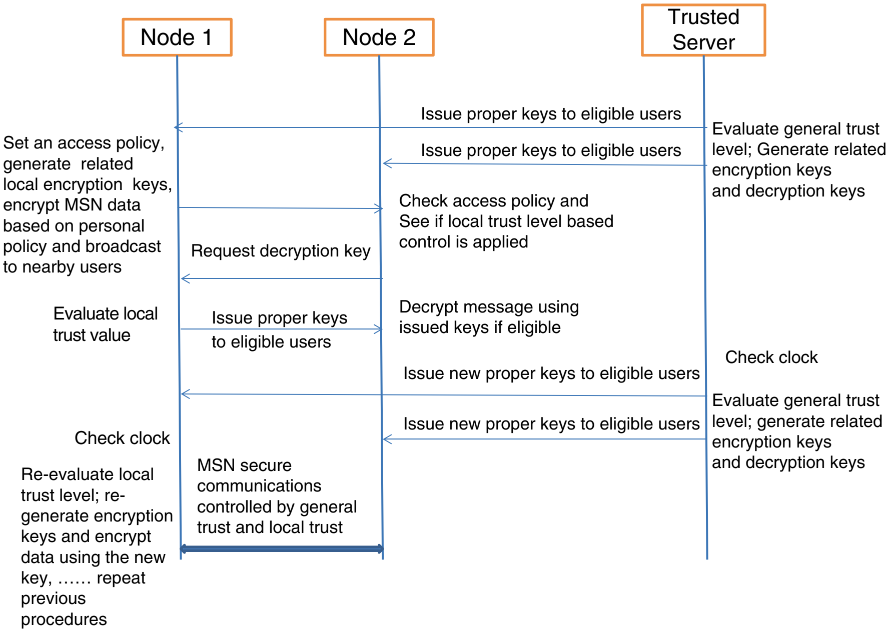


Fig. 9.6 Secure MSN communications controlled by both general trust and local trust

In the case that the trusted server is available and the node would like to control communication data access based on both trust levels (the general trust level and the local trust level), based on periodically issued keys from the server, the node further controls the access of its data by generating new public/secret keys based on the server-issued keys and its own issued keys. The procedure is illustrated in Fig. 9.6.

Issue keys: TS evaluates the general trust of nodes; it generates encryption public keys and personalized decryption secret keys of attributes, and then issues the proper keys (both public encryption keys and personalized decryption keys) to eligible nodes.

Protect MSN data: a node u sets an access policy, generates related encryption keys, encrypts its data in MSN based on the personal policy policies using corresponding public keys (issued by both TS and itself), and broadcast to nearby nodes. After detecting the broadcast message from the node, other nodes check the local trust value of the node and decide whether to communicate with it. If the decision is positive, the other node sends a key request to the node u with identity. The node u evaluates the local trust level of requesting nodes; it generates personalized decryption secret keys of local trust, and then issues the proper keys (personalized decryption keys) to eligible nodes.

Secure MSN communications: the eligible node u' decrypts the message of node u based on personal access policies, using corresponding secret keys issued by both TS and the node u.

Quit from MSN: a node may quit the MSN at any time by sending a request to TS. The TS confirms it by not sending any new keys to this node after its current keys are expired.

Re-issue keys: TS checks the clock frequently; if the current keys' valid period has expired, it will re-evaluate the general trust, then re-generate encryption public keys and personalized decryption secret keys for eligible nodes, set their valid period and re-issue the new keys to the eligible nodes.

Re-generate keys by nodes: The node u checks the clock frequently. If it is time to re-evaluate the local trust, the node re-evaluates the local trust of other nodes. If the results have a big difference from previous local trust, the node u will re-generate encryption keys based on a new access policy. It will use the new keys to protect later MSN communication data.

The MSN nodes continue secure communications with the new keys when their valid period starts. The communication data are protected based on both the general trust and the local trust.

9.6.4 Unwanted Content Control

The unwanted content control at the node in mobile social networking is based on content monitoring, the behavior of unwanted content handling, and the collection of broadcast complaint on unwanted contents. The control procedure contains two parts in parallel: local detection and unwanted content control.

As shown in Fig. 9.7, local detection is conducted automatically in a MSN node. It monitors the node inbound and outbound traffic in order to detect if the local node is infected (when outbound traffic is sharply increased for the purpose of intruding other nodes) or intruded (if inbound traffic is unaffordable by the node considering its remaining power and processing capability). Meanwhile, the node also monitors node behavior on content maintenance, which is an important clue to indicate unwanted content by a content consumer that can be referred by other nodes.

The procedure of unwanted traffic control in a MSN node is depicted in Fig. 9.8. The process is triggered by the events that the node receives a complaint or the local detection is negative. Firstly, the node evaluates the trust values of other nodes based on local detection and the received complaint. If the trust value of a MSN node is below a threshold, the node controls the content from that node in MSN, especially for the contents that were complained about by other nodes. Detailed algorithms and evaluation results of the above solution are reported in Yan et al. (2012b).

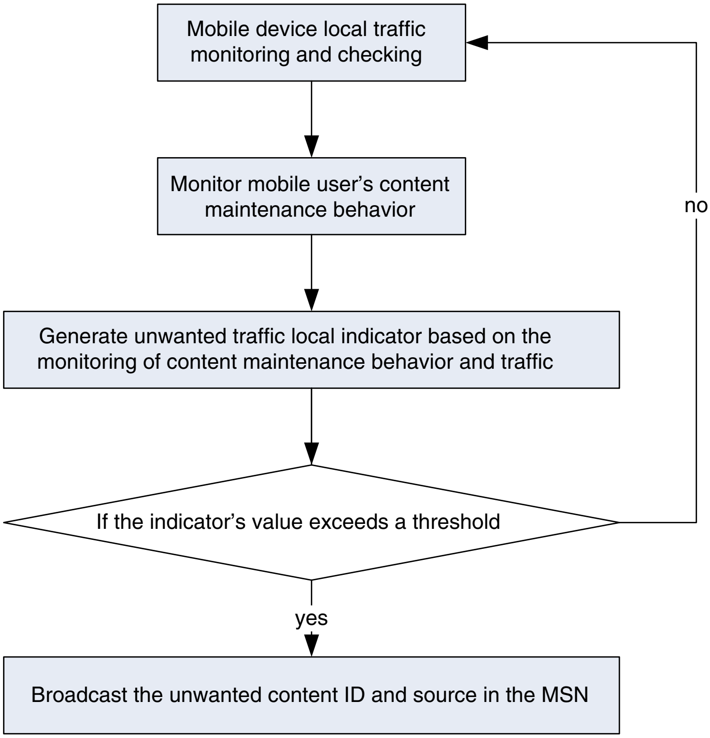


Fig. 9.7 Procedure for local unwanted content detection

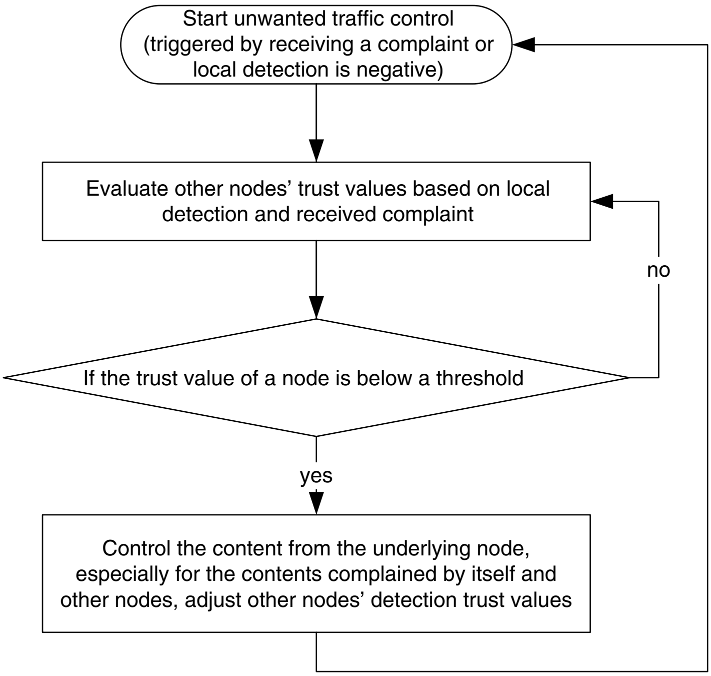


Fig. 9.8 Procedure for unwanted content control at the MSN node

9.6.5 User Privacy Recommendation and Preservation

Based on the AwareTrust structure, we have designed a number of algorithms for generating context-aware user privacy recommendation as described below:

- When TS is accessible, the recommendation vector is calculated at the TS and provided to a requesting node.
- The recommendation vector is fine-tuned/calculated at the node with/without the input from TS.

