# An Empirical Investigation of Disclosure of Personal Information in Ubiquitous Social Computing

Antonio Sapuppo and Boon-Chong Seet

*Abstract*—**Privacy has already been identified as the main threat to long-term success of ubiquitous computing, especially in environments, which target at promoting ubiquitous social networking. Notably, these environments are founded on disclosure of personal information and thus, the amount of data disclosed is directly proportional to potential networking benefits. The networking advantages would be maximized by sharing all available personal data, however this would result in jeopardizing of users' privacy and a compromise is necessary. Consequently, privacy management systems of ubiquitous computing must be capable of disclosing only personal data, which is relevant, however not sensitive in specific circumstances. In this paper we provide insight into human personal data sensitivity and disclosure decisions by presenting results of an online survey regarding respondents' willingness to share their personal information under different circumstances. We believe that our findings provide relevant inputs for the design of management privacy models in ubiquitous computing.**

*Index Terms*—**Information Disclosure, Privacy, Social Networking, Ubiquitous Computing.**

## I. Introduction

Even when ubiquitous computing was just a vision [1], privacy threats were already identified as the greatest barrier to the long-term success [2], [3]. Nowadays, the technological development is moving towards people-centric era, where humans are the main focus of sensing. In people-centric sensing, users are parts of mobile sensor networks, where mobile devices are conceptually tied to individuals. New mobile phones, called smartphones, are capable of acquiring not only environmental data, but obtaining users' personal information as well, thanks to their sensing components such as accelerometer, Bluetooth, microphone, etc. Therefore, mobile devices are considered to be key elements in the development of ubiquitous social computing as they are ideally suited to provide insight into social behavior patterns [4].

Ubiquitous social computing (in the following referred to as socUbicomp) environments such as local social networks [5], [6] and other sociable opportunistic networks [7]-[11] target at developing possible advantageous relationships (e.g. friendships, partnerships, business relations) between their participants during physical meetings. Specifically, these environments are based on exchange of personalized profiles not only among friends, but especially between strangers. Thus they lead to new opportunities to leverage interpersonal affinities for personal benefits between people who do not know each other, but probably they should [6], [7]. Indisputably, socUbicomp must be capable of providing a secure and safe exchange and dissemination of users' personal information. This challenge arises due to the fact that the foundation of socUbicomp is based on automated sharing of participants' personal data, which can provoke potential privacy threats. If not addressed responsibly, these threats could motivate users to detain their personal information due to mistrust in socUbicomp [12], [13].

In previous works, it has been already discussed that the central challenge of socUbicomp is shifting from hiding personal data to ensuring accuracy of selective disclosure of users' personal information [14]. Consequently, privacy management systems in socUbicomp must be capable of following the human data sensitivity evaluations and attempt to act as the real user would [2], [15]-[17]. In order to facilitate the development of privacy management systems, the influential factors of human decisions must be taken into consideration. In [2], [6], [18], [19] the sensitivity of the personal information was assumed to vary depending on the inquirer and the situation determinants. The inquirer is considered to be the individual that the user is interacting with and the situation is defined according to the circumstances at that time. Lederer et al [19] determined the identity of the inquirer to be the most important value, influencing the users' privacy choices, followed by the situation as parameter of secondary significance. However socUbicomp advances the attention to the current circumstances as crucial influential factor, due to its primary target to initialize relationships between strangers. Thus, in this paper we present results of a survey, which investigates the sensitivity of different kinds of personal information under different circumstances. Further, we provide insight into the influential factors such as location familiarity and current activities that impact users' personal data disclosure decisions. We believe that our findings provide relevant contributions for understanding human data disclosure choices in order to facilitate further development of privacy management systems in ubiquitous social computing.

## II. Survey design

In order to gain insight into users' perceptions about personal data sensitivity in different circumstances, we asked users to indicate personal information that they would like to

Manuscript received April 15, 2012; revised May 10, 2012. This work is supported by Nokia and developed as a part of the Converged Advanced Mobile Media Platforms (CAMMP) project, funded by the Danish Advanced Technology Foundation. Antonio Sapuppo is with the Aalborg University, Sydhavnsgade 17-19, Frederikskaj 12, Copenhagen 2450, Denmark (e-mail: antonio@ cmi.aau.dk).

