A Methodological Assessment of Location Privacy Risks in Wireless Hotspot Networks

Marco Gruteser and Dirk Grunwald

Department of Computer Science
University of Colorado at Boulder
Boulder, CO 80309
{gruteser, grunwald}@cs.colorado.edu

Abstract. Mobile computing enables users to compute and communicate almost regardless of their current location. However, as a side effect this technology considerably increased surveillance potential for user movements. Current research addresses location privacy rather patchwork-like than comprehensively. Thus, this paper presents a methodology for identifying, assessing, and comparing location privacy risks in mobile computing technologies. In a case study, we apply the approach to IEEE 802.11b wireless LAN networks and location-based services, where it reveals significant location privacy concerns through link- and application-layer information. From a technological perspective, we argue that these are best addressed through novel anonymity-based mechanisms.

1 Introduction

Pervasive computing promises near-ubiquitous access to information and computing services. Indeed, advances in mobile computing and networking, both wireless and the higher penetration of wired networks, have increased the spatial and temporal coverage of computing services. Users can access computing services from virtually anywhere, anytime.

Wireless networks also provide the ability to approximately track the location of users. Moreover, specialized location sensing technology has seen a dramatic reduction in price. These developments spurred an immense interest in exploiting this positional data through location-based services (LBS) [1–4]. For instance, LBS could tailor their functionality to the user's current location, or vehicle movement data could improve traffic forecasting and road planning.

The ability to track users' location, however, also creates considerable privacy concerns. For example, these privacy risks received attention through discussions about the use of IMSI Catchers [5], which can identify and locate GSM mobile phones, and several cases of monitoring the movements of rental cars through GPS receivers.[1]

[1] At least in one case [6], a rental car company used GPS technology to monitor the driving speed of their customers. When the company-set threshold speed limit of 79mph was exceeded, the company automatically charged a USD 150 speeding fee per occurrence on the customer's credit card. The fines were later found illegal, however the company is still allowed to track its cars [7].

Many different technologies are affected by these location privacy risks. Moreover, an adversary can potentially derive location information at different layers of the network stack. For example, at the physical layer through triangulation of the wireless signal, or at the network layer through DNS names of intermediate routers.

Effective privacy-enhancing technologies need to address these privacy risks according to their significance. At the data collection stage, this requires a framework and methodology for comparing and evaluating the privacy risks associated with different technologies. Specifically, this paper provides the following key contributions:

- A preliminary framework and methodology for identifying, comparing, and evaluating the significance of location privacy risks in network-related technologies.
- A case study that applies this methodology to IEEE 802.11b wireless LAN hotspot networks.
- A discussion of research directions to address the identified privacy challenges.

The remainder of this paper is structured as follows. Section 2 provides background in wireless networks, location-based services, and the resulting location privacy challenges. Section 3 then details the methodology for identifying and assessing location privacy risks. This methodology is applied to WLAN networks in Sect. 4. After covering related work that enhances location privacy in Sect. 5, we describe research directions for improving location privacy in WLAN networks in Sect. 6.

2 Background

In recent times, digital wireless networks revolutionized the telecommunications and networking industry. Wireless networks offer at least two key advantages over their wired counterparts. First, they allow for user mobility by untethering users from fixed communications stations such as desktop PCs or landline telephones. Second, they can achieve broad spatial coverage with less deployment effort due to the reduced need for installing cables.

Successful digital wireless networking standards include the Global System for Mobile Communications [8] for mobile phones and IEEE 802.11b [9] for wireless local area networks (WLAN). While the telecommunications industry is struggling to absorb the high initial investment costs for deploying higher bandwidth, packet-switched mobile phone networks (3G networks), WLAN technology has emerged as a cost-effective alternative for providing wireless network access in restricted areas.