We apply a trusted server that knows the node's real identity to calculate a recommendation vector based on long-term historical experience. The vector can be also generated at each MSN node according to recent experiences accumulated based on node pseudonyms, while this vector could be fine-tuned when the recommendation server is accessible. We further designed a couple of algorithms to generate context-aware recommendations for MSN nodes. Concretely, the recommendation vector is calculated based on a comprehensive number of factors such as data-sharing behaviors and behavior correlation, service popularity and context, personal data type, community information of nodes, and trust value of each involved party. It represents recommendation values on different data-sharing settings in a given context.

The privacy recommendation helps the node to make a decision as to whether it is safe and proper to provide a specific piece of personal information to a third party in a MSN service. The recommendation provided is personalized based on the correlation of past data-sharing behaviors of nodes in various situations. It represents recommendation values on different nodes data sharing in a specific context. Thus, AwareTrust contributes context-aware personalized recommendations on user data privacy in MSN services. For detailed algorithms and solutions, refer to Yan and Chen (2010); Yan and Zhang (2011).

9.6.6 Other Trust and Privacy Protection Technologies

In this section we list other technologies that can be used for privacy and trust purposes in MSN.

1. *Privacy-Triggered Networking*

Sometimes mobile social communications and networking do not require instant delivery. The main idea of this technology is to "hide in the crowd" in the following sense: a user is able to dynamically monitor her environment and her derived privacy level (using an appropriate metrics), and she gets involved in communications only if the privacy level is sufficiently high. The protection has been designed for an adversary model where an eavesdropper can relate the source of a radio message to a certain location but only with a limited accuracy. Details about the technology, implementations, and experiments can be found in Jadliwala et al. (2011).

2. *Pseudonyms*

One strategy to protect identities is to use pseudonyms that are changed often enough. In the Nokia Instant Community trial, which is explained later in the chapter, a pseudonym change scheme was used that takes into account information about user's environment and utilizes a "mix zone" concept. This ensures that pseudonym changes are done in a coordinated manner, in order to increase the confusion created for a potential eavesdropper. In the trial, the eavesdropper was implemented using a network of tens of WLAN access points located on the EPFL campus. One main finding of Bindschaedler et al. (2012) was that such a powerful attacker is able to track users with a reasonably high success rate (while they stay under the coverage of the attacker network) unless pseudonyms are changed very often and in such a manner that the device stays silent for a relatively long period (e.g., 30 s) while making each change. On the other hand, a much more straightforward pseudonym change strategy is sufficient in cases where the user moves out of the coverage of the attacker network (and returns later).

3. *Usage Control*

Once sensitive information about user context has been given to a service provider, it is seldom transparent to the user how the service provider actually uses this information. One way to increase trust between the service provider and the service user is to apply usage control mechanisms on the collected data. An important ingredient of usage control is precise definitions of policies about data usage. The formal policies are typically converted from informal ones, written in natural language. Metric First-order temporal logic (MFOTL) has been successfully applied for this purpose (Basin et al. 2008). Another part of usage control are the mechanisms which monitor that the formal policies are also followed. Preferably, there should also be ways to enforce compliancy with the policies. All these mechanisms have been tried out with data from the trials mentioned in Sect. 9.6.4. The usage control service is often provided in a distributed manner, and in this case the usage monitoring should also be done in distributed manner as far as possible (Biswas et al. 2011).

4. *Context-Aware Policy Management*

Sometimes sensitive context information can be used to enhance user privacy. In Gupta et al. (2011), a mechanism is presented that utilizes sensor information from a mobile device, and helps the user by making automatic policy decisions about access control to the device. For instance, when the device infers from the context information that the user is probably in a potentially unsafe place (e.g., a bus stop), then the device is locked quickly and a PIN is required to open it. On the other hand, in a familiar context (e.g., in the office), the time-out is longer and the screen-saver may be removed by a single tap. Similar mechanisms can be used also for other policy management tasks.

9.7 Simulations, Prototypes, and Trials

9.7.1 A Prototype of MSN Chatting and User Study

1. *Prototype of MSN Chatting*

The system is implemented by extending the functions of the Twin application developed by Tampere University of Technology (Chen et al. 2011). The Twin application is an ad hoc chatting application based on the Nokia Instant Community platform – an energy-efficient and fully distributed pervasive social networking platform developed by the Nokia Research Center, Helsinki (Ahtiainen et al. 2009). We developed MSN nodes using Nokia N900 with Python and GTK binding. The MSN communications are based on wireless LAN. The TS is implemented with Apache and PHP in Linux platform (Ubuntu 9.04). The connection between the TS and nodes is also based on wireless LAN, which could also be extended to cellular networks (Chen 2010).

The implemented prototype system has three modules: mobile social chatting, reputation management and privacy/identity management. The system supports both node-to-node chatting and community chatting. Any user can create a community by indicating the community name and its importance (i.e., the crucial level of chatting topic) through the UI, as shown in Fig. 9.9. After creating a community, other people in the vicinity can find the community in their device and join the community chatting. AwareTrust allows on-chat voting and reputation visualization during chatting. Figure 9.10 shows a community chatting UI with personalized reputation visualization, and on-chat voting with comments (e.g., “You DOWN node 3: Too expensive” and “You UP node 3: Good”.). Particularly, an AwareTrust

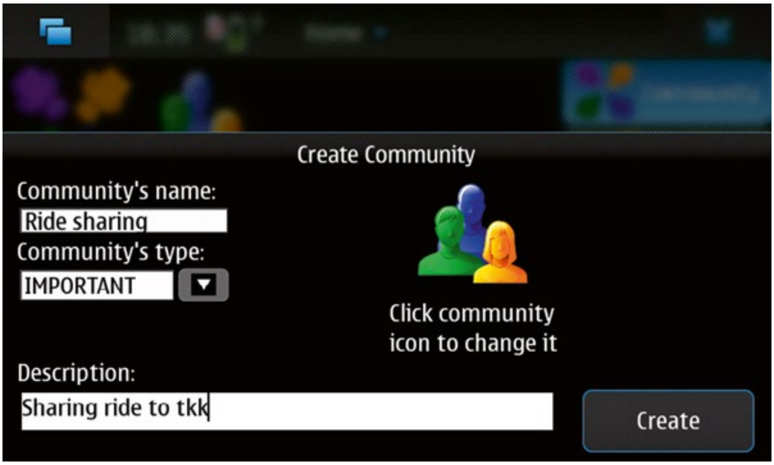


Fig. 9.9 Create a chatting community (Yan and Chen 2011)

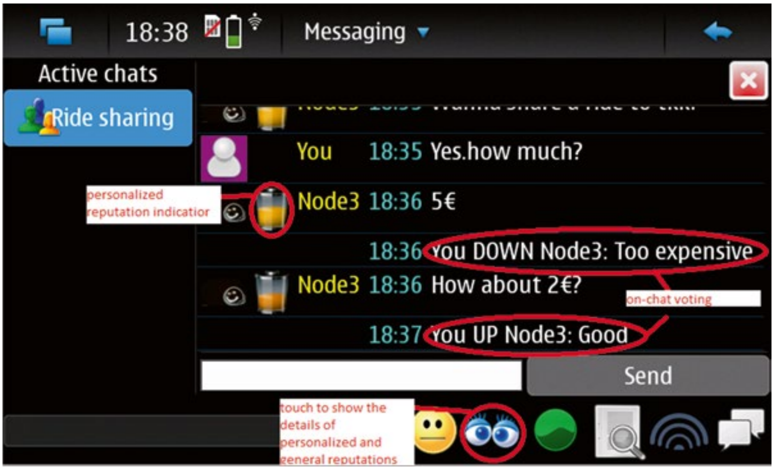


Fig. 9.10 Community chatting UI and on-chat voting (Yan and Chen 2011)

user can select a preferred visualization scheme and activate or deactivate it. In Fig. 9.10, a reputation visualization scheme is shown with battery volume (which can be personalized by the user). AwareTrust also provides detailed information of personalized and general reputations by touching the 'eyes' icon and the user photos in Fig. 9.10.

