http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 8 (6): 910-916, 2009
ISSN 1812-5638

Privacy Preserving in Ubiquitous Computing: Architecture

Tinghuai Ma, SenYang, Wei Tian and Wenjie Liu
School of Computer and Software, Nanjing University of Information Science and Technology, China

Abstract: In this study, we summarize the main principles of privacy-aware system and present a new architecture. We preserve the person's location privacy by using the methods of spatiotemporally-based anonym and location information disturbing. In the spatiotemporally-based anonym method, the space and time are divided into pieces. While, an entity hands over from one domain to another, its ID will be refreshed. In the location information disturbing method, there are two methods to disturb coordinate data. One is transferring coordinate to a random data, the other is transferring coordinate to a fixed data.

Key words: Ubiquitous computing, privacy system, architecture, spatiotemporally-based, anonymity, location information disturbing

INTRODUCTION

Ubiquitous computing represents the concept of seamless everywhere computing and aims at making computing and communication essentially transparent to the users. In ubiquitous computing environment, we will be surrounded with a comfortable and convenient information environment that merges physical and computational infrastructures into an integrated habitat (Weiser, 1998). Privacy in ubiquitous computing has been a contentious issue and the privacy concerns that have been raised suggest that privacy may be the greatest barrier to the long-term success of ubiquitous computing (Hong *et al*., 2004; Alexander, 2007; Marc, 2007; Cardoso and Issarny, 2007).

Privacy preserving in the ubiquitous computing mainly focuses on two aspects. One is the anonymous/ pseudonym oriented, the other is the policy based.

For the anonymous based privacy enhancing technology, there are k-anonymities that they can't be distinguished from each other. In other words, there are k entities having the same ID or in the same location at the same time. The higher value of k is, the higher level of anonymity (Sweeney, 2002). For some ID based services, the virtual identities are used to conceal the real identity of the user (Papadopoulou *et al*., 2008). Anonym and authentication are always conflict. Diep *et al.* (2007) presented a scheme using anonymous user ID, sensitive data sharing method and account management to provide a lightweight authentication while keeping users anonymously interacting with the services in a secure and flexible way.

For the policy based privacy enhancing technology, the main concept is storing privacy-compliant rules to process personal information. The stored privacy policies describe the allowed recipients, uses and storage duration of users' data. Also, a policy engine is used to reason the compatible privacy policy. Now, privacy policy is described in XML (Langheinrich, 2002; Myles *et al*., 2003) and XACML (Zheng *et al*., 2007). Most privacy policy based Privacy Enhanced Technologies (PETs) are focus on one or more scenarios, the configuration of policy for each scenario is a huge work (Pallapa *et al*., 2008).

Using the PETs, several privacy awareness prototypes have been built in Ubiquitous Systems (US). Langheinrich (2002) presented a privacy awareness system named pawS for the Ubiquitous Computing Environment (UCE). The pawS has four core concepts:

- Machine-readable privacy policies to provide choice and consent
- Policy announcement mechanisms to give notice
- Privacy proxies for supporting access
- Privacy-aware databases for recourse

Xiaodong *et al.* (2002) proposed another framework for the UCE. Their framework presented a key objective called the Principle of Minimum Asymmetry, which seeks to minimize the imbalance between the people whose data is being collected and the systems and people that collect and use those data. Jalal *et al.* (2002) suggested a communication infrastructure named MIST, allowing simultaneously authentication and privacy protection. Cardoso (2007) proposed taxonomy for privacy invasion attacks, classifies existing privacy enhancing technologies according to the protection provided for those attacks and presented a service-oriented privacy-enhanced architecture for pervasive computing.

Corresponding Author: Tinghuai Ma, Ningliu Road 219, Pukou, Nanjing, Jiangsu, 210044, China Tel: 86-25-5873-1400

The privacy preserving architectures are studied by several researchers. Bhatti *et al.* (2007) presented an enforcement architecture for the healthcare information access control, which is privacy policies based. The requirements specification is used for describing Clinical Document Architecture (CDA) and specification language is used to encode the CDA-based requirements into privacy and disclosure policy rules. Kaya *et al.* (2009) Public Key Cryptography (PKC) architecture to ensure that the access rights of the RFID tags are preserved based oil the spatial and temporal information collected from the RFID readers. Gedik and Liu (2008) presented architecture includes the development of a personalized location anonymization model and a suite of location perturbation algorithms. This flexible privacy personalization framework can support location k-anonymity for a wide range of mobile clients with context-sensitive privacy requirements. This framework enables each mobile client to specify the minimum level of anonymity that it desires and the maximum temporal and spatial tolerances that it is willing to accept when requesting k-anonymity-preserving LBSs.