WLAN hotspots are deployed in many residential homes, universities, and industry campuses. More recently, commercial wireless access services have also been offered to the public at coffee shops, airport lounges, and hotels. The term hotspot illustrates that WLAN provides higher bandwidth than 3G networks—11 Mbps compared to a target of 2 Mbps for 3G networks in Europe—but focuses the signal on a much smaller area of interest, typically with a radius of approximately 150 feet. Thus, WLAN networks achieve less spatial coverage than mobile phone networks, but concentrate the coverage on key areas. However, researchers (e.g. Negroponte [10]) envision WLAN access points to route messages between different hotspots. Thus, access points would form a mesh network that greatly increases coverage beyond a single WLAN network. Larger

distances between wireless networks could be bridged through wired connections or specialized antenna equipment.

2.1 Location-based Services

Mobile computing has enabled users to access computing services at many different locations. In addition, location sensing or positioning technology allows automated tracking of users' positions. Particularly well-known is the Global Positioning System [11], which determines location through reference signals emitted by satellites. In 2002, Motorola unveiled a GPS chip that is small and cost-effective enough to be included in a wide range of consumer devices. The chip only measures about 50 square millimeters and costs $10 in volume quantities [12]. According to Tim McCarthy, business director for GPS at Motorola's Automotive Group's Telematics Division, position awareness has a bright future: "All of a sudden, starting 10 or 15 years ago, every electronics device had a clock. I see position awareness going down that same path. It's just a question of how long it takes" [12]. Ambuj Goyal, IBM, goes further in predicting prices for GPS chips to drop to a few cents [13]. Many other location sensing technologies are described in a survey by Hightower [14]. In the United States, widespread deployment of such technologies is encouraged by the Federal Communications Commission since it mandated that locating mobile phones must be possible for emergency purposes [15].

Thus, location-based services (LBS) have emerged, which combine positioning technology with the mobile Internet. Location information becomes an additional input parameter to customize the service. For example, the Webraska Corporation [16] offers applications ranging from navigational services that provide driving directions over point-of-interest or accommodation finders to automotive fleet management. These applications transmit the user's current position over a wireless network to the server hosting the location-based service. Location information also proved useful for customizing functionality to the user's current situation. For example, context-aware tour guides [17] automatically adjust the presented information and menu options to the user's current location.

2.2 Location Privacy

These advances in location sensing and communication technology have significantly decreased the effort to track an individual's movements. Historically, the current location and the history of movements of an individual were little known to other parties. Only through cooperation of the individual or intensive investigative effort could this information be obtained. Today, positioning technologies can easily determine a subject's position, improved database and storage technology permits permanent recording and wide distribution of this data, and data mining enables the detection of movement patterns.

Not surprisingly, information privacy concerns have mounted globally [18–20]. Issues range from detailed, publicly available satellite imagery over data collection on the Internet to DNA databases. In surveys, consumers reiterate their concern for privacy. For example, according to one survey 94% of web user have denied a request for personal information and 40% have provided fake data [21].

In the United States, Privacy risks related to *location* information have been identified in the Location Privacy Protection Act of 2001 [22]. While public disclosure of location information enables a variety of useful services, such as improved emergency assistance, it also exhibits significant potential for misuse. For example, location information can be used to *spam* users with advertisements or to learn about users medical conditions, alternative lifestyles, or unpopular political views. Inferences can be drawn from visits to clinics, doctors' offices, entertainment districts, or political events. Such conclusions can be particularly annoying for subjects if they are inaccurate. In extreme cases, public location information can lead to physical harm such as in stalking or domestic abuse scenarios. Karger and Frankel provide a more detailed discussion of security and privacy risks in Intelligent Transport Systems [23].

Location information is valuable for location-based services because it implicitly conveys characteristics that describe the situation of a person. However, the foregoing examples illustrate how adversaries can exploit the same information to cause harm to a person. Phil Agre also warns us of such location privacy issues [24]. Specifically, he is concerned about a widespread deployment of automatic face recognition technology. He fears "spottings markets" that trade information about the times and locations where people have been identified.[2] Wireless networking could provide an even easier means for spotting people.