We performed a between-subject study to investigate the impacts of this prototype on mobile users. We selected two groups of participants from two student villages (each had seven persons). All participants were university students aged between 23 and 28 years old. Among seven participants in each group, three of them were female and four were male. They didn't know with each other. All of them had Internet chatting experiences, but none of them had experience of MSN chatting. Group 1 used this system during chatting, while group 2 did not (i.e., turned off reputation visualization).

2. *User Experiment*

We designed the experiment in a board game style in order to organize the study and make the results of two tests comparable. We asked the participants to simulate three chatting scenarios as described in Sect. 9.3. Before the experiment, each participant got a card that indicated his/her roles and tasks in chat. The participants tried to make a decision with regard to their chatting purpose. For each scenario, they chatted in a community. During the tests, the chatting information, such as chatting time, contents, length, on-chat voting, and afterwards voting were automatically logged for future analysis.

Additionally, we conducted an interview after the experiment to evaluate the perceived usefulness, perceived ease of use, interface, playfulness, and user attitude in terms of MSN chatting. Our interview was designed based on the technology acceptance model (TAM) and its extension, which indicates that usefulness, ease of use, and playfulness lead to user acceptance (Davis 1989; Venkatesh and Bala 2008). This theory also indicates that good interface leads to better perceived usefulness and ease of use; playfulness causes better acceptance (attitude). Finally, we randomly talked to some participants in order to get additional comments. After the test, each participant was awarded a movie ticket.

Investigating the chatting time and length, we observe that displaying reputation information could encourage participants to chat more and become more social (refer to chatting record length), and help them chat in a more efficient way (i.e., chatting time was shorter) than the situation without reputation visualization. We also note that participants became more serious and took a longer time to make a decision in a more crucial chatting scenario (e.g., Scenario 3 – car riding), when the reputation value was visualized (refer to chatting time).

The prototype system had satisfactory evaluation scores with regard to perceived ease of use, perceived usefulness, interface, playfulness, and user attitude. In terms of perceived ease of use, we note that visualizing reputation in MSN chatting made it easier for participants to select a person they like from many candidates during chatting than without reputation visualization. The result showed that the prototyped system is a very useful and interesting (playful) application that can aid user decisions in MSN chatting. Its UI (especially reputation visualization) got good feedback and comments from the participants. They liked using it. Based on the TAM, we conclude that this system was well accepted by the participants.

9.7.2 A Prototype of MSN Content Recommendation

We have implemented a prototype system using Nokia N810 tablets as the MSN nodes and an Apache server acting as the TS for MSN content recommendations. We developed the mobile node part on Nokia N810 using Python with GTK binding, and the server part on Linux (Ubuntu 9.04) together with Apache, MySql and PHP. MSN communication is based on Wireless LAN. There is no guarantee for MSN nodes to connect to TS. The implementation satisfies all requirements as specified in Sect. 9.1. We attempt to achieve efficient power consumption by controlling the message length of node communication within 100 bytes and applying an Awareness ad hoc networking platform developed by the Nokia Research Center (Ahtiainen et al. 2009). This is because the message length of node communications will greatly influence power consumption. The longer the length, the more power will be consumed (Ahtiainen et al. 2009). The prototype system provides essential security protection on node–server communications with Open SSL and node–node communications by utilizing a community symmetric key. The user interfaces of query and query response are shown in Figs. 9.11 and 9.12.

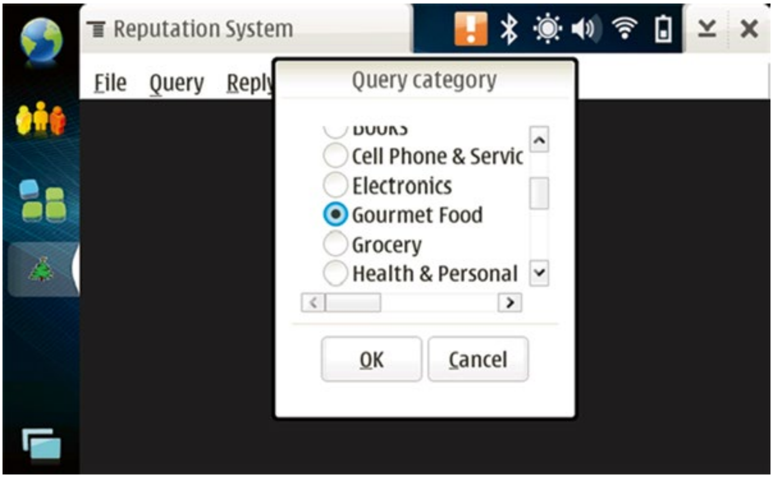


Fig. 9.11 User interface of node query for content recommendations (Yan and Chen 2010)

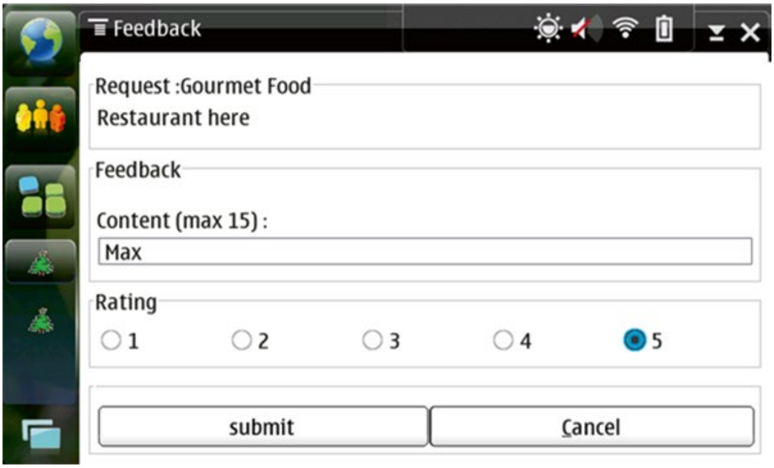


Fig. 9.12 User interface for query response for content recommendations (Yan and Chen 2010)

9.7.3 Simulations

We also conducted a series of simulations to evaluate the accuracy, efficiency, and robustness of the designed algorithms for MSN trust/reputation generation, trustworthy content recommendations, unwanted traffic control, and user privacy recommendations. Detailed simulation results are presented in (Yan and Chen 2010, 2011; Yan et al. 2012b; Yan and Zhang 2011).

9.7.4 Trials

Nokia Research Center has carried out two major trials with regard to MSN where emphasis was put on privacy aspects.

In the first trial, approximately 200 volunteers from the Lake Geneva region carried a mobile device with a tailor-made data collection client application. The campaign lasted for more than 1 year (starting in September 2009), and participants used the device as their primary mobile phone. Many types of sensory data were collected, including data from GPS, microphone, and wireless interfaces. Information about communications (calls, short messages) was also collected, but not the actual contents of communications. The main purpose was to learn about participants' socio-geographical behavior and provide a rich data set for researchers (Kiukkonen et al. 2010). Lots of collected data were privacy-sensitive, and several protection mechanisms were needed (Aad and Niemi 2010).