Boon-Chong Seet is with the Auckland University of Technology, 92006, Auckland 1142, New Zealand (e-mail: bseet@aut.ac.nz).

share in different situations of their lives. The participants were informed that sharing of personal data is motivated by potential personalized networking services, provided in return to information disclosed. Naturally, the benefits would be directly proportional to the amount of information shared, thus the respondents were asked to compromise between privacy risks and expected benefits. The chosen personal dataset, composed of 28 different types of personal information, was selected in accordance to data categorization in popular online social network sites. The full dataset, chosen for this survey, is shown in Table 1. In order to determine possible circumstances, the most common life situations were grouped into five categories:

1) Family places: these environments can be considered to be places where the user or her family members live, e.g. parents' apartment, uncles' apartment, etc. Thus, it was assumed that the user would encounter family members as well as family members' acquaintances;
2) Social environments: these environments can be considered to be places where the user is spending her leisure time, e.g. restaurants, bars, theaters in the city of the user. Thus, it was assumed that the user would encounter friends and strangers;
3) Holiday: similarly to the social environments, holiday environments were social leisure places, however the user's encounters and activities were occurring outside his home city;
4) Work environments: these environments can be considered to be the employment places of the users, such as universities, offices, etc. Thus, users would mainly encounter co-workers and as well strangers, associated to the user's employment activities;
5) Work Trip: similarly to work environments, during work trips the user was assumed to encounter colleagues and strangers, associated to his employment activities, however these encounters and activities were occurring outside the regular work place.

For example, if the participants of the survey indicated "Name" only under "Family places" and "Work environments", they accepted to share their name among people (i.e. both friends and strangers) as well as service providers in those selected circumstances. The disclosure of personal data was assumed to be limited to the physical surroundings of the user. Further, since they did not indicate "Name" in the remaining three socUbicomp environments (i.e. "Holiday", "Social environments" and "Work Trip"), they implied that sharing the name would jeopardize their privacy in those circumstances. Finally, the respondents also had the opportunity to indicate "Never", which would express that "Name" is too sensitive to be disclosed in any environment, even having taken into consideration the potential benefits.

## III. Survey participants

The distribution of the questionnaire was limited to online social networking users, based on the expected validity of their answers. Particularly, even if the perceptions of data disclosure might vary between virtual and physical worlds, we determined this category to be the most relevant due to their advanced experience with personal data disclosure in online social networking sites. In total we received 121 complete answers, which composed the sample. Following we present the demographic characteristics of the survey sample:

- Gender: 54,5% of the respondents were males and 45,5% were females;
- Age: 64,5% of the respondents were between 26 and 35 years old, 26,5% were in the range of 19 to 26 years old, 5,8% were between 35 and 50 years old and 1,6% were less than 19 years old or more than 50 years old;
- Occupation: 69,4% of the respondents were working, 5,8% were unemployed and the 24,8% were studying at the time of the survey;
- Education: 52,9% of the respondents had a master degree, 19% had a PhD degree, 15,7% had a bachelor degree and 12,4% had a high school degree.

Additionally to the demographic information, respondents were also asked to reflect on their own data disclosure decisions in online social networking sites. We asked them to indicate their preferences on visibility of their own personal data, such as user profile, pictures, posts, etc. Based on these answers, we were able to indicate patterns among data disclosure attitudes and, consequently, investigate whether users' privacy preferences in online social networking sites would reflect to socUbicomp environments. Similarly to Westin [20], our participants were classified into three main clusters: 10,7% as privacy fundamentalists, 74,4% as privacy pragmatists and 14,9% as privacy unconcerned. The clustering was based on the following descriptions:

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

## IV. Survey results and discussion

Before investigating the influential factors of data disclosure in socUbicomp, we provide insight into variation of personal data sensitivity by investigating the most sensitive personal information as implied by the different clusters of respondents. The results are shown in Table 1, in which the percentage indicates the fraction of respondents in the user cluster who considered that the personal information was too sensitive to be shared under any circumstances. Based on these results, it is important to notice that none of the kinds of personal data was indicated as too sensitive to be shared in any circumstances by all the respondents.