3 Methodology

A location privacy threat describes the risk that an untrusted party can locate a transmitting device *and* identify the subject using the device. We make the assumption that a

[2] Excerpted from Phil Agre [24]: "My candidate for Privacy Chernobyl is the widespread deployment in public places of automatic face recognition. [...] And that's just the start. Wait a little while, and a market will arise in "spottings": if I want to know where you've been, I'll have my laptop put out a call on the Internet to find out who has spotted you. Spottings will be bought and sold in automated auctions, so that I can build the kind of spotting history I need for the lowest cost. Entrepreneurs will purchase spottings in bulk to synthesize spotting histories for paying customers. Your daily routine will be known to anyone who wants to pay five bucks for it, and your movement history will determine your fate just as much as your credit history does now. Prominent firms that traffic in personal movement patterns will post privacy policies that sound nice but mean little in practice, not least because most of the movement trafficking will be conducted by companies that nobody has ever heard of, and whose brand names will not be affected by the periodic front-page newspaper stories on the subject. They will all swear on a stack of Bibles that they respect everyone's privacy, but within six months every private investigator in the country will find a friend-of-a-brother-in-law who happens to know someone who works for one of the obscure companies that sells movement patterns, and the data will start to ooze out onto the street.

Then things will really get bad. Personal movement records will be subpoenaed, irregularly at first, just when someone has been kidnapped, but then routinely, as every divorce lawyer in the country reasons that subpoenas are cheap and not filing them is basically malpractice. Then, just as we're starting to get used to this, a couple of people will get killed by a nut who been [sic] predicting their movements using commercially available movement patterns. Citizens will be outraged, but it will indeed be too late ..."

subject sporadically sends and receives messages, likely from different locations in the area covered by the wireless network. The adversary seeks to obtain information from these messages. An assessment of location privacy risks should proceed accordingly. Based on the data subject's messages, it should analyze how location information can be obtained, how originators can be identified, and who has the means to do so.

3.1 Locating

The originator can be located through a variety of mechanisms, for example eavesdropping when the originator explicitly reveals location information to a LBS or triangulating the wireless signal. A careful analysis of the information contained in each layer of the network stack reveals possible approaches for location determination. These approaches are then characterized according to the following privacy-enhancing criteria.

User Choice. Location systems differ widely on how much control a user has over the system. An ideal system would allow users to hide while still providing full functionality; however, in practice a user usually experiences inconveniences. For example, the signal of a mobile phone may be located through triangulation. The user can prevent this by switching the phone off; however, the user is then unable to receive or originate phone calls.

Restricted Coverage. Location sensing technologies are often restricted to function only in certain areas (spatial coverage) or during certain times (temporal coverage). Spatial coverage can range from near universal for the GPS system to occasional for a credit card-based location tracking. In the credit card case, location information is only available at point of sale terminals, which are sparsely distributed over populated areas. Furthermore, it is restricted to the occasions when a subject uses a credit card at the terminal.

Lower Resolution and Accuracy. In their areas of coverage, location sensing technologies achieve different resolutions. Higher resolution conveys significantly more information. Consider a system that achieves 1 km resolution versus a system with 1 meter resolution. The first system reveals information such as the city name and district. The latter system can additionally disclose the exact building and even room, where a subject is located.

3.2 Identifying

Identification of the subject means that an adversary learns the real-world name of a subject, such as legal name and address of residence. Network addresses, for instance, are not necessarily considered identifying information, since not every party can correlate a network address to a real-world name and address. Rather, we consider network addresses *pseudonyms*, information that can help identifying a subject.

Note that locating and identifying are not completely independent tasks, since distinct location information helps in identifying subjects. Assume that a user does not disclose her identity but includes precise location information in a transaction. The recipient could correlate the location with location information obtained through other means to identify the user. For example, when the transaction originates from a private

residential driveway, it can be easily linked to a public address database. This likely reveals the originator and violates location privacy. Even if the location information itself is less sensitive, it can be used to link other private information (e.g., the content of the transaction) to the user.

Location information can identify the sender of an otherwise anonymous message, if the information is correlated with public knowledge or observations about a subject's location. Consider the case where a subject sends a message M to a location-based service and an adversary A gains access to the subject's location information L. Then, sender anonymity is threatened by location information in the following ways.