A carefully selected and post-processed data set was extracted from the whole data set. The data set was made available widely to researchers around the world, and a global mobile data challenge was held based on the data set. The best research results were presented in a specific workshop attached to the Pervasive 2012 conference; see Mobile Data Challenge Nokia (2012).

Another trial was based on Nokia Instant Community technology. This is a platform built by Nokia Research Center which can be used for various services based on local communications. The implementation of the platform used in the trial was based on WLAN ad hoc mode, but it is possible to realize the concept on top of other local radio technologies as well. In the spring of 2011, around 100 EPFL students and researchers were using a device including the Nokia Instant Community platform and several applications on top of it. The trial lasted for 3 months, during which data about several privacy features were collected.

One of the main results was related to the use of pseudonyms in ad hoc networks. The trial included an observing network, i.e., an infrastructure of a few dozen installed fixed WLAN access points. The role of this network in the trial was to act as an adversary that tried to track the users, who continuously changed their pseudonyms using state-of-the-art techniques. It was found out that an adversary of that scale would be able to track the users quite well based on their movement patterns, regardless of the ever-changing pseudonyms; see Bindschaedler et al. (2012).

Another interesting finding was that it was possible to deduce social relationships between the participants, based on the data about their movements only; see Bilogrevic et al. (2012).

9.8 Further Discussions

Social networking nowadays is changing the style of people's social life. With the rapid growth and usage of mobile devices, mobile social networking will play an important role of social activities in the future. Trust, security, and privacy have become extremely important and relevant in mobile social networks. This section further discusses additional issues related to MSN trust, security, and privacy.

9.8.1 Additional Issues

We try to answer the following questions based on our research experiences. Our opinions shown below only represent the personal points of view of the authors.

1. How can we infer trust based on social relationships?

   Mobile social networking has the characteristic of being pervasive. People who are mostly strangers would like to communicate in a pervasive and private way in order to pursue instant reciprocal benefits. In this case, it is hard to infer trust based on existing social relationships. On the other hand, mobile social networking can operate as an important complement for Internet social networks by extending our daily feelings and relationships with strangers that we may regularly observe but do not interact with in public places. In addition, strangers could become acquaintances later on with the support of mobile social networking. In this case, AwareTrust could be applied to share MSN experiences by reporting MSN activities to TS, and then it could be possible to plug in new social relationships into Internet social networking if users would like to.
2. Can privacy be accomplished based on social groups so that only those people who you were with within the mobile social network can access that content?

   It is hard to ensure user data privacy based on social groups. It is obvious that any malicious social group members could disclose confidential contents shared within a social group. Thus, additional control should be provided. Cryptographic solutions will play an important role in solving this issue. AwareTrust offers flexible MSN data access based on two-dimensional trust levels using attribute-based encryption. Thus, it can protect user data privacy in a flexible manner.
3. Are the issues in MSN different than in regular online social networks, or are they extensions?

   The issues of trust, security, and privacy in MSN are different from those in regular online Internet social networks. The reason is obvious, since MSN is based on self-organized mobile networks, not Internet or mobile Internet. The issues related to trust, security, and privacy should be mostly solved in a distributed manner, not a centralized way. The framework for trust, security, and privacy should follow the system structure of MSN, supporting the mobility of each MSN node. AwareTrust adopts a hybrid trust management framework that can support MSN trust in a distributed way with the support of a centralized trusted server. It is also compatible and easy to merge into Internet social networking.
4. Do we need new models for this?

   New models are needed to support the new system structure of MSN, which is obviously different from the traditional Internet-based online social networks. This could open a new research area of trust, security, and privacy for mobile social networking.

9.8.2 *Future Research Trends*

This chapter has only discussed some issues of trust, security, and privacy, and provided a limited number of solutions based on our past work. To move further towards real trustworthy mobile social networking, more issues and challenges should be considered and overcome. Herein, we simply list a number of challenges for the readers who are interested in the research in this area:

- Efficient private data preservation in a distributed manner in MSN
- The merging of MSN with online social networks with privacy preservation
- Sound user experience support for trustworthy MSN
- The extension of MSN for trustworthy physical social networking based on MSN experiences

9.9 Conclusions

This chapter has discussed the issues involved in moving towards trustworthy mobile social networking. Based on the literature review, we proposed AwareTrust, a trust management framework that supports context-aware trust/reputation generation, trustworthy content recommendations, secure communications, unwanted traffic control, user privacy recommendation and preservation, and other trust and privacy technologies. Simulations, prototype implementation, and trial experiments further prove the effectiveness of our proposed solutions. Obviously, mobile social networking is a new application area that acts as a significant complement for Internet social networking. In MSN, trust, security, and privacy are the most crucial issues that will influence the success of MSN. Nowadays, academia and industry have initiated a number of research and development activities in this area. We believe more work will be conducted in order to achieve trustworthy mobile social networking and lead to its real success.

Acknowledgments This work is sponsored by the Fundamental Research Funds for the Central Universities with the grant number K5051201032 and the initial grant of Chinese Educational Ministry for researchers returned from abroad. The authors would like to thank Nokia Research Center. Part of the work for this chapter was conducted and sponsored by Nokia Research Center, Helsinki and Nokia Research Center, Lausanne.

References

Aad, I., & Niemi, V. (2010). NRC data collection and privacy by design principles. In *Proceedings of PhoneSense* (pp. 41–45). Zurich, Switzerland.

Adler, BT., & Alfaro, L. (2007). *A content-driven reputation system for the Wikipedia*. In WWW (pp. 261–270). Banff, Alberta, Canada.

Ahn, J., & Amatriain, X. (2010). Towards fully distributed and privacy-preserving recommendations via expert collaborative filtering and RESTful linked data. In *WI-IAT'10* (pp. 66–73). Toronto, Canada.

Ahtiainen, A., et al. (2009). Awareness networking in wireless environments: Means of exchanging information. *IEEE Vehicular Technology Magazine, 4*(3), 48–54.

Bansal, S., & Baker, M. (2003). Observation-based cooperation enforcement in ad hoc networks (Technical Report No. NI/0307012). Stanford University, CA, USA.

Basin, D., Klaedtke, F., Muller, S., Pfitzmann, B. (2008). Runtime monitoring of metric first-order temporal properties. In *Proceedings of FSTTCS* (pp. 49–60), Bangalore: India.

Beach, A., Gartrell, M., Han, R. (2009). Solutions to security and privacy issues in mobile social networking. In *International conference on computational science and engineering, 2009, CSE '09* (Vol. 4, pp. 1036–1042). Vancouver.

Becchetti, L., Castillo, C., Donato, D. et al. (2006). Using rank propagation and probabilistic counting for link-based spam detection. In *Proceedings of WebKDD*. Philadelphia, August 20, 2006.

Berkovsky, S. (2010). Putting things in context: Challenge on context-aware movie recommendation. In *Proceedings of the workshop on context-aware movie recommendation* (pp 2–6). Barcelona.

Bethencourt, J., Sahai, A., Waters, B. (2007). Ciphertext-policy attribute-based encryption. In *Proceedings of the 2007 IEEE symposium on security and privacy* (pp. 321–334). Oakland.

Bi, J., Wu, J., Zhang, W. (2008). A trust and reputation based anti-spam method. *IEEE INFOCOM,* 2485–2493.

Bilge, A., & Polat, H. (2010). Improving privacy-preserving NBC-based recommendations by preprocessing. WI-IAT (pp. 143–147). Toronto.

