An Effective Method for Location Privacy in Ubiquitous Computing

Gunhee Lee^{1,*}, Wonil Kim², and Dong-kyoo Kim³

¹ Graduate School of Information Communication, Ajou University, Suwon, Korea
{icezzoco, dkkim}@ajou.ac.kr

² College of Electronics and Information, Sejong University, Seoul, Korea
Tel.: +82-2-3408-3795
wikim@etri.re.kr

³ College of Information Communications, Ajou University, Suwon, Korea
dkkim@ajou.ac.kr

Abstract. In ubiquitous computing environment, every service should have the characteristic of context-awareness. Since physical location is important information to grasp a user's context, the service provider should maintain the location information of the user to provide necessary service. When a malicious user acquires the location information of any other user, it is possible that he/she deduces the context of the user. Although there have been many researches for the problem, previous methods impose complex computation and many subjects requires on the system in order to gain acceptable anonymity. In this paper, we propose an effective method that protects the location privacy in ubiquitous computing environment. This method prohibits the malicious user from hijacking the location information by diffusing the information. It also confuses the attacker with transmitting dummy messages as normal users do.

1 Introduction

In ubiquitous computing environment, computing resources are embodied in everywhere around the people. In street, driving car, flying airplane, office, and even body, the resources are working to provide many services for the people. Further more, these services are provided timely and appropriately according to the user's context. For example, when a person falls down on a street and hurts himself/herself, someone at the vicinity calls 911 for an ambulance in the existing environment. Whereas, in the ubiquitous computing environment, the nearest computing resource, such as camera and his/her watch that includes medical sensor, is aware of the situation, and it sends an alert to the nearest hospital. The ambulance arrives in several minutes, and doctors prepare medical treatment for the injured person. To provide this kind of service environment, the system should maintain the location information of the user. Physical location of a user is important information to grasp the context of the user [1]. For example, when a professor enters the classroom at school hour, he/she may

* Corresponding author.

have a class. Thus the system provides the materials for the class. In this service environment, intercepting of the location information is a serious threat on the privacy. For instance, when a malicious third party knows that a person is in a mental hospital, he/she is able to infer that the victim has a mental problem. If he/she informs the boss of the fact, the victim will be discharged. Moreover, in the ubiquitous computing environment, it is easier to intercept the message than the wired network, and the location information is acquired without the consent of the user from the intercepting message. The attacker can easily infer the context of the victim by collecting and analyzing the victim's location information within a certain period of time [2].

To protect location information, there are many researches on location privacy in various disciplines such as telemetric, mobile computing, and ubiquitous computing [3], [4], [5]. These researches try to control the usage of the location information. However existing methods have some limitations. They need complex computation to gain acceptable anonymity. Fewer subjects than the number desired by the approach do not support appropriate anonymity. Therefore, in this paper, we propose an effective approach that protects the privacy in ubiquitous computing environment. The proposed approach employs information diffusion method that scatters the user's location information in order to confuse the attacker. In addition, the base station or the access point transmits the dummy message that looks like a real traffic but has no meaning.

This paper is organized as follows. We describe private invasions, existing approach to handle it, and its limitations in section 2. Section 3 explains the proposed information diffusion approach in detail. This is followed by the describing the system architecture of the proposed method in section 4. Section 5 concludes.

2 Privacy Invasion in Ubiquitous Computing

2.1 Information Leaking vs. Information Gathering

In location-aware application, there are two kinds of method to hijack other's location information; one is unauthorized information access, called information leaking, another is inference with observation of the user's location, called information gathering. In the former case, the malicious user acquires the victim's location information from the system. In the latter case, the attacker is able to obtain location information by observing the traffic of a network, and then he/she infers what the victim does or how the victim's status is. Moreover, in this case, since both the victim and the system do not notice that the private information was indeed leaked, it is impossible to prohibit gathering the private information.

In case of the information leaking, the existing access control methods can be used to control the unauthorized access on the location information. Each system holds own policy to control the access on the information. Every request is sent to the authority that decides whether it is allowed or not. In addition to this, the monitoring system logs all the usage of the location information. When the unauthorized access is exposed, the logged data can be used to trace the malicious user.

In the case of the information gathering, since the eavesdropping of the signal does not detected by both the sender and the receiver, the malicious user is able to collect any user's location information without notice. Moreover, the malicious user can infer the locations and the user's context from the collected set of information. It is difficult to control the observation and inference attack properly. In this paper, we focus on the information gathering problem in order to prohibit the malicious user from obtaining the location information of others.