TABLE I: SENSITIVITY OF RESPONDENTS' PERSONAL INFORMATION IN ALL CIRCUMSTANCES

| Personal Information | % Fund. | % Prag. | % Unco. |
|---|---|---|---|
| Political views | 69,2% | 42,2% | 33,3% |
| Smoking and drinking | 61,5% | 38,9% | 22,2% |
| Working hours | 53,8% | 30,0% | 5,6% |
| Religion | 46,2% | 37,8% | 22,2% |
| Sexual orientation | 46,2% | 37,8% | 22,2% |
| Personal phone number | 46,2% | 25.6% | 5,6% |
| Home address | 38,5% | 31,1% | 16,7% |
| IM screen names (e.g. Facebook) | 38,5% | 23,3% | 5,6% |
| Gender | 30,8% | 16,7% | 0% |
| Living with (e.g. alone, parents) | 23,1% | 28,9% | 5,6% |
| Interested in (e.g. partner) | 23,1% | 20,0% | 11,1% |
| Work phone number | 23,1% | 10,0% | 0,0% |
| Relation status (e.g. single, etc.) | 15,4% | 26,7% | 16,7% |
| Web site | 15,4% | 22,2% | 5,6% |
| Job position | 15,4% | 15,5% | 0% |
| Work employer | 15,4% | 14,4% | 5,6% |
| Birthday | 15,5% | 8,9% | 0% |
| Favourite books, game, etc. | 15,4% | 8,9% | 0% |
| Personal email address | 7,7% | 16,7% | 0% |
| Interests | 7,7% | 8,9% | 0% |
| Career interests and skills | 7,7% | 5,6% | 5,6% |
| Food tastes | 7,7% | 5,6% | 5,6% |
| Work email address | 7,7% | 4,4% | 5,6% |
| Home city | 7,7% | 4,4% | 0% |
| Education details | 0% | 5,6% | 0% |
| Languages that I speak | 0% | 3.3% | 0% |
| Name | 0% | 2,2% | 0% |
| Nationality | 0% | 2,2% | 0% |

Even if focusing on the fundamentalist cluster, the majority of the data was not preferred to remain confidential by more than 30% of the respondents. In regard to the pragmatists, it can be observed that only 6 out of 28 data types were preferred to remain undisclosed by more than 30% of the survey participants. Finally, only the "political views" data type was considered to be too sensitive for sharing by more than 30% of the unconcerned respondents.

The results presented in Table 1 lead to a conclusion that no data is commonly preferred to remain confidential in all the circumstances. Respondents present general inclination to prefer not to miss potential ubiquitous social networking benefits over privacy concerns by deciding to share their personal information in at least one environment. Notably, in socUbicomp the sensitivity of personal data continuously varies depending on different situations, thus we further research the relevant influential factors by investigating the impact of current users' location familiarity and activity.

### A. Location as influential factor

In this section we investigate whether the disclosure of personal data is influenced by the familiarity of users' locations. Figure 1 shows the average responses of sharing users' personal data in different socUbicomp environments.

Particularly, the respondents tend to share more personal information in familiar locations such as "Family places" and "Work environments". This inclination can be explained by the fact that the users spend the majority of their time in these places and thus they develop an unconscious trust in these environments. In fact, social and work environments also had commonly higher sharing rate in comparison to respectively holiday and work trip environments, even if both circumstances are considered to comprise similar conditions. These results indicate the importance of location familiarity factor as a determinant for selecting personal information to be shared in socUbicomp.

Comparing the responses of different participant clusters, we can notice that all the clusters are willing to share more than 60% of their personal information in family places. This inclination drastically decreases directly proportionally to the unfamiliarity with the environments. However, the decline of data shared is more significant among fundamentalists in comparison to the other clusters.