Bilogrevic, I., Jadliwala, M., Lam, I., et al. (2012). Big Brother knows your friends: On privacy of social communities in pervasive networks. In *Proceedings of the 10th International Conference on Pervasive Computing*. Newcastle, UK

Bindschaedler, L., Jadliwala, M., Bilogrevic, I., Aad, I., Ginzboorg, P., Niemi, V., Hubaux, J.P. (2012). Track me if you can: On the effectiveness of context-based identifier changes in deployed mobile networks. In *Proceedings of the 19th Annual Network & Distributed System Security Symposium (NDSS*). Hilton San Diego Resort & Spa, US.

Biswas, D., Nefedov, N., Niemi, V. (2011). Distributed usage control the 8th International Conference on Mobile Web Information Systems (MobiWIS 2011). In *Proceedings. of MobiWIS. Procedia Computer Science 5*, 562–569. Niagara Falls, Ontario, Canada

Blaze, M., Feigenbaum, J., Lacy, J. (1996). Decentralized trust management. In *Proceedings. of the IEEE Conference on Security and Privacy*. (pp. 164–173). Oakland, CA

Buchegger, S., & Boudec, J.L. (2002). Performance analysis of the CONFIDANT protocol. In *Proceedings of the ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)* Lausanne, Switzerland. (pp. 226–236).

Buchegger, S., & Boudec, J.Y.L. (2003). The effect of rumor spreading in reputation systems for mobile ad-hoc networks. In *Proceedings of WiOpt Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*. Sophia-Antipolis, France.

Chen, Y. (2010). *A privacy-enhanced reputation system for mobile ad hoc services*. Master thesis, Aalto University, Espoo, Finland.

Chen, G., & Rahman, F. (2008). Analyzing privacy designs of mobile social networking applications. In *IEEE/IFIP International Conference on Embedded and Ubiquitous Computing, EUC'08: Shanghai, China. Vol. 2.* (pp. 83–88).

Chen, Y., Yan, Z., Niemi, V. (2011). Implementation of a reputation system for pervasive social networking. In *IEEE TrustID 2011* (pp. 857–862). Changsha, China.

Chuong, C., Torabi, T., Loke, S.W. (2009). Towards context-aware task recommendation. In *JCPC'09* (pp. 289–292). Taiwan.

Corritore, C. L., Kracher, B., & Wiedenbeck, S. (2003). On-line trust: Concepts, evolving themes, a model. *International Journal of Human–Computer Studies, Trust and Technology, 58*(6), 737–758.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340.

Douceur, J.R. (2002). The Sybil attack. In *IPTPS, LNCS 2429* (pp. 251–260). Cambridge, MA, USA.

EZSetup. (2012). http://research.microsoft.com/en-us/groups/wn/mssn.aspx. Accessed 20 June 2012.

Goyal, V., Pandey, O., Sahai, A., et al. (2006). Attribute-based encryption for fine-grained access control of encrypted data. In *Proceedings of the 13th ACM Conference on Computer and Communications Security* (pp. 89–98). Alexandria, VA, USA.

Gupta, M., Judge, P., Ammar, M. (2003). A reputation system for peer-to-peer networks. In *Proceedings of the 13th international workshop on network and operating systems support for digital audio and video (NOSSDAV '03)*, Monterey (pp. 144–152). New York: ACM.

Gupta, A., Miettinen, M., Asokan, N. (2011). Using context-profiling to aid access control decisions in mobile devices. In *Proceedings of PerCom Workshops* (pp. 310–312). Seattle, WA, USA.

Hancock, J.T., Toma, C., Ellison, N. (2007). The truth about lying in online dating profiles. *Proceedings of the ACM CHI 2007* (pp. 449–452). San Jose, California, USA.

Hu, J., & Burmester, M. (2006). LARS: A locally aware reputation system for mobile ad hoc networks. In *Proceedings of the 44th ACM Annual Southeast Regional Conference* (pp. 119–123). Melbourne, FL, USA.

Hyytia, E., Virtamo, J., Lassila, P., Kangasharju, J., Ott, J. (2011). When does content float? Characterizing availability of anchored information in opportunistic content sharing. In *Proceedings of IEEE INFOCOM* (pp. 3137–3145). Shanghai, China.

Jadliwala, M., Freudiger, J., Aad, I., Hubaux, J.P., Niemi, V. (2011). Privacy-triggered communications in pervasive social networks. In *Proceedings of IEEE WoWMoM Workshop on Autonomic and Opportunistic Communications* (pp. 1–6). Lucca, Italy.

Jøsang, A., Ismail, R., & Boyd, C. (2007). A survey of trust and reputation systems for online service provision. *Decision Support Systems, 43*(2), 618–644. Elsevier.

Junction. (2012). *Harvard MobiSocial Group*. http://openjunction.org/. Accessed 20 June 2012.

Katzenbeisser, S., & Petkovic, M. (2008). Privacy-preserving recommendation systems for consumer healthcare services. In *ARES'08* (pp. 889–895). Barcelona , Spain.

Kikuchi, H. (2009). Privacy-preserving collaborative filtering schemes. In *ARES'09* (pp. 911–916). Fukuoka, Japan.

Kiukkonen, N., Blom, J., Dousse, O., Gatica-Perez, D., Laurila, J. (2010). Towards rich mobile phone datasets: Lausanne data collection campaign. In *Proceedings of ICPS*. Berlin, Germany.

Kolan, P., & Dantu, R. (2007). Socio-technical defense against voice spamming. *ACM Transactions on Autonomous and Adaptive Systems*, 2(1), Article 2(44).

Krishna, P., Puttaswamy, N., Zhao, B. (2010). Preserving privacy in location-based mobile social applications. In *Proceedings of the Eleventh Workshop on Mobile Computing Systems & Applications (HotMobile '10)* Annapolis, Maryland, US (pp. 1–6). New York: ACM.

Kujimura, K., & Nishihara, T. (2003). Reputation rating system based on past behavior of evaluators. In *Proceedings of the 4th ACM Conference on Electronic Commerce* (pp. 246–247). San Diego, CA, USA.

Li, J., Li, R., Kato, J., et al. (2008). Future trust management framework for mobile ad hoc networks. *IEEE Communications Magazine, 46*(4), 108–115.

Li, T., Gao, C., Du, J. (2009). A NMF-based privacy-preserving recommendation algorithm. In *ICISE'09* (pp. 754–757). Nanjing, China.

Liiv, I., Tammet, T., Ruotsalo, T. et al. (2009). A: Personalized context-aware recommendations in SMARTMUSEUM: Combining semantics with statistics. In *SEMAPRO '09* (pp. 50–55). Sliema, Malta.

Liu, Z., Yau, S.S., Peng, D. et al. (2008a). A flexible trust model for distributed service infrastructures. In *Proceedings of 11th IEEE Symposium on Object Oriented Real-Time Distributed Computing* (pp. 108–115). Orlando, Florida, USA.

Liu, D., Meng, X., Chen, J. (2008b). A framework for context-aware service recommendation. In *ICACT 2008* (pp. 2131–2134). Gangwon-Do, Korea (South).

Luo, Y., Le, J., Chen, H. (2009). A privacy-preserving book recommendation model based on multi-agent. *WCSE '09* (pp. 323–327). Xiamen, China.

Michiardi, P., & Molva, R. (2002). Core: A COllaborative REputation mechanism to enforce node cooperation in mobile ad hoc networks. In *Advanced Communications and Multimedia Security, LNCS 2828* (pp. 107–121). Portorož, Slovenia.

MicroBlog. (2012). *SyNRG*, Duke University. http://synrg.ee.duke.edu/microblog.html. Accessed 20 June 2012.

Miluzzo, E., Lane, N., Fodor, K. et al. (2008). 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 (SenSys '08)* (pp. 337–350). Raleigh, NC, USA.

Muller, S., Katzenbeisser, S., Eckert, C. (2008). Distributed attribute-based encryption. In *Proceedings of the 11th Annual International Conference on Information Security and Cryptology* (pp. 20–36). Seoul, Korea.