2.2 Previous Approaches Against the Privacy Invasion

M. Langheinrich introduces the requirements for the privacy in the ubiquitous computing environment [6]. It allows data collectors to both announce and implement data usage policies, as well as providing data subjects with technical means to keep track of their personal information as it is stored, used, and possibly removed from the system. E. Sneekenes identifies concepts that may be useful when formulating the privacy policy [7]. This policy makes that the individual should be able to adjust the accuracy depending on parameters such as the intended use and the identity of the recipient. Urs Hengartner et al. suggest the system that controls the access on the location information in the Wireless LAN based people location system [8]. It has the hierarchy of the location system and the access control mechanism delegates the request to the lower level system. After all, the victim user's device authenticates and authorizes the requester according to the victim's policy. The policy and mechanism are based on the SPKI/SDSI (Simple Public Key Infra Structure/Simple Distributed Security Infrastructure). P3P (Platform for Privacy Preference) provides negotiation between server and client [9]. At the negotiation, when a user wants to access the information, he/she receives the server's policy and compares it with the local policy. If the received policy is acceptable, he/she accesses to the information.

Bugra Gedik et al. introduce the cloaking method that provides vague temporal and spatial information in order to conceal a user within a group of k people [10]. The system employs an anonymous server act as the mix node. It prevents a malicious observer from linking ingoing and outgoing messages at the server. Different with any other k -anonymous system [5], it uses a customizable k that is changed by the environment. F. Stajano et al. solves the location privacy problem with the pseudonym that changes randomly and frequently [11]. It classifies every area into two zones; one is application zone where a user has registered for a service, another is mix zone that is a connected spatial region of maximum size in which none of users has registered for any services. Since applications do not receive any location information but pseudonym when users are in a mix zone, the identities are mixed.

Although there are many researches on the anonymity based location privacy protection method as we mentioned above, they have still some drawbacks. Indeed, it is very hard to prohibit the malicious user from gathering the location information. In the next section, this paper will discuss on the drawbacks on the previous approaches. That will give the reason why we should propose an effective method that enhances the location privacy.

2.3 Drawbacks on the Previous Approaches

Existing approaches provide the reasonable anonymity service if only there are enough subscribers in the network since they employ k anonymous system. If the number of subjects is smaller than the constant k , anonymity may be corrupted by the attacker.

In addition, there is large computational complexity since it changes the pseudonym frequently for the secure anonymity service. When a few subjects are in the anonymous area, the number of pseudonym change is not large since the number of location change is small. The attacker, however, may be able to perceive the pattern of pseudonym change quickly than we thought [12]. In practice, many subjects are in an area in order to achieve the reasonable anonymity. In this case, the system suffers from the extremely complex computation.

In practice, every person has his/her own pattern of movement. For example, a person, who is a salaried worker, is almost in the office at the morning. Between twelve and fourteen o'clock, he is in the restaurant for the lunch. At 4:00 p.m., he/she is in the meeting room with his/her co-worker. Thus the attacker can easily infer where the victim's signal comes from.

In the Mix zone, the area is represented as the combine of application zone and mix zone. The various users are in an area in general. When user A's registered service s and user B's registered services are different, there are two different application zones and mix zones are constructed. In practice, the subjects are generally more than 100 in the street. Then there may be too many different zones in the same area. It makes the system will forgive providing the anonymity service.

3 Location Information Diffusion and Confusion

3.1 Diffusing the Location Information

The basic idea of the diffusion is that it divides the information and inserts each fraction of the information into the normal packet. Most of service request packet size does not mount up to maximum packet size. Every packet may have many optional fields, which are not practically used, in the TCP/IP header parts. Moreover, the practical size of the internet packet is 576 bytes. That is, every packet has many unused space. The proposed system uses some parts of them. To do so, firstly the system chooses the multiple parts among the unused space. Both client and server should know the selected part consistently, but position should be changed whenever the private information is transmitted. Thus, we use selection table that stores selected part of unused space. Fig. 1 shows that how the position is selected with the table and how the information is diffused into a packet.

The size of the selection table is 8×8 . Each cell of the table contains a list of the unused positions in the packet. Every list includes three positions. Each position may be one of the following candidate positions; IP option field (40 bytes), TCP option field (40 bytes), and data field (1380 bytes). The proposed

Fig. 1. The diffusion of the location and related identity information

system divides them into 365 positions of 4 bytes. Each selected position can be represented with 9 bits ($365 < 2^9$). Since each cell contains 3 positions, the size of the cell is 27 bits that can be expressed as 1 word of the 32-bit machine.