Restricted Space Identification. If A knows that space L exclusively belongs to subject S then A learns that S is in L and S has sent M. For example, consider the owner of a suburbian house sending a message from his garage or driveway. The coordinates can be correlated with a database of geocoded postal addresses, such as provided by Geocode [25], to identify the residence. An address lookup in phone or property listings then reveals the owner and likely originator of the message.

Observation Identification. If A has observed the current location L of subject S and finds a message M from L then A learns that S has sent M. For example, the subject has revealed its identity and location in a previous message and then wants to send an anonymous message. The later message can be linked to the previous one through location information.

These identification approaches require high-resolution location information. Specifically, the resolution needs to be high enough to distinguish a subject from other persons and to pinpoint her at a restricted space or to uniquely match the observation.

A more sophisticated identification approach depends on the ability to link messages, but allows lower resolution location information. Linking messages means determining that two or more message stem from the same originator, whose identity is not necessarily known. The approach accumulates location data over periods of time to prepare a movement profile. For example, an adversary could learn from the profile that the morning commute takes a subject from a certain suburb to a certain work location. By filtering residential addresses and work address the subject might be identified.

The ability to link messages provides another advantage for the adversary. If the originator of a set of messages is already identified it can be easier to link new messages to the existing set than to use other identification mechanisms on the new messages. One technique for linking messages is based on pseudonyms. If two messages contain the same pseudonym, they most likely stem from the same originator.

If the subject transmits her location with high resolution and frequency, the adversary can, at least in less populated areas, also link messages based on physical constraints and knowledge of the area. For example, maximum speeds of data subjects are known. Furthermore, an adversary might use road maps, which describe likely travel paths and give clues about the expected speed. The adversary could then link messages received from different locations based on this spatio-temporal analysis.

3.3 Data Collectors

Finally, an important consideration is *who* has access to the location data. We characterize the data collectors as follows:

Dispersion. Measurements stations can be geographically distributed and potentially belong to different organizations. This increases the effort of collecting data, because for effective cross-organizational surveillance, data must be compiled into a central database. We characterize relationships according to the degree of cooperation and data sharing between different network operators or LBS providers.

Trust and Recourse. Trust relationships between data subjects and service providers can differ substantially. Here, we consider only trust related to location privacy, not other concerns such as reliability or security. The relationship can range from complete mistrust to a contractual relationship. In the complete mistrust scenario, the network operator is likely unknown to the data subject or even a known privacy offender. A legally binding privacy contract with a reputable company typically establishes a higher level of trustworthiness compared to an unknown provider.

4 Wireless LAN Risk Assessment

In a case study, we apply the described methodology to a wireless LAN and location-based services environment. The following subsections provide an analysis of possible approaches to determine location information, link messages, or identify subjects.

4.1 Determining Location

Resolution. Most easily, location information can be obtained, when the subject explicitly includes it in the messages. For example, the current location of the subject, as determined by a GPS receiver, could be sent to a LBS that provides driving directions to a specific destination. In this case, both LBS providers and network operators have access to the location information. However, even if location information is not explicitly included in the message, the following mechanism can gather it.

Coarse location information can be obtained at the network layer. While IP-addresses are not assigned according to geographic locations, the hierarchical structure leads to a geographic concentration of address ranges. For example, all addresses within the subnets assigned to a university campus are typically collocated. However, proxies and firewalls are a common obstacle, since they can hide the true address.

The techniques to determine the location of regular internet hosts can be categorized dependent on their information source: DNS name clues and registration information, network delay measurements [26], and web service registrations. DNS names of routers usually include geographic clues such as city names. Therefore *traceroute*-based mechanisms [27] reveal an approximate location. Another widely used approach queries WHOIS servers for the registered contact address information. The geoclustering algorithm [26] combines a clustering technique based on routing information with location information gathered from user registrations at popular web services. Its median error distance varies from below 50 kilometers to several hundred kilometers.

At the physical and link layer, access points can estimate the position of a transmitter based on proximity. If packets are received from the transmitter, it must be within range of a typical 802.11b system - around 50-100 meters. Higher resolution is provided by triangulation mechanisms of which several systems for indoor WLAN installation have been developed [28–30]. They measure the signal strength of received packets to determine position. Together with Bayesian signal processing approaches, resolutions up to 1m can be achieved. However, it is not clear, whether these high resolutions can also be obtained in outdoor mesh networks because the distance between WLAN base stations is likely greater than in an indoor setting.