Nokia Instant Community. (2010a). https://lausanne.nokiaresearch.com/nic/. Accessed 20 June 2012.

Nokia Instant Community. (2010b). http://conversations.nokia.com/2010/05/25/nokia-instant-community-gets-you-social/. Accessed 20 June 2012.

Nokia. (2012). Mobile Data Challenge. *Workshop in Pervasive 2012*. Newcastle, UK. http://pervasiveconference.org. Accessed August 31, 2013.

O'Donovan, J., & Smyth, B. (2006). Trust in recommender systems. In *IUI'05* (pp. 167–174).

Ott, J., Hyytiä, E., Lassila, P. E., Kangasharju, J., & Santra, S. (2011). Floating content for probabilistic information sharing. *Pervasive and Mobile Computing, 7*(6), 671–689.

Polat, H., & Du. W.L. (2005). Privacy-preserving top-n recommendation on horizontally partitioned data. In *The 2005 IEEE/WIC/ACM International Conference on Web Intelligence* Compiegne University of Technology, France. (pp. 725–731).

Paulos, E. and Goodman, G. (2002). http://www.paulos.net/research/intel/familiarstranger/index.htm. Accessed 20 June 2012.

Raya, M., Papadimitratos, P., Gligory, VD. et al. (2008). On data-centric trust establishment in ephemeral ad hoc networks. In *IEEE INFOCOM*, Phoenix, AZ, US. 1912–1920.

Resnick, P., & Varian, H. R. (1997). Recommender systems. *Communications of the ACM, 40*(3), 56–58.

Resnick, P., & Zeckhauser, R. (2002). Trust among strangers in Internet transactions: Empirical analysis of eBay's reputation system. In M. Baye (Ed.), *The economics of the Internet and e-commerce* (Advances in applied microeconomics, Vol. 11, pp. 127–157). Amsterdam: Elsevier.

Resnick, P., Kuwabara, K., Zeckhauser, R., et al. (2000). Reputation systems. *Communications of the ACM, 43*(12), 45–48.

Sadeh, N., Hong, J., Cranor, L., et al. (2009). Understanding and capturing people's privacy policies in a mobile social networking application. *Personal and Ubiquitous Computing, 13*(6), 401–412.

Sahai, A., & Waters, B. (2005). Fuzzy identity-based encryption. In *Proceedings of 24th International Conference on the Theory and Application of Cryptographic Techniques* (pp. 457–473). Aarhus, Denmark.

Seetharam, A., & Ramakrishnan, R. (2008). A context-sensitive, yet private experience towards a contextually apt recommendation of service. In *IMSAA 2008* (pp. 1–6). Bangalore.

Song, S., Hwang, K., Zhou, R., et al. (2005). Trusted P2P transactions with fuzzy reputation aggregation. *IEEE Internet Computing, 9*(6), 24–34.

Stuedi, P., et al. (2008). *Demo abstract ad hoc social networking using MAND*. http://www.iks.inf.ethz.ch/publications/files/mobicom08_demo.pdf, Accessed 25 June 2010

Su, X., & Khoshgoftaar, T.M. (2009). A survey of collaborative filtering techniques. *Advances in Artificial Intelligence* 19p. doi:10.1155/2009/421425.

Sun, Y., Yu, W., Han, Z., et al. (2006a). Information theoretic framework of trust modeling and evaluation for ad hoc networks. *IEEE Journal on Selected Area in Communications, 24*(2), 305–317.

Sun, Y., Han, Y.Z., Yu, W. et al. (2006b). A trust evaluation framework in distributed networks: Vulnerability analysis and defense against attacks. *IEEE INFOCOM* (pp. 1–13). Barcelona, Spain.

Sun, Y., Han, Z., & Liu, K. J. (2008). Defense of trust management vulnerabilities in distributed networks. *IEEE Communications Magazine, 46*(2), 112–119.

Tada, M., Kikuchi, H., Puntheeranurak, S. (2010). Privacy-preserving collaborative filtering protocol based on similarity between items. *AINA'10* (pp. 573–578). Perth, Australia.

Tang, Y., Krasser, S., He, Y. et al. (2008). Support vector machines and random forests modeling for spam senders behavior analysis. In *IEEE GLOBECOM* (pp. 1–5). New Orleans, LA, US.

Theodorakopoulos, G., & Baras, J. S. (2006). On trust models and trust evaluation metrics for ad hoc networks. *IEEE Journal on Selected Areas in Communications, 24*(2), 318–332.

Trifunovic, S., Legendre, F., Anastasiades, C. (2010). Social trust in opportunistic networks. *IEEE INFOCOM Workshops,* 1–6. San Diego, CA, US.

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences, 39*(2), 273–315.

Walsh, K., & Sirer, E.G. (2005) Fighting peer-to-peer SPAM and decoys with object reputation. In *Proceedings of P2PECON* (pp. 138–143). Philadelphia, Pennsylvania, USA.

Wang, J. (2010). Mining context-related sequential patterns for recommendation systems. In *CAMP'10* (pp. 270–275). Shah Alam, Malaysia.

Wang, G., Liu, Q., Wu, J. (2010). Hierarchical attribute-based encryption for fine-grained access control in cloud storage services. In *Proceedings of the 17th ACM Conference on Computer and Communications Security (Poster)* (pp. 735–737). Hyatt Regency Chicago, Chicago, IL, USA.

Wu, B. et al. (2006). Topical TrustRank: Using topicality to combat web spam. In *WWW '06: Proceedings of the 15th international conference on World Wide Web* (pp. 63–72). Edinburgh, Scortland.

Xiao, H. (2010). An approach for context-aware service discovery and recommendation. In *IEEE ICWS'10* (pp. 163–170). Miami, Florida, USA.

Xiong, L., & Liu, L. (2004). PeerTrust: Supporting reputation-based trust for peer-to-peer electronic communities. *IEEE Transactions on Knowledge and Data Engineering, 16*(7), 843–857.

Yan, Z. (Ed.). (2010). *Trust modeling and management in digital environments: From social concept to system development*. Hershey: IGI Global.

Yan, Z., & Chen, Y. (2010). AdContRep: A privacy-enhanced reputation system for MANET content services. In *UIC 2010, LNCS 6406*, (pp. 414–429). Xi'an, China.

Yan, Z., & Chen, Y. (2011). AdChatRep: A reputation system for MANET chatting. In *SCI2011 in UbiComp2011* (pp. 43–48). Beijing, China.

Yan, Z., & Holtmanns, S. (2008). Trust modeling and management: From social trust to digital trust. In R. Subramanian (Ed.), *Computer security, privacy and politics: Current issues, challenges, and solutions* (pp. 290–323). Hershey: Idea Group.

Yan, Z., & Niemi, V. (2009). A methodology towards usable trust management. In *Proceedings of ATC'09, LNCS 5586* (pp. 179–193). Brisbane, Australia.

Yan, Z., & Zhang, P. (2011). AdPriRec: A context-aware recommender system for user privacy in MANET services. In *UIC2011, LNCS 6905* (pp. 295–309). Banff, Canada.

Yan, Z., Liu, C., Niemi, V., Yu, G. (2009). *Trust information indication: Effects of displaying trust information on mobile application usage* (Technical Report No. NRC-TR-2009-004). Nokia Research Center. http://research.nokia.com/files/NRCTR2009004.pdf.

Yan, Z., Zhang, P., & Deng, R. H. (2012a). TruBeRepec: A trust-behavior-based reputation and recommender system for mobile applications. *Journal of Personal and Ubiquitous Computing, 16*(5), 485–506. Springer.

Yan, Z., Kantola, R., Shen, Y. (2012b). Unwanted traffic control via hybrid trust management. In *IEEE TrustCom 2012*. Liverpool, UK.