### B. Activities as influential factor

In this section we investigate whether the disclosure of personal information is also influenced by the current users' activities. In this analysis, we focus on two different subsets related to social and work activities and their associated environments. Firstly, we show respective responses of the fundamentalists. The results in Figure 2 show a significantly different sharing ratio between the two analyze environments. In Figure 2-B, the data related to work activities is shown. Specifically, this set of personal data reached common sharing acceptance in "Work environments", while it attained mainly denial sharing rates in "Social environments". On the contrary, responses regarding personal information, related to social activities, are presented in Figure 2-A. It can be noticed that the fundamentalists are generally less open to sharing personal data in social environments. However, the influence of activity factor, even if minimized, can be still observed. In fact, fundamentalists recognized the impact of particular data types related to social activities by presenting higher sharing tendency in social environments than in work environments.

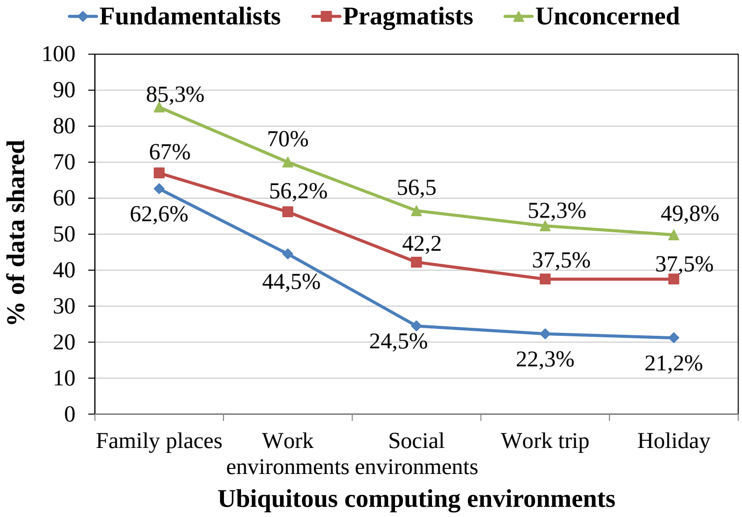

Fig. 1. Average data sharing in different ubiquitous social computing environments