Coverage. Information included in the application layer and network layer are visible to servers and all parties whose routers a packet traverses. However, only parties in range of the transmitter can access information in the link and physical layer. Therefore, accessibility to link layer information is typically restricted to a smaller set of wireless network operators. However, these risks cannot be ignored considering the astonishing density of access points found in major metropolitan areas. Figure 1 shows the density of access points in downtown Chicago. The map was obtained from the Wireless Geographic Logging Engine [31], which collects and merges data from the *Netstumbler* WLAN monitoring software [32]. Judging by the density of access points, it is difficult to use a WLAN network without being detected by other parties.

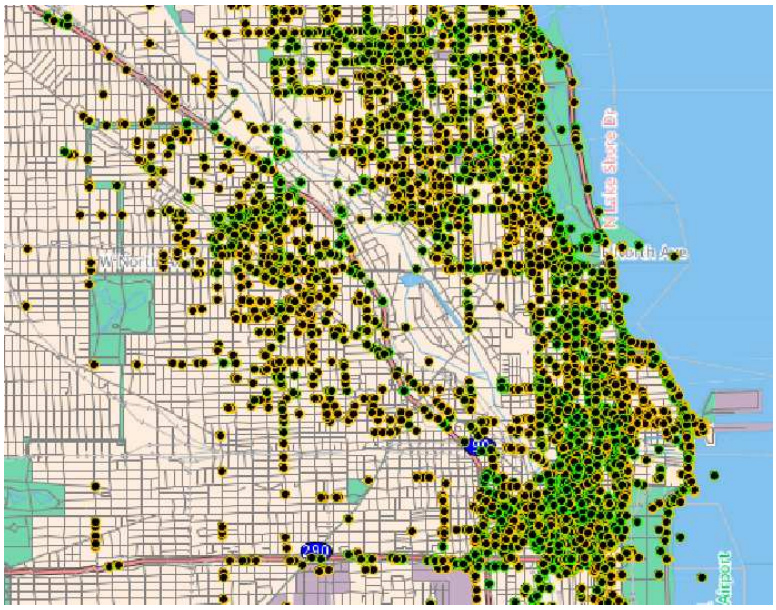

Fig. 1. IEEE 802.11b Access Points in Downtown Chicago

User Choice. Most of the described mechanisms do not offer the user a choice whether he allows others to locate the device. For the physical and link layer approaches, the choice is made implicitly, when the user activates the wireless network card. Even when the card is not actively used for data traffic, it will transmit beacons that are sufficient for the location determination mechanisms. At the network layer, the choice is similarly implicit. When the user communicates with any network entity, his network address will be visible to intermediate routers and the communication partner.

GPS location information differs in that position is calculated in a client-side receiver. No signals are emitted from this receiver, which would enable remote parties to determine the location of the receiver. In the GPS case, the user can make an explicit choice to disclose the locally determined location information to another party, for instance to a LBS.

4.2 Identifying a Subject

Pseudonym-based approaches for linking messages are possible on all layers of the network stack. At the application layer user or login names help in linking messages to the same originator, even if the user hides his real identity. The IP-address also provides an easily accessible pseudonym at the network layer. However, IP-addresses tend to change more frequently than service logins, because many clients are configured with dynamic IP addresses, which are assigned on a temporary basis through the Dynamic Host Configuration Protocol. More static identifiers are also available at the link layer, where MAC addresses provide a simple way to identify packets originating from the same transmitter.

4.3 Data collectors

As illustrated in figure 2, several entities can be distinguished to clarify whose location privacy should be protected from which adversaries:

Data subject. The data subject is the person to whom the private information relates. The subjects access services over the network through client devices such as laptops, PDAs, or in-car PCs. They typically move among different cells of the wireless network as they go about their daily life.