Yang, Y., Sun, Y., Kay, S. et al. (2002). Defending online reputation systems against collaborative unfair raters through signal modeling and trust. In *SAC'09, 2009* (pp. 1308–1315). Hawaii, USA.

Yap, G. (2007). Discovering and exploiting causal dependencies for robust mobile context-aware recommenders. *IEEE Transactions on Knowledge and Data Engineering, 19*(7), 977–992.

Yu, Z., Zhou, X., Zhang, D., et al. (2006). Supporting context-aware media recommendations for smart phones. *IEEE Pervasive Computing, 5*(3), 68–75.

Zhang, J. (2010). MailTrust: A mail reputation mechanism based on improved TrustGuard. In *International Conference on Communications and Mobile Computing (CMC)* (pp. 218–222). Shenzhen, China.

Zhang, D., & Yu, Z. (2007). Spontaneous and context-aware media recommendation in heterogeneous spaces. In *IEEE VTC'07* (pp. 267–271). Baltimore, MD, USA.

Zhang, H., Duan, H., Liu, W. et al. (2009a). IPGroupRep: A novel reputation-based system for anti-spam. In *Symposia and Workshops on Ubiquitous, Autonomic and Trusted Computing* (pp. 513–518). Brisbane, QLD.

Zhang, X., Han, B., Liang, W. (2009b). Automatic seed set expansion for trust propagation based on anti-spamming algorithms. In *WIDM'09, Proceeding of the Eleventh International Workshop on Web Information and Data Management* (pp. 31–38). HongKong, China.

Zouridaki, C., Mark, B.L., Hejmo, M. et al. (2006). Robust cooperative trust establishment for MANETs. In *SASN '06: Proceedings of the fourth ACM Workshop on Security of Ad Hoc and Sensor Networks* (pp. 23–24). Alexandria, VA, USA.

Chapter 10
Conclusions

Alvin Chin and Daqing Zhang

Abstract Mobile social networking is the next chapter in social networking. It bridges and fuses offline and online together to create one homogeneous social network. This chapter summarizes the objectives of MSNs, provides our vision about future MSNs, and identifies the new challenges and opportunities ahead.

Mobile social networking is the next evolution of social networking. It bridges the offline and online and fuses them together, which is accessible via mobile devices. Mobile social networking goes beyond online social networking and mobile ad hoc networking, thus bringing new challenges and opportunities. This book has been written with the following objectives:

1. Identify the problems that exist in mobile social networking and propose possible solutions
   Each chapter in the book addresses a different topic or perspective of mobile social networking, and with it, a different set of problems. For capturing offline social networks, Chap. 2 addresses the main research challenges pertaining to the capture, processing, and identification of the sensing data from smart phones, and proposes a social aware computing framework, while Chap. 3 examines the problem of opportunistic social networking, that is, how to capture the missed opportunities of social interaction, where the authors propose the concept of an ephemeral social network. Chapter 4 addresses the problem of characterizing user behavior, and the authors present two real-life applications using

A. Chin (✉)
Xpress Internet Services, Nokia, Beijing, 100176, China
e-mail: alvin.chin@nokia.com; ubiquitousdude@gmail.com

D. Zhang
Institut Mines-Telecom/Telecom SudParis, 9, rue Charles Fourier,
91011 Evry Cedex, France
e-mail: daqing.zhang@it-sudparis.eu

A. Chin and D. Zhang (eds.), *Mobile Social Networking: An Innovative Approach*,
Computational Social Sciences, DOI 10.1007/978-1-4614-8579-7_10,

face-to-face contacts to demonstrate how to characterize roles, links, and communities as user behavior. In Chap. 5, the authors address the design of mobile social services, and propose a design framework for creating mobile social services using two examples, a large-scale exhibition and a local group buying application. Chapter 6 addresses the classification of context, and how context is used to define mobile social networking. The authors present personal and community context, and create taxonomies from these two contexts to explore how they could be used in the different phases of the mobile social networking lifecycle. In Chap. 7, the authors identify the problem of the gap between the current Web and smart phone application-based social networking services and the next generation of pervasive computing services. The authors use the SOCIETIES project to help bridge this gap by explaining the concepts and architecture of their system. Chapter 8 addresses human mobile behavior in popular location-based social networks by using machine learning and data analytic techniques to describe its distinct properties, and presents two illustrative examples to show the application of data mining to real-world location-based social networks. Finally, in Chap. 9, the authors address how to build up trust in mobile social networking by proposing a trust management framework that supports context-aware trust/reputation generation, trustworthy content recommendations, secure communications, unwanted traffic control, user privacy recommendation and preservation, and other trust and privacy enhancement technologies.

2. Provide examples of real-life applications that illustrate different aspects of mobile social networking
   Each chapter in our book has examples of real-life applications that the authors have designed and implemented. For example in Chap. 2, a campus application was developed to help students to find a suitable place to study and locate his/her friends based on Wi-Fi positioning technology. In Chaps. 3 and 4, a conference application for finding and connecting attendees through proximity encounters was developed, while in Chap. 5 a group buying application for products on mobile phones was created. Finally, in Chap. 9 a trust management application for mobile social networks was presented.
3. Demonstrate real-life data extracted from deploying the applications in the field
   For many of the chapters that the authors created their applications, they also recorded the real-life data and conducted experiments and measurements to analyze this data and interpret it. In Chap. 3, data that was collected from the conference application included the sessions that the conference attendees went to (from their positioning), which people that they encountered or were nearby with, other attendees that they added as contacts with, papers that they viewed and shared with others, and messages that they communicated with others. Social network analysis was done to understand the attendees' behavior and to determine the effect that offline had on online and how ephemeral social networks were created. In Chap. 4, real-life data was obtained from the attendees' behavior in conferences in order to characterize the communities and roles of the

attendees such as Ph.D. students, session chairs, paper authors, and conference committee members. In Chap. 5, data was collected from two user studies, a large-scale exhibition and a group purchasing application through surveys and application usage. In Chap. 8, data was obtained from existing location-based social networking applications, and analysis was done on the user's social–historical ties in check-in behavior for location prediction, and to address the "cold-start" check-in problem. In Chap. 9, data was extracted from sensors in an application including volunteers' communications, and from another application that recorded volunteers' privacy features.

4. Challenge the widely accepted notion of what mobile social networking is within the industry and academic fields
   Many people view mobile social networking as just accessing and connecting to an online social network by using their mobile phones, such as reading and posting to Facebook directly from their phones. However, in this book, we challenge that notion, and have presented chapters where mobile social networking involves recording the offline activity through sensors (Chaps. 2, 6, and 7) and extending social networking to other areas besides the well-known areas of business meetings but into conference environments (Chaps. 3 and 4), campus environments (Chap 2), public and private places (Chaps. 5 and 9), and ad hoc environments where activities have not been organized in advance (Chaps. 3, 7, and 9).

Even though we have presented a foundation for mobile social networking, there is still much yet to be done. What are the next steps in mobile social networking? What is the research agenda? What challenges have still not been addressed? What are some of the opportunities in this area? How can researchers from academia and industry work together to achieve the common goal of mobile social networking? We elaborate on some of these issues below.

10.1 Next Generation of Mobile Social Networking

Social networks (SNs) have revolutionized the way people communicate and interact in the last decade. The impact and popularity of SNs has demonstrated the power of instantaneous multimedia communication among people independent of the distances among them, as well as the importance of enabling one-to-one, one-to-group, and broadcast communication using multimedia within the same framework. Leveraging the enormous reach of mobile phones equipped with various sensors, the pervasiveness of wireless communication infrastructure and sensor networks, the next generation of social networks is expected not only to *facilitate communication and interaction among people* in a more intelligent and effective manner, but to have the potential of *matchmaking the supply and demand among people* in terms of *information, services*, and *physical goods* (Jain and Sonnen 2011).