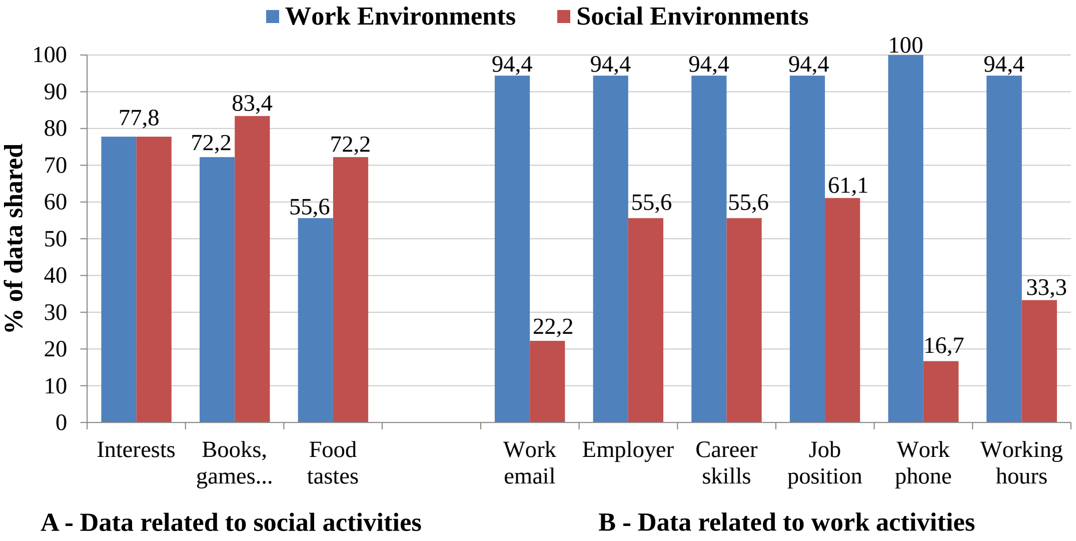

Fig. 2. Extent of data sharing among fundamentalists

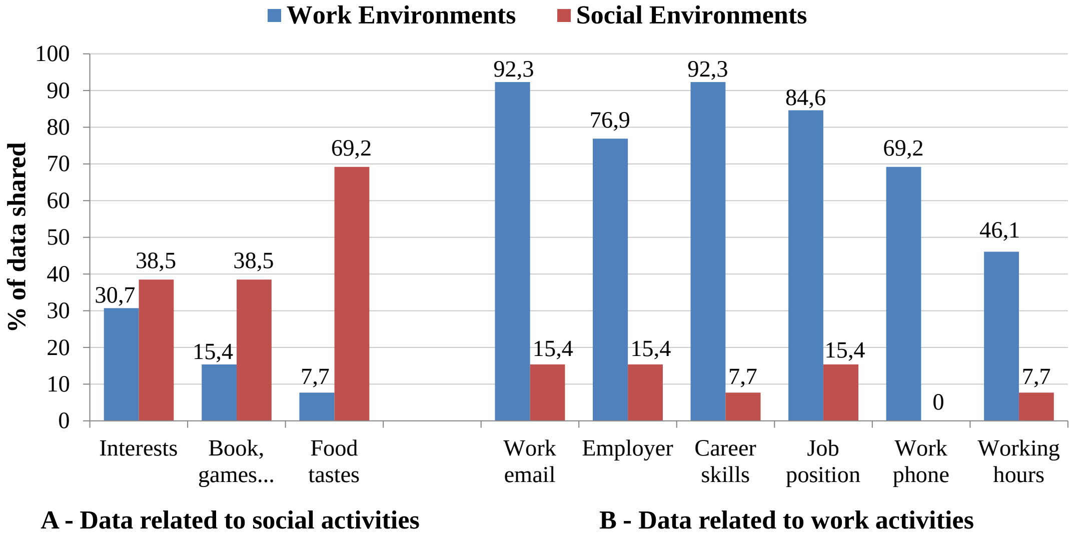

Fig. 3. Extent of data sharing among pragmatists

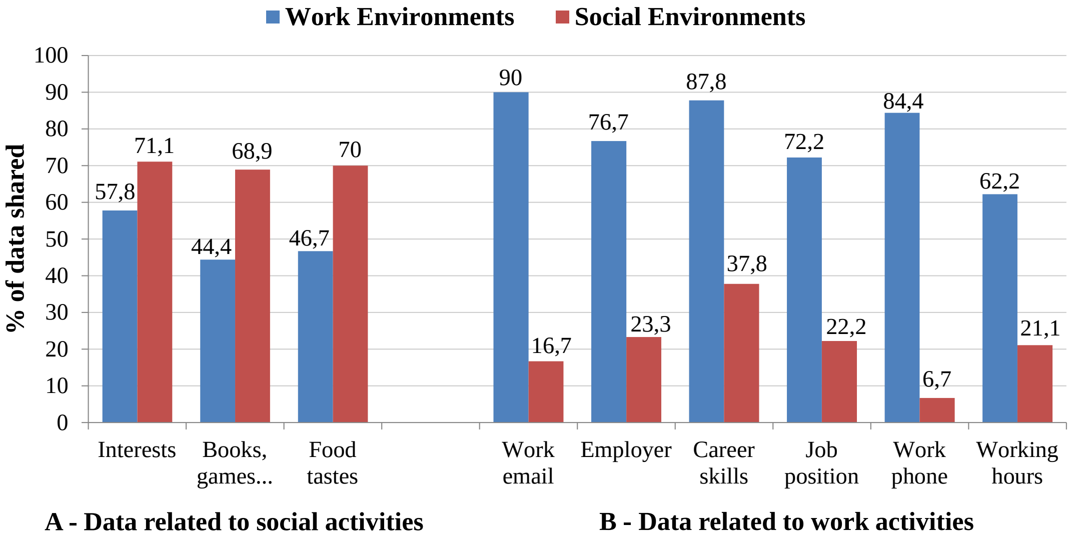

Fig. 4. Extent of data sharing among unconcerned respondents

Fig. 3 shows results regarding the pragmatists privacy cluster. In comparison to the responses of the fundamentalists, the pragmatists are generally willing to share more personal information - slightly more data related to work activities in work environments and significantly more data related to social activities in the social environments. These choices present strategies that guide the data disclosure of pragmatists, based on the evaluation of the current activity as a crucial determinant.