Wireless Network Operator. Wireless network operators maintain base stations, which form the wireless network. An operator can maintain any number of base stations. Some operators will only have a single base station, while others set up multiple clusters of base stations. Wireless network operators have access to network and application layer information if they act as routers. In addition, they can also access link and physical layer information if packets are directly received from the client device. Even if packets are intended for another access point, wireless network operators may be able to eavesdrop on the information.

Location-based service providers. Location-based services can be provided by arbitrary servers connected to the wireless or wired network. Over the network, LBS providers can typically access only network and application layer information. However, they may receive partial location information from other sources.

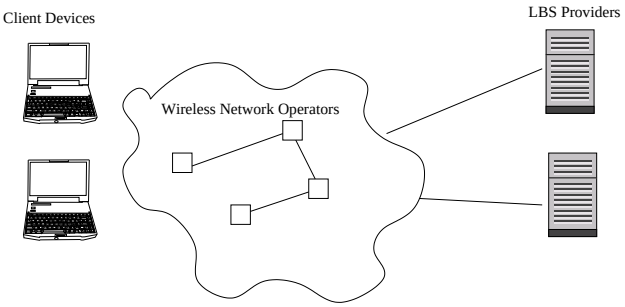

Fig. 2. Architectural context

The same party might exhibit multiple roles. For example, a person could install a set of base stations, provide a location-based service, and use the wireless network himself. However, more typically, different companies act as the LBS provider and the wireless network operators.

Since wireless access points are relatively inexpensive, many different parties will act as wireless network operators. Thus, trust relationships will also differ. In some cases, a formal contract may be established between data subjects and the network operators. This contract is likely part of the service agreement under a subscription-based plan. Such a plan typically provides wireless coverage at lucrative hotspot locations, for example in coffee shops.[3] In these cases, service providers may be trustworthier. However, network operators can also enter into roaming agreements, which allow data subjects to use the same access plans on different networks, but also to share collected data. Additionally, data subjects are then likely to use services from network operators without a direct contractual agreement. Moreover, we expect also many little-known entities to act as wireless network operators. Service and privacy will most likely not be governed by a contract. The relationship between data subjects and LBS providers exhibits a similar range of trust characteristics.

Unfortunately, WLAN location privacy issues are further complicated because other parties can easily overhear communications. Everybody can install an access point and thus act as a network operator. Even when data subjects' wireless cards are not associated with the access point, link layer information is visible. In fact, not even an access point is required for monitoring wireless users. Some wireless cards, those based on Prism chipsets, can emulate an access point through software and thus provide the ability to monitor other subjects. When such access points are operated in the vicinity of medical clinics or other privacy sensitive areas, major privacy issues arise.

4.4 Summary

Table 1 summarizes the location determination mechanisms. The most accurate location information is either included at the application layer or can be determined through triangulation at the link layer. Based on this information, identification of a subject is

[3] Such as the service delivered by T-Mobile to many US Starbucks coffee shops

possible for both wireless network operators and LBS providers (if the subject chooses to reveal the location). This leads to significant privacy concerns, especially because of the ease with which untrusted network operators can overhear and locate a subject's communications.

Method	Accuracy	Coverage	Choice
GPS	10m	Near-universal	explicit
IP-Address, DNS	50km +	address dependent	implicit
WLAN Proximity	50-200m	Densely populated areas	implicit
WLAN Triangulation	1-10m	Densely populated areas	implicit

Table 1. Characterization of location determination mechanisms

5 Related Work

Prior work on privacy aspects of telematics and location-based applications has mostly focused on a policy-based approach [33, 34]. Data subjects need to evaluate and choose privacy policies offered by the service provider. These policies serve as a contractual agreement about which data can be collected, for what purpose the data can be used, and how it can be distributed. Typically, the data subject has to trust the service provider that private data is adequately protected. In contrast, the anonymity-based approach depersonalizes data before collection, thus detailed privacy-policies and safeguards for data are not critical. Specifically, the IETF *Geopriv* working group is addressing privacy and security issues regarding the transfer of high resolution location information to external services and the storage at location servers. It focuses on the design of protocols and APIs that enable devices to communicate their location in a confidential and integrity preserving manner to a location server.