In this study, based on the anonymous/pseudonym methods, a privacy preserving architecture is presented. We design the service consumption process for privacy preserving. The spatiotemporally-based anonym and location information disturbing method are adopted for avoiding traced. The architecture is more powerful for hiding the identity than before.

PRIVACY OF UCE

According to the view of Gritzalis (2004), privacy is defined as the right to be left alone. Generally speaking, personal information can be divided into two parts: identification and profile. Identification is the part that can identify one individual uniquely such as identification number, name and so on. Profile is the part affiliated to the identification, such as hobbies and habits. Their aggregation represents a physical entity, whether it is a person or an objective. Any individual has the right to know what, where and how some devices collect and use these data. Zugenmaier (2003) proposed a privacy diamond as shown in Fig. 1. That shows us an intuitionist comprehension on the relationship between devices, individuals and relevant data.

The device, namely smart device such as RFID scans perpetually to require information directly or indirectly, who is entering the scope, where exactly he is standing, what actually he is doing. In general circumstances, devices here in the ubiquitous environment are often

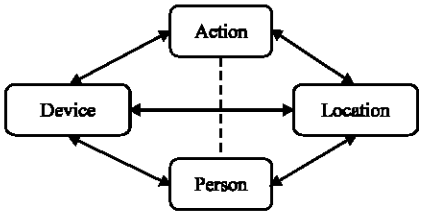

Fig. 1: The privacy diamond

light-weighted, characterized as limited processing capacities and insufficient power, it straight leads to the lack of security measurements such as authentication, complex verification computation.

In the practical UCE, since privacy has no strict definition, people can hardly reach a consensus on this issue. Various service providers need diverse information to deliver distinct services, whether temporally-based or spatially-based. A fat lot of regulations that does legislation in some cases: clandestine intrusion, purposive invasion. Consequently, only rely on legislation is not necessarily able to solve anything concerning privacy. To minimize such risks, we should employ more sophisticated techniques as another vehicle to counterpoise privacy and services.

ARCHITECTURE

Combine all respects, we focus on three main principles in our Privacy-Aware System (PAS), they can be described as follows:

- **Anonymity:** Anonymity is a broad conception, it not only refers to identities of an individual, but also covers the places, hobbies, habits and such information concerning to internal and external aspects of an individual
- **Untraceability:** Almost everyone in the UCE doesn't want to be traced. For instance, John requests a location-based service at his office, but he is not necessarily willing the service provider to record his location where he is after he is off duty
- **Confidentiality:** All data must be kept security

We might as well commence with the fundamental structure, based on the current study.

The structure of PAS: The structure of PAS is shown by Fig. 2. The working process of PAS can be stated as follows:

- **Step 1:** Entity (e) creates a Random Serial Number (RSN) and queries services available in the ambient environments through access base. Access base here is supposed to be trustworthy and RSN is dynamic, namely, each query has different RSN
- **Step 2:** Access base utilizes the Random Serial Number (RSN) it received as contact identity and directly contacts with available service providers
- **Step 3:** Service providers send service list to the Access Base, using RSN as contact ID. In practice, millions of people query at a time and the RSNs do not carry any individual identifiable information. So, it is impossible for any provider to trace a certain individual and create their profiles
- **Step 4:** Access base transfers the list to the requester via., RSN, until now, RSN become invalid. On the assumption that the requester needs the service on the list, such as a location-based service, he selects a service. If there is no favorite service on the list, then the process is terminated
- **Step 5:** Entity generates a new RSN (called e-RSN) as identity to query service, as well as pre-configured Privacy Policy List (PPL) and sends it to Privacy System (PS) this time instead of access base
- **Step 6:** Entity sends his identification to Authentication Center (AC) for being authenticated
- **Step 7:** Privacy system generates another RSN (called ps-RSN) and correlates it with e-RSN, then PS contacts with relevant Service Provider (SP)
- **Step 8:** Service provider communicates with AC to verify its validity, sending the ps-RSN
- **Step 9:** If the SP is genuine, then AC transfers its notarization to PS, taking ps-RSN as identification
- **Step 10:** Service provider claims its PPL to PS, PS compares two privacy lists and determines whether they match well. If they do not match well, the system notices entity to determine what to do next, this just like Langheinrich's structure (Langheinrich, 2002)
- **Step 11:** If they match well, PS applies for the service on behalf of entity, utilizing its own ps-RSN
- **Step 12:** If PS gets the service successfully, then transfers it to the entity, via., relationship of e-RSN and ps-RSN