Finally, Figure 4 presents survey results of the unconcerned respondents. Notably, the amount the amount of personal data shared is the highest in regard to both activities in comparison to the other clusters of respondents. Moreover, the unconcerned cluster was not presenting the same extent of data disclosure based on the activity determinant as influential factor. Particularly, while sharing of personal data related to work activities still present relevant variation between the two activities (Figure 4-B), the relevance of the activity determinant is reduced or even not taken into account any more in disclosure of personal data related to social activities (Figure 4-A). In fact, personal data related to social activities reached high sharing ratio not only in social environments but as well in work environments.

## V. Conclusions

In this paper we investigated users' perceptions of data sensitivity and influential factors that impact users' personal information disclosure decisions, by conducting a survey based on the respondents' willingness to share their personal information in exchange for networking benefits in different socUbicomp environments. The survey results did not indicate any personal data that would be commonly defined as too sensitive to be shared in any circumstances. Furthermore, the location familiarity factor was commonly approved by all the respondents who presented tendency to be more open to share their personal information in more familiar locations. The investigation of the activity factor, instead, presented different behaviors in disclosure of personal information among the three privacy clusters. While fundamentalists and pragmatists had different behaviors upon different activities, the privacy unconcerned cluster were less influenced by evaluation of the current activity as a crucial determinant, especially in regard to data related to social activities. These results strongly encourage further research on privacy of socUbicomp, focusing not only on the inquirer, but also on familiarity of the users' location and current activities as crucial parameters for selecting personal data to be disclosed. Finally, we also noticed relation between users' personal privacy preferences in online social networks and in socUbicomp. Consequently, knowing online social networking privacy preferences and crossmatching them with relevant influential factors such as familiarity of user location and current activity would provide relevant input for the design of privacy management systems. Thus we would also recommend further investigation into application of online social networking users' privacy preferences in socUbicomp.

## References

[1] M. Weiser, "The computer for the 21st century," *Scientific American*, vol. 272, no. 3, pp. 94-104, 1991.

[2] C. Bünnig, "Simulation and analysis of ad hoc privacy control in smart environments," *Intelligent Interactive Assistance and Mobile Multimedia Computing*, pp. 307-318, 2009.

[3] J. I. Hong, J. D. Ng, S. Lederer, and J. A. Landay, "Privacy risk models for designing privacy-sensitive ubiquitous computing systems," in *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques*, 2004, pp. 91-100.

[4] E. Miluzzo, N. D. Lane, K. Fodor, R. Peterson, H. Lu, M. Musolesi, S. B. Eisenman, X. Zheng, and A. T. Campbell, "Sensing meets mobile social networks: the design, implementation and evaluation of the cenceme application," in *Proceedings of the 6th ACM conference on Embedded network sensor systems*, 2008, pp. 337-350.

[5] A. Sapuppo, "Spiderweb: a Social Mobile Network," in *Wireless Conference (EW), 2010 European*, 2010, pp. 475-481.

[6] A. Sapuppo and L. T. Sørensen, "Local social networks," in *International Proceedings of Computer Science and Information Technology – Computer Communication and Management*, vol. 5, pp. 15-22, 2011.

[7] N. Eagle and A. Pentland, "Social serendipity: Mobilizing social software," *IEEE Pervasive Computing*, vol. 4, no. 2, pages 28-34, 2005.

[8] A. Gupta, A. Kalra, D. Boston, and C. Borcea, "Mobisoc: a middleware for mobile social computing applications," *Mobile Networks and Applications*, vol. 14, no.1, pp. 35-52, 2009.

[9] V. Kostakos and E. O'Neill, "Cityware: Urban computing to bridge online and real-world social networks," *Handbook of Research on*

*Urban Informatics: the Practice and Promise of the Real-Time City*, pap. 195-204, 2008.