Snekkenes [35] presents concepts for specifying location privacy policies. These concepts enable access control decisions based on the identity of the requestor and purpose of the request. In addition, time, location, speed, and identity of the located object influence the decision. In the policy specifications, accuracy of temporal, location, and identity information is modeled using a lattice. However, the author concludes by expressing doubt that the average user will specify such policies.

The Mist routing project for mobile users [36] combines location privacy with communication aspects. It focuses on the problem of routing messages to a subject's location while keeping the location private from the routers and the sender. To this end, the system is comprised of a set of mist routers organized in a hierarchical structure. The leaf nodes have knowledge of the location of users but not their identity. They refer to them through handles (or pseudonyms). Each user selects a higher-level node in the tree, which acts as a semi-trusted proxy. It knows the identity of the user but not his exact location. The paper then presents a cryptographic protocol to establish connections between users and their semi-trusted proxies and mechanisms to connect to communication partners through their proxies. The paper does not address the problem of sending anonymous messages to external location-based services.

Mobile IP enables hosts to transparently migrate between different networks by registering the current location with a home agent that tunnels all traffic to the current network. Thus, adversary can track the location of a host by observing the registration messages and the payload messages through the tunnel. The Non-Disclosure-Method [37] method places several rerouting security agents between home and foreign network. Security agents forward messages in encrypted form; therefore, it is hard to trace the path of a message if the security agents are spread over several administrative domains. This method hides the care-of network address of a host from intermediary routers, but it does not address explicit application layer information, provide anonymity with respect to a server, or anonymity with respect to the foreign agent.

Narten and Draves propose privacy extensions for stateless address autoconfiguration in IPv6 [38]. Server-less address autocofiguration in IPv6 can use the MAC address of network interfaces as part of the network layer IP-address. Thus the MAC address becomes visible to servers and intermediaries outside the local area network. This enables such outside entities to track movements of mobile nodes between different networks. The proposed solution switches the network address periodically. New addresses are generated through iterative applications of the MD5 message digest algorithm on the previous network address and the actual MAC address.

Location privacy has also influenced the design of location sensor systems. The Cricket system [39] places location sensors on the mobile device as opposed to the building infrastructure. Thus, location information is not disclosed during the position determination process and the data subject can choose the parties to which the information should be transmitted. However, this solution does not provide for anonymity. Similarily, Smailagic and Kogan [40] addressed privacy in a WLAN location sensing system.

6 Research Directions

We believe that the multi-faceted issue of location privacy must be addressed through a variety of means encompassing legislative, self-regulation, and technological approaches. This discussion focuses on the technological approaches at the data collection stage. As alluded to before, an invasion of location privacy requires that an untrusted party can locate a person *and* identify the person. This suggests at least three alternative approaches to enhance location privacy. First, the data subject establishes trust in the unknown party before it reveals location information, for example through an exchange of privacy policies and privacy preferences. Second, the untrusted party can learn the identity of the subject, but is unable to locate the subject. And third, the untrusted party can locate a subject, however, the subject remains anonymous.

The first approach requires the data subjects to read, understand, and evaluate the privacy policy of every service provider, before they interact with this service. Especially for spontaneous usage of services, this poses a significant burden on the data subject. Moreover, a privacy policy does not provide protection of the data. For example, company insiders could steal private data or data might be inadvertently disclosed over the Internet during software maintenance operations. If a company wishes to employ data protection technology, this significantly complicates the computing architecture

and poses a high processing overhead. This approach is most appropriate for services offered by a well-known, reputed company, with which a customer whishes to engage on a longer-term basis. It cannot address privacy issues related to access points operated by unknown parties.

The second approach does not satisfy the requirements of most LBS, since a LBS needs to receive location information to perform its function. However, it could be useful for certain classes of services that work with aggregate data from a large number of individuals. The individuals could hide their true location information through data perturbation, while the service provider can still draw inferences about the distribution of values over a large population. Agrawal and Srikant [41] demonstrated such an approach for data mining applications. However, it is unclear whether the population size would be large enough in the context of LBS. In addition, it is difficult to hide location information from network operators, since they can have access to physical or link layer information. Directional antenna designs could provide a degree of protection, but would require hardware changes to wireless network cards.