Every list in the table is generated randomly by the position handling server. For a list, the server randomly selects 3 positions among 365 candidates whenever it creates the position selection table. To communicate with the server, the positions of the inserted parts are maintained consistently between the client and the server. Thus, in the proposed system, we employ the same position selection table in the both sides. The management issue will be discussed at section 3.3.

Fig. 2 shows the detailed process of the transmission of the location information. In order to select a cell having position list, the client calculates a 6-bits hash value from the time stamp. The client selects the row of the table by the first 3-bits of the hash value, and it selects the column of the table by the last 3-bits of the hash value. According to the selected list, the fractions of the location and the related identity information are inserted into a normal packet. The identity information is inserted into the packet as the form of hashed value, and the location information is inserted as the form of ciphertext that enciphered with the hashed value of the identity. The other empty positions are filled with random numbers.

When the service provider receives the packet from the client, it just transmits the packet to the position handling server. The position handling server extracts the location information, and it transmits the information to the service provider. The extraction process will be performed in the opposite order of the diffusion. The position handling server generates 6-bit hash value from the timestamp in the packet header, and then the position handling server selects a cell of the position selection table. According to the position list in the cell, the server gathers all the fragments of the location and identity information. The server checks whether the hashed identity exists or not, and then the server decrypts the location information with the hashed identity only if the identity exists. If the position handling server succeeds to extract, then it sends the location information to the service provider. Otherwise, it discards the message and notifies the fact.

Fig. 2. Detailed process of the transmission of the location information

3.2 Confusing the Malicious User

When a user's client sends the location information, the user's traffic can be easily captured or intercepted by the malicious user in general. To confuse the malicious user, anonymity approach has been used until now. The anonymity approach, however, has high computation complexity as we previously discussed. In the proposed approach, the location and identity information are already diffused in the normal packet. Even if the attacker intercepts the packet, he/she does not know whether the packet includes the location information. In addition, he/she does not know which positions are used for the location information. Thus, instead of the complex anonymity method, we employ simple confusing method that employs the dummy user instead of the anonymity system.

The dummy user is a part of the system that acts as the normal user. The dummy user can be a base station or an access point. They periodically emit the dummy message that has no meaning but acts as normal traffic. Thus the real traffic hides behind the floods of dummy traffic. It is difficult that the attacker distinguishes the real traffic from the dummy traffic. Fig. 3 shows the schematic view of the environment that employs several dummy users. In this figure, since the dummy user and the client send the message to someone, many packets move to somewhere else. Thus the status of the network is very dizzy.

To make the dummy message, the dummy user fills the contents of the packet with several random numbers. The address may be selected randomly among several neighbor hosts in the destination list. The destination list is managed by each dummy user. When the position handling server receives a message, it is unnecessary to determine whether it is normal message or not. Since the position handling server fails to extract the location and related identity information with the dummy message filled with random values, the dummy message will be discarded at the server. As shown in Fig. 3, the position handling server s0 is unsuccessful to extract the location information since the received message is a

Fig. 3. Dummy users on the networks; the confused area filled with numerous dummy messages. Dummy message is discarded at the server, while normal message is properly used for providing location information.

dummy message, while the position handling server s_1 succeeds in abstraction of the location information. The period of the emitting the dummy message is randomly changed.

3.3 The Position Selection Table Management

If the table is always the same, the malicious user can infer the selected position with the brute force analysis since the rate that selects the same position may comply with the uniform distribution. Although the malicious user does not know the location in real time, he/she can find out the set of the location. They collect that information for a day or a week, and then they can infer his private status such that he has a mental disease. Therefore the position selection table should be updated within reasonable amount of time t_{update} , called *update time*. That is, the proposed system will update the selection table if and only if $t_{used} \geq t_{update}$, where the t_{used} is the used time of the selection table and t_{update} is the update time of the table. For the decision of the update time t_{update} , there is no deterministic method but practical one. According to the service, the required degree of strength of the location privacy is various. The higher degree is required, the lower update time is decided.

The proposed system employs two table update methods; one is pull method, another is push method. In the push method, the position handling server transfers the new position selection table when a user uses the service. The detailed steps are as follows.

1. When the position handling server extracts the location information from a request, the server checks whether the used time t_{used} is larger than the predefined update time t_{update} .
2. If $t_{used} \geq t_{update}$, then the position handling server sends the table update message to the service provider. The service provider relays the message to the client who sent the request.

3. The client sends the ready message to the service provider. The service provider relays the message to the position handling server.
4. The position handling server creates a new selection table, and sends the table to the client via the service provider. The table is encrypted with the temporal symmetric key. The temporal key is a hash value from the combination of the location information and the user's id.
5. The client decrypts the received selection table and uses the table at the next position selection process.