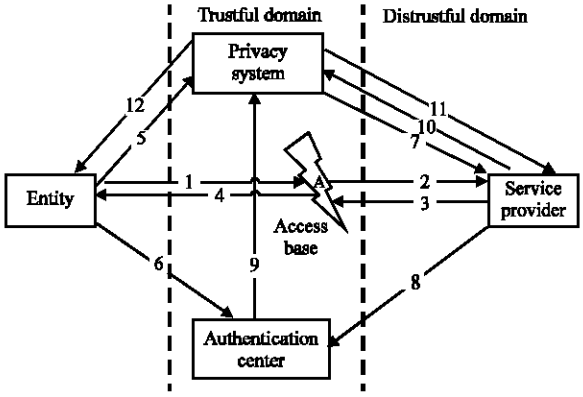

Fig. 2: The Structure of PAS

Here, we introduce access base, which is software and responsible to deal with the query. The access base is separated from privacy system, because it can deal with all entities with no authentication. It will respond entity's query quickly, as avoiding complexity authentication in privacy system. At the same time, the malicious entities will be filtrated outside the privacy system.

As we see, while privacy system receives the e-RSN from entity in step 5, it will generate new identification ps-RSN and sends it to service provider. So, a mapping relation of e-RSN and ps-RSN needs to be established. The same problem is encountered by the entity identification and e-RSN in the AC as well as the ps-RSN and service required in the SP.

We formally describe the mapping processing as follows:

$$\text{e-RSN} = f_1(\text{Entity}) \quad (1)$$

$$\text{ps-RSN} = f_2(\text{e-RSN}) \quad (2)$$

$$\text{ps-RSN} \rightarrow \text{SP} \quad (3)$$

where, f_1, f_2 are random functions. Equation 3 means giving ps-RSN to a SP as its identifier.

Here, we use two-dimensional forms to record those mapping relations, shown in Table 1-3, respectively as follows:

Table 1: The mapping relation of entity identification and e-RSN in the AC

Entity identification	e-RSN
Bob	Iete32455j
Alice	32jijanrfe
John	faee5rg5y

Table 2: The mapping relation of e-RSN and ps-RSN in the PS

e-RSN	ps-RSN
Iete32455j	e45fdq3etr
32jijanrfe	6778trefw
faee5rg5y	tggrytr45

Table 3: The Mapping Relation of ps-RSN and service required in the SP

ps-RSN	Service required
e45fdq3etr	LBS 1
6778trefw	LBS 2
tggrytr45	LBS 1

According to above discusses, we can answer the following question:

- Why we should trust privacy system or AC in the trustful domain, how can we make sure they will never betray us?

It is a quite hard question to answer. Different from former computing environment, trust in ubiquitous environment is dynamic and temporary. No one can be sure that another peer is trustworthy or trustless if we do not know him better. As a matter of fact, in such environment, trust level is hard measured towards each other. However, if we do not set a trust domain to peers in the environment, it would probably not be able to process any computing. Traditional certificates center based on known trustful entity is not suitable for ubiquitous environment.

- Why Privacy System generates another RSN called ps-RSN and correlates it to the e-RSN?

In this architecture, privacy system in fact represents the entity. Vulgarly speaking, it is a proxy of entity. Since privacy system is trustworthy, it contains sensitive information of entity. If the system utilizes e-RSN to communicate with SP, SP may exploit the e-RSN to trace the entity's locus. The position of PS is public, while the position of entity is private; we design present system using public position of system instead of private position of entity.

- What if we integrate the privacy system into authentication center?

It is feasible indeed, however, we separate them into two in that privacy system usually acquires profile information of an entity rather than identification information, while authentication center acquires identification information of an entity rather than profile information. Integration may cause leakage of the entire information if there is any loophole in the integration system.

Spatiotemporally-based anonymous: As there is no absolute safety in the world, however strict rules are. We introduce a strategy named spatiotemporally-based anonymous matching strategy. It can be designed like this: system divides time into pieces, carves out the entire domain into some certain areas. When a person's random data stream is due or out of the confined area, the random data stream updates itself and distributes a new data stream to represent the person, then invalidates the old random data stream. Doing this is able to ultimately avoid the long-time attacking to a certain people. In this case, attacking is not efficient in that there is no connection between old random data stream and new random data stream. The method of data stream changed is showed in Fig. 3.