In order to achieve the vision of the next generation of mobile social networks (MSNs), the existing mobile social networks should be enhanced from the following perspectives:

1. Extending the sensing capability of MSN by linking the MSN to the Internet of Things (IoT)
   By leveraging the sensors embedded in the mobile phones, the devices, and sensor network installed in our surroundings, the users' home pages, and digital footprints left in the virtual space (Zhang et al. 2011), huge amounts of context information about the MSN users and interaction space can be acquired to enhance the social awareness, ambient awareness, situation awareness, etc. of the MSN. By enabling context-awareness in MSN, more intelligent interaction, information sharing, resource matchmaking among mobile users can be supported in next-generation MSNs.
2. Extending the communication capability of MSN by seamlessly bridging the mobile ad hoc network to the infrastructure-based network
   Nowadays, millions of people worldwide use social network services to share information, build relationships, engage with others, buy and sell products, and support interactions mainly in virtual space; however, those social network services lack effective support for people's face-to-face interaction, especially when there is no infrastructure support available. This calls for research into offline social network creation, management, and migration to the online social networks, as well as seamless transition between offline and online social networks. In other words, future mobile social networks should be infrastructure- independent, supporting spontaneous social interaction as well as long-term relationships. In addition, future MSNs should be transparent to communication networks, protocols, and media.
3. Extending the computing capability of MSN by connecting to the cloud.
   As mobile devices have limited computing power and storage, users might request storage and processing for the huge amount of data generated by the MSN. In this regard, the MSN itself might not possess the capability to handle everything locally; therefore, a viable way is to tap into the capability of the cloud for storage and computation.
4. Extending the service platform of MSN by adding the capability of context-aware service discovery, generalized service matchmaking between service producers and consumers, and open development tools for MSN services
   Generally speaking, the MSN is a service platform to facilitate user interactions, information sharing, service discovery and usage, and matchmaking of user's personal desires. In order to make MSN a convenient and effective service platform for mobile users, it is highly desirable to enhance each key feature of the platform. For example, when users are looking for services, events, activities, objects, or information, context-aware discovery is a nice feature. In addition to context-aware discovery, matchmaking of different kinds of producing and consuming entities in terms of voice, text, video is crucial for the economy, because people can use MSN as a trading platform among users directly. Last but not

least, the MSN should provide an open platform that lets developers freely create new applications to address specific needs of different users.

Based on the above requirements, we need to work in the following directions:

- Support mobile access in a heterogeneous network
- Infer context and detect situations from heterogeneous data sources
- Enable large scale, heterogeneous, and multi-modal data fast indexing and query using event, activity, time, and place
- Support multi-modality interaction according to user capability and context
- Provide an open platform for application development
- Facilitate matchmaking between user supply and request of information, services, and goods
- Provide mechanisms and incentives to motivate users to contribute resources, services, and information to social networks
- In the future, we wish to bring MSN to people's daily lives for more effective communication and interaction; link MSN to the Internet of Things to extend the entities from people to objects; use MSN as a platform to bridge the physical and virtual worlds, in both distant and face-to-face communication; and use MSN to facilitate the supply and demand of not only information and services, but also goods.

10.2 Research Challenges Ahead

In this book, we have addressed several research issues; however, there are still many other research challenges yet to be addressed. First, there are challenges related to privacy. There is a need to balance between the benefit derived from data sharing and the risk which results from this. Robust models and mechanisms are needed to safeguard user privacy during the sharing and usage of sensing data. Second, there are also challenges regarding how to apply the technologies already developed for one domain to another domain. In Chap. 2, mobile social networking applications for education were discussed, but what was not discussed was how this framework could be applied to other applications, such as public health, urban transportation management, and environment monitoring. Third, there are challenges regarding detecting and inferring user behavior. From the sensing technologies, Chaps. 3 and 4 addressed how we can detect proximity or ephemeral social networks using RFID and Wi-Fi, but how can we do it accurately? How can we create a user model for representing online to offline and vice versa? How can we use pure peer-to-peer technologies for implementing proximity and ephemeral social networks without the need for server infrastructure?

Fourth, mobile social service design also poses a research challenge; since this is new, there is no agreed upon standard for this user-centered design. Chapter 5 addressed some of the design challenges in terms of scenario, activity, deployment environment, and trial use and support when designing a mobile social service.

Fifth, there are challenges in enabling future context-aware mobile social networks, as discussed in Chap. 6. These include challenges in data itself such as data quality which differs according to the source it is obtained from (device or infrastructure), data collection and context identification, the uncertainty of data, heterogeneous data management that comes from different sources, data fusion from independent sensing sources, and data visualization. Sixth, there are also challenges involved in security, privacy, and user control. Seventh, in terms of community challenges, these include community management, extracting community preferences, mining the underlying structure of MSNs, identifying and storing historical and live contexts in ad hoc or long-term communities, as well as social concerns, both positive such as friendships and negative such as social selfishness. Eighth, once a framework has been developed, a challenge is how to deploy and commercialize the framework into a product that can be exploited by businesses and the public, and also form an open source community around its development. These are the challenges that Chap. 7 has identified.

Ninth, with regards to location-based social networks (LBSNs) in Chap. 8, there are certain challenges dealing with how to better utilize social network information on LBSNs, how to handle the check-in sparseness of LBSNs, and how to efficiently make use of user-generated content on LBSNs. Tenth, with regards to security and trust, there still exist many challenges, as posed in Chap 9. How to ensure efficient private data preservation in a distributed manner in MSN? How to merge MSN with online social networks with privacy preservation? How to have sound user experience support for trustworthy MSN? Finally, how can MSN be extended for trustworthy physical social networking based on MSN experiences?

We have still not addressed all the issues in mobile social networking. Some particular issues that have not been explained in this book include energy efficient solutions, new ways of HCI, future networking standards, etc. Therefore, the area of mobile social networking provides many opportunities to tackle these challenges, especially in application domains that have not been well explored, such as healthcare, elderly care, transportation, security, food, water, and goods trading. These challenges can be addressed through event reporting, crowd-sourcing, citizen computing, and information sharing and aggregation.

10.3 How Researchers from Industry and Academia Can Work Together to Achieve the MSN Vision

This book intends to be the catalyst for mobile social networking research, and we hope that the readers prescribe to our vision of mobile social networking. There are many research issues and challenges that can be solved from industry and academia, and some issues can be tackled better than others from industry or from academia. However, we must remember that the solutions that we create are for the betterment of society, and not for personal gain. Therefore, researchers from industry and academia must create solutions that are extensible, non-proprietary, and open so that others can re-use existing frameworks without re-inventing the wheel. We should have open collaboration and dialogue through projects that allow for creative research and innovation, and the research environment should help support that. This means that we need to have constant dialogue through industry and university partnerships, workshops, and conferences. We also must not be afraid to fail or take risks, because failure is the secret to success. In industry, companies are usually risk-averse, so it is difficult for them to undergo more risks. However, corporate research labs need to do that in order to advance research. There should also be more pilots and demonstrations of MSN technology to the public to help spread the MSN vision and make MSN a full reality. For industry labs, there should be a research focus on MSN, with full support from the research leadership team and executive members. For university, there should be courses on MSN, and we encourage professors to use our book to teach the basic and advanced technologies about MSN, and design projects to help students expose the exciting research in MSN.

We, the editors, have attempted in this book to provide the foundational tools for helping individuals to get into MSN research, and for laying out the MSN vision. Now, it is time for industry researchers and academics to work together to address these issues, and to make the MSN vision a reality. To start along the path of dialogue, we have created a Facebook page at http://www.facebook.com/MobileSocialNetworkingBook, as well as a Twitter page at http://twitter.com/msn_book(@msn_book), so we encourage everyone to use this social media to pose and discuss the research issues.

References

Jain, R., & Sonnen, D. (2011). Social life networks. *IT Professional, 13*, 8–11.

Zhang, D., Guo, B., & Yu, Z. (2011). The emergence of social and community intelligence. *IEEE Computer, 44*(7), 21–28.