[10] P. Persson and Y. Jung, “Nokia sensor: from research to product,” in *Proceedings of the 2005 conference on Designing for User eXperience*, 2005, pp.53.

[11] A. K. Pietiläinen, E. Oliver, J. LeBrun, G. Varghese, and C. Diot, “Mobiclique: middleware for mobile social networking,” in *Proceedings of the 2nd ACM workshop on Online social networks*, 2009, pp. 49-54.

[12] G. D. Abowd and E. D. Mynatt, “Charting past, present, and future research in ubiquitous computing,” *ACM Transactions on Computer-Human Interaction (TOCHI),* vol. 7 no. 1, pp. 29-58, 2000.

[13] J. Bohn, V. Coroama, M. Langheinrich, F. Mattern, and M. Rohs, “Social, economic, and ethical implications of ambient intelligence and ubiquitous computing,” *Risk*, vol. 10, no. 5, 2004.

[14] M. Langheinrich, “Privacy by design - principles of privacy-aware ubiquitous systems,” in *Ubicomp 2001: Ubiquitous Computing*, 2001, pp. 273-291.

[15] C. Bünnig, “Smart privacy management in ubiquitous computing environments,” *Human Interface and the Management of Information. Information and Interaction*, pp 131-139, 2009.

[16] A. Kapadia, T. Henderson, J. Fielding, and D. Kotz, “Virtual walls: Protecting digital privacy in pervasive environments,” *Pervasive Computing*, pp. 162-179, 2007.

[17] G. Yee, “Using privacy policies to protect privacy in ubicomp,” in *Proceedings of the IEEE 19th International Conference on Advanced Information Networking and Applications*, 2005, vol.2, pp. 633-638.

[18] U. Jendricke, M. Kreutzer, and A. Zugenmaier, “Pervasive privacy with identity management,” in *Proceedings of the Workshop on Security in Ubiquitous Computing, Ubicomp*, 2002.

[19] S. Lederer, J. Manko and A. K. Dey, “Who wants to know what when? privacy preference determinants in ubiquitous computing,” in *CHI’03 extended abstracts on Human factors in computing systems*, 2003, pp. 724-725.

[20] A. F. Westin, “Harris-equifax consumer privacy survey 1991,” *Atlanta, GA: Equifax Inc*, 1991.

**Antonio Sapuppo** is a Ph.D. student from 2009 at the Aaborg University in Copenhagen (Denmark). He holds a master degree in Software Engineering from the Aalborg University and a bachelor degree in Computer Science at the University of Studies of Catania (Italy). During 2008 and 2009, he has been working as research assistant at the Copenhagen University College of Engineering (Denmark) to design, implement and test a mobile social network application, called Spiderweb. Antonio Sapuppo has been visiting during fall 2011 the Auckland University of Technology (New Zealand) focusing on social ambient intelligence research area. He is currently involved in the CAMMP project - Converged Advance Mobile Media Platform. His main interest areas are privacy, social networks, mobile applications, mobile services and ubiquitous computing.

**Boon-Chong Seet** obtained his PhD in Computer Engineering from Nanyang Technological University, Singapore, in 2005. Upon graduation, he was employed as a Research Fellow under the Singapore-Massachusetts Institute of Technology Alliance (SMA) program at the National University of Singapore, School of Computing. In 2007, he was awarded a visiting scholarship to the Technical University of Madrid, Spain, to pursue research under an EU-funded project on multi-disciplinary advanced research in user-centric wireless network enabling technologies (MADRINET). Since December 2007, he is with the Auckland University of Technology, New Zealand, where he is currently a Senior Lecturer in its Department of Electrical and Electronic Engineering. He was also a visiting faculty at the University of British Columbia, Canada. He has served as Guest Editor for special issues in IEEE Wireless Communications Magazine and ACM/Springer Journal of Personal and Ubiquitous Computing. His recent research activities include sensor networks, ambient intelligence, and mobile computing with applications in healthcare and education. He is a member of ACM and a senior member of IEEE.