Table 4: The rules of boolean calculation

VLs expire (valid location)	VLt expire (valid time)	Refresh
Yes	No	Yes
Yes	Yes	Yes
No	Yes	Yes
No	No	no

Definition
A sub-domain is a set containing following data:
Sb{valid location set{}, valid time set{}}
where, Valid Location set contains following data:
Vls{[x_1,x_2], [y_1, y_2], [z_1, z_2]}
And Valid Time set includes following data:
Vlt{[t_1, t_2]}
Its Boolean calculation is:
Vls×VLt

Fig. 3: The method of data stream changed

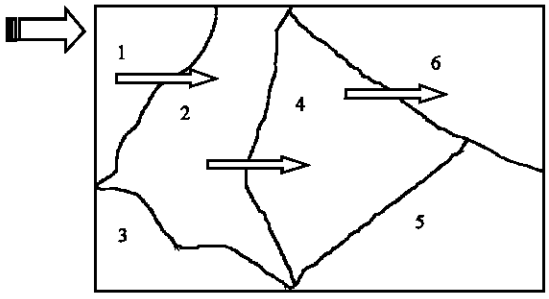

Fig. 4: Sub-domains in a domain

It can be seen that the data changed switch is based on VLs×VLt. Its operation rules are shown in Table 4.

We construct a scenario to illustrate this spatiotemporally anonymous method.

As shown in Fig. 4, we divide a domain into 6 sub-domains. Bob is in sub-domain 1 at 1:00, he is admeasured a Random Serial Number (RSN), representing his real identity. For the sake of handiness, we consolidate every sub-domain configuration. In practical circumstance, every sub-domain can have its own configuration. Supposing the time interval is 30 min, there are two circumstances:

- If he enters sub-domain 2 from sub-domain 1 within the prescriptive time, namely, within the lifespan of RSN, 30 min here, the RSN refreshed itself, replacing the old random serial number
- If he stays in sub-domain 1 until the expiry of prescriptive time, more than 30 min here, the RSN refreshed itself, replacing the old random serial number

Definition
Spatiotemporal Service is a set of spatiotemporal elements and type and it can be defined as a five-element set:
SS{ $[x_1, x_2]$, $[y_1, y_2]$, $[z_1, z_2]$, $[t_1, t_2]$, s}
Where x, y, z is the spatial information of services, t is the temporal information of services and *s* is the type of available service

Fig. 5: The definition of services

Definition
Entity Spatiotemporal Information (ESI) is a set of spatiotemporal elements and type and it can be defined as a five-element set:
ESI {x, y, z, t, rs}
Where x, y, z is the spatial information of service, t is temporal information of service and rs is the required service type

Fig. 6: The definition of entity

According to the security level, we can adjust the area of the sub-domain and the time interval. If we need high security level, we shorten the area and the time interval, but it may cause a slight more consumption of system resources. If we need low security level, we can prolong the area and time interval.

Service matching: While, entity applies for services after being authenticated, the privacy system will check if the services can meet entity's requirements. All the services in ubiquitous computing have the limits of spatial and temporal. Only the service's spatial and temporal preferences strictly match with the entity's requirements, the service is valid for entity.

The services can be defined as shown in Fig. 5. We can see that every service has its service area ($[x_1, x_2]$, $[y_1, y_2]$, $[z_1, z_2]$), the valid time domain ($[t_1, t_2]$) and the service type s.

For example, service can be SS {[2, 5], [1, 7], [3, 8], [1:00, 5:00], {GPS, taxi service, download service, e-book service}}, which means that there are GPS, taxi, download and e-book services are valid in [2, 5], [1, 7], [3, 8] domain from 1:00 to 5:00.

When entity wants to apply for services, its requirements should match with service's preferences. So, the entity could be described in the same way as shown in Fig. 6.

For example, an entity's information can be ESI {2, 3, 5, 3:10, taxi service}.

The service matching can be processed as follow:

$$mSer = x \in [x_1, x_2] \wedge y \in [y_1, y_2] \wedge z \in [z_1, z_2] \wedge t \in [t_1, t_2] \wedge rs \in s \quad (4)$$

If mSer is true, the service matching is successful.

The process mentioned above is rational and can be receivable. But, the entity's position (expressed by x, y, z) should be provided precisely. Unfortunately, the position is the location of entity, it needs protecting. There are two methods to disturb entity's location information, but not affect the service matching.