In the pull method, according to the request of a user, the position handling server creates a new selection table, and then the server transfers it to the user's client. To transfer securely, the client and the service provider are temporally connected with the secure channel such as SSL, and the table should be encrypted by the user's password that is created by the user at the registration time.

Fig. 4. Schematic diagram of the proposed system; three thick white arrows mean the flow of the location information. The shaded hexagon represents the hidden location information, while white hexagon represents extracted location information.

4 System Architecture

Fig. 4 shows the schematic view of the proposed system. It consists of three components and two different communication environments. The client is a user's mobile device such as PDA, cellular phone, and handheld PC. The service provider is a server that supplies any service. Through the service provider, the client acquires necessary service such as location for a friend and showing the way to the destination. The position handling server authenticates and authorizes the requested user. It acts as trusted third party in the public key system. The confused area is located between the client and the service provider, and all the messages from them should be transferred through the area. It helps that the system hides the user's traffic. Both the service provider and the position handling server are connected with the secure channel such as SSL or VPN since the traffic between them contains the user's identity. The location information is diffused into a message from client to service provider. The location information is extracted by the position handling server and is sent to the service provider.

Every transaction, including the location and the related identity information, between the client and the provider is handled as following brief description. The client creates a message that notifies a user's location. It is diffused and inserted into the normal packet. Through any mobile network such as wireless LAN, CDMA, or GSM, the client sends the packet to the service provider, then the provider just passes the packet to the position handling server. In the position handling server, the location and the related identity information are extracted from the received packet, and then the information is transferred to the service provider.

5 Conclusion

In the ubiquitous environment, every service providing system may have the characteristic of the context-awareness. One of the most important information is the user's location information in order to aware his/her context. Thus the location information and the related identity are often sent to the service provider though the network. Since the attacker may intercept the information, sending the information without protection is very dangerous.

In this paper, we propose an effective approach to protect the location and related identity information. In the proposed approach, the sender diffuses the location information into the normal traffic and dummy user emits dummy message that confuses the malicious user for higher secure level. With this approach, the location information can be transferred to the service provider securely. In the future work, the diffusion will be applied to multiple packets, and the dummy user will act like a real user.

References

1. Garlan, D., Siewiorek, D.P.: Project Aura: Toward Distraction-Free Pervasive Computing, *IEEE Pervasive Computing* (2002) 22-31
2. Rodden, T., Friday, A., Muller, H., Dix, A.: A Lightweight Approach to Managing Privacy in Location-Based Services, *Proceedings of Equator Annual Conference* (2002)
3. Duri, S., Gruteser, M. et al.: Framework for Security and Privacy in Automotive Telematics, *WMC02, Atlanta, Georgia USA* (2002) 25-32
4. Gruteser, M., Grunwald, D.: Enhancing Location Privacy in Wireless LAN Through Disposable Interface Identifiers: A Quantitative Analysis, *WMASH03, San Diego, California USA* (2003) 46-55
5. Gruteser, M., Grunwald, D.: Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking. In *MobiSys, USENIX* (2003) 31-42
6. Langheinrich, M.: A Privacy Awareness System for Ubiquitous Computing Environments, In *Proceedings of the 4th International Conference on Ubiquitous Computing, LNCS No. 2498, Springer-Verlag* (2002) 237-245
7. Sneekenes, E.: Concepts for Personal Location Privacy Policies. In *Proceedings of the 3rd ACM Conference on Electronic Commerce* (2001) 48-57
8. Hengartner, U., Steenkiste, P.: Implementing Access Control to People Location Information, *SACMAT 2004, NewYork, USA* (2004) 11-20

9. Cranor, L., Langheinrich, M., Marchiori, M., Presler-Marshall, M., Reagle, J.: The Platform for Privacy Preferences 1.0 (P3P1.0) Specification. (2004) <http://www.w3.org/TR/P3P/>
10. Gedik, B., Liu, L.: A Customizable k-Anonymity Model for Protecting Location Privacy, CERCS Technical Reports, GIT-CERCS-04-15 (2004)
11. Beresford, R., Stajano, F.: Location Privacy in Pervasive Computing. Pervasive computing, IEEE CS and IEEE Communications Society (2003) 46-55
12. Beresford, A., Stajano, F.: Mix Zones: User privacy in location-aware services, Proceedings of First IEEE International Workshop on Pervasive Computing and Communication Security, A Workshop in PerCom 2004 (2004)