Random coordinates: Since, the service is available in a relatively larger domain, the privacy system masks the entity's precise spatiotemporal information, instead of it, system randomly chooses a geographic location coordinates within the valid area and sent it to the relevant SP.

Using above example, the entity's information is ESI {2, 3, 5, 3:10, taxi service}, the service information is SS {[2, 5], [1, 7], [3, 8], [1:00, 5:00], {GPS, taxi service, download service, e-book service}}. The privacy system can disturb the entity's location information to a random data as:

ESI {2, 3, 5, 3:10, taxi service} → ESI {4, 7, 3, 3:10, taxi service} (5)

The new entity information is within the valid area.

Fixed coordinates: That is, privacy system provides a fixed coordinate to whoever enters the service area regardless of any different coordinates, if the enters are in the service valid area.

ESI {2, 3, 7, 3:10, taxi service} → ESI {3, 6, 5, 3:10, taxi service} (6)

ESI {4, 2, 4, 4:30, download service} → ESI {3, 6, 5, 3:10, download service} (7)

ESI {5, 5, 6, 13:12, e-book service} → ESI {3, 6, 5, 3:12, e-book service} (8)

Privacy system acts as a proxy to any entity who or which enters the service area.

DISCUSSION

Security and privacy analysis: Present proposed scheme has some nice security properties as follows.

Anonymity: The anonymity is the most significant character in our proposed scheme. Firstly, users are anonymously while in querying, applying and utilizing service, as only RSNs used to indicate the users. Secondly, the RSNs can be changed using spatiotemporally-based anonymous matching strategy. Thirdly, in the service matching, the users' coordinate will be replaced as a fake one.

Protection of user preference: In present scheme, there is only one step needs users' preference for authentication. We assume the authentication center is

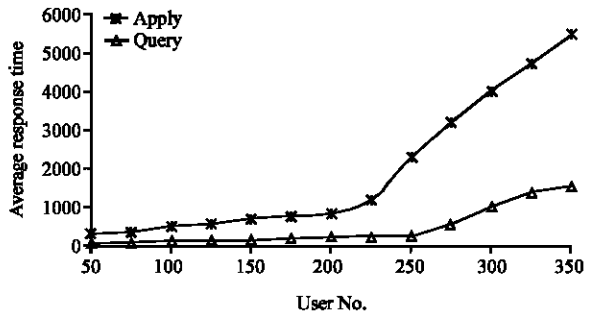

Fig. 7: Time cost of query and apply service

trustful. So, we can say users' preference won't be leaked if the authentication center is security.

Non-linkability: Ideally, nonlinkability means that, for both users (i.e., entities) and service providers, (1) neither of them could ascribe any session to a particular user and (2) neither of them could link two different sessions to the same user (Diep *et al*., 2007). In this scheme, non-linkability is achieved for both users and service providers. For each session, every user has different temporal RSN as identifier under spatiotemporally-based anonymous. Therefore, there is no relationship among user and service provider.

Performance evaluation: In present provided scheme, there are several steps for service consumption transaction. In some steps, the spatiotemporally-based anonymous is adopted. And at last, the coordinator is disturbed in the service mapping step. The computation cost is most important aspects of performance. We divide the steps into two processes. One is query service process, which includes step 1 to 4. The other is applying service process, which include step 5 to 12. Figure 7 shows the average time cost of these processes according to different user number. We evaluate this time cost on normal PC.

Figure 7 shows that present scheme is suit for users not more than 200. The query process cost little time in this scheme. It proof that we design access base to deal with query service process. This architecture is much more complexity while two anonymous technologies are adopted. So, it is not suit for wide range domain.

CONCLUSIONS

With the development of ubiquitous computing technologies, its applications will cover every aspects of our life. The user's willingness of acceptance is the key to success. If users can manage the privacy risks of exposing their personal information to the UCE, they can accept the ubiquitous environments well. The intelligent environments that provide benefits to the user without invading their privacy will be much more attractive than the traditional security and sensing environments. In this study, based on some earlier study, we propose architecture of privacy protection in ubiquitous computing.

This study is the first step to integrate privacy protection technologies into access control architecture for ubiquitous computing. Right now, we are working on designing the ubiquitous software system based on service-oriented. Privacy architectures are crucial for the definition of a software design where technologies do not overlap, do not have conflicting requirements and cooperate to provide multi-level privacy protection and effectively protect the user's personal data.

ACKNOWLEDGMENTS

Present study is partly supported by Natural Science Foundation from Nanjing University of Information and Science Technology (20080302), the overseas study scholarship of Jiangsu government and Jiangsu youth project, Foundation of President of the Chinese Academy of Sciences (O65001H).

REFERENCES

Alexander, A., 2007. Privacy Issues and Concerns-From a Ubiquitous Computing Point of View. IT-University of Gothenburg, Ronneby, Sweden.

Bhatti, R., A. Samuel, M.Y. Eltabakh, H. Amjad and A. Ghafoor, 2007. Engineering a policy-based system for federated healthcare databases. IEEE Trans. Knowledge Data Eng., 19: 1288-1304.

Cardoso, R.S. and V. Issarny, 2007. Architecting pervasive computing systems for privacy: A survey. Proceedings of the WICSA '07 the Working IEEE/IFIP Conference on Software Architecture, Jan. 6-9, IEEE Computer Society, pp: 26-26.

Diep, N.N., S.Y. Lee, Y.K. Lee and H.J. Lee, 2007. A privacy preserving access control scheme using anonymous identification for ubiquitous environments. Proceeding of the 13th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications, Aug. 21-24, Daegu, Korea, pp: 482-487.

Gedik, B. and L. Liu, 2008. Protecting location privacy with personalized k-anonymity: Architecture and algorithms. IEEE Trans. Mob. Comput., 7: 1-18.

Gritzalis, S., 2004. Enhancing web privacy and anonymity in the digital era. Inform. Manage. Comput. Secur., 12: 255-287.

Hong, J.I., J.D. Ng, S. Lederer and J.A. Landay, 2004. Privacy risk models for designing privacy-sensitive ubiquitous computing systems. Proceedings of ACM Conference on Designing Interactive Systems, Aug. 1-4, Cambridge, Massachusetts, pp: 91-100.

Jalal, A., R. Anand, C. Roy and M. Dennis Mickunas, 2002. A flexible, privacy-preserving authentication framework for ubiquitous computing environments. Proceedings of the 22nd International Conference on Distributed Computing Systems Workshops, Jul. 2-5, IEEE Computer Society, pp: 771-776.

Kaya, S.V., E. Savas, A. Levi and O. Ercetin, 2009. Public key cryptography based privacy preserving multi-context RFID infrastructure. Ad Hoc Netw., 7: 136-152.

Langheinrich, M., 2002. A privacy awareness system for ubiquitous computing environments. Proceedings of the 4th International Conference on Ubiquitous Computing, Sept. 9-Oct. 1, Goteborg, Sweden, pp: 237-245.

Marc, L., 2007. RFID and Privacy. In: Milan Petkovic, Willem, J. (Ed.). Security, Privacy and Trust in Modern Data Management, Springer, pp: 433-450.

Myles, G., A. Friday and N. Davies, 2003. Preserving privacy in environments with location-based applications. Pervasive Comput. IEEE., 2: 56-64.

Pallapa, G., N. Roy and S.K. Das, 2008. A scheme for quantizing privacy in context-aware ubiquitous computing. Proceeding of IET 4th International Conference on Intelligent Environments, Jul. 21-22, Seattle, USA., pp: 1-8.

Papadopoulou, E., S. McBurney, N. Taylor, M.H. Williams, K. Dolinar and M. Neubauer, 2008. Using user preferences to enhance privacy in pervasive systems. Proceeding of 3rd International Conference on Systems, Apr. 13-18, Cancun, Mexico, pp: 271-276.

Sweeney, L., 2002. k-Anonymity: A model for protecting privacy. Int. J. Uncertain. Fuzziness Knowl.-Based Syst., 10: 557-570.

Weiser, M., 1998. The future of ubiquitous computing on campus. Commun. ACM, 41: 41-42.

Xiaodong, J., I. Jason Hong and A. James Landay, 2002. Approximate Information Flows: Socially-Based Modeling of Privacy in Ubiquitous Computing. In: UbiComp 2002, Borriello, G. and L.E. Holmquist (Eds.). Springer-Verlag, Berlin, Heidelberg, pp: 176-193.

Zheng, Y., D. Chiu, H. Wang and P. Hung, 2007. Towards a privacy policy enforcement middleware with location intelligence. Proceedings of the 11th IEEE International Enterprise Distributed Object Computing Conference Workshop, Oct. 15-16, Maryland, pp: 97-104.

Zugenmaier, A., 2003. Anonymity for Users of Mobile Devices through Location Addressing. Rhombos Verlag, Berlin.