The anonymity-based approach seems most promising in the WLAN and LBS context. If data remains anonymous, privacy policies and technological safeguards are not necessary. In addition, service users can reveal their true location provided that the data is not sufficiently distinctive to disclose the identity of the user. Location information could be slightly perturbed, so that it preserves anonymity but is still useful for LBS. Protection from network operators could be achieved through address switching approaches. When network and MAC address change, linking of messages becomes more difficult. Thus, network operators would be unable to track users' movements. However, address switches must be well timed so that network operators cannot link new and old addresses based on location information. If a subject is relatively alone and stationary, address switching might not be effective.

7 Conclusions

This paper presents an approach for identifying and evaluating location privacy risks in network technologies. The methodology distinguishes between location determination mechanisms, identification mechanisms, and the trust relationships between the data subjects and potential data collectors. It characterizes location determination mechanisms according to data resolution, coverage, and user choice.

In the WLAN case study, we found obtaining high-resolution location information relatively easy, provided that the user activates his network card. In addition, access points are densely distributed in highly populated areas, and in many cases, the operators of access points are not known to the data subjects and thus not trustworthy. This causes significant location privacy concerns.

Specifically, we draw the following conclusions. First, accurate location information obtained through GPS or WLAN triangulation can be sufficiently distinctive to identify subjects. Second, the presented methodological approach proved useful in identifying and comparing the location privacy risks in the WLAN case study. WLAN exhibits significant privacy risks because of the availability of inexpensive hardware and the determination of position with high resolution. Finally, anonymity mechanisms that

reduce the identification risks of location information itself and hide identifiers from untrusted access points are a promising research direction.

Acknowledgments

Paul Chou and my colleagues at the IBM T.J. Watson Research Center encouraged me to research location privacy challenges. The anonymous referees provided useful comments on a draft of this paper.

References

1. Mike Spreitzer and Marvin Theimer. Providing location information in a ubiquitous computing environment. In *Proceedings of the 14th ACM SOSP*, pages 270–283, 1993.
2. Andy Harter, Andy Hopper, Pete Steggles, Andy Ward, and Paul Webster. The anatomy of a context-aware application. In *Mobile Computing and Networking*, pages 59–68, 1999.
3. Rui Jose and Nigel Davies. Scalable and flexible location-based services for ubiquitous information access. In *In Proceedings of First International Symposium on Handheld and Ubiquitous Computing, HUC'99*, pages 52–66. Springer Verlag, 1999.
4. C. Bisdikian, J. Christensen, J. Davis II, M. Ebling, G. Hunt, W. Jerome, H. Lei, and S. Maes. Enabling location-based applications. In *1st Workshop on Mobile commerce*, 2001.
5. Dirk Fox. Der imsi-catcher (in german). *Datenschutz and Datensicherheit*, 21(9), Sep 1997.
6. Robert Lemos (ZDNet News). Car spy pushes privacy limit. http://zdnet.com.com/2100-11-530115.html?legacy=zdnn, Jun 2001.
7. Robert Lemos (ZDNet News). Car rental gps speeding fines illegal. http://www.mvcf.com/news/cache/00400/, Jul 2001.
8. M. Rahnema. Overview of the gsm system and protocol architecture. *IEEE Communications Magazine*, 31(4):92–100, 1993.
9. IEEE. IEEE standard 802.11b - wireless LAN medium access control (MAC) and physical layer (PHY) specications: High speed physical layer(PHY) in the 2.4 GHz band, 1999.
10. Nicholas Negroponte. Being wireless. *Wired Magazine*, Oct 2002.
11. I. Getting. The global positioning system. *IEEE Spectrum*, 30(12):36–47, December 1993.
12. John Spooner (CNET News). Motorola: New chip will bring gps to all. http://news.com.com/2100-1040-959085.html, Sep 2002.
13. Ambuj Goyal. In talk given at NIST pervasive computing conference. 2001.
14. J. Hightower and G. Borriello. A survey and taxonomy of location sensing systems for ubiquitous computing. UW CSE 01-08-03, University of Washington, August 2001.
15. J. Reed, K. Krizman, B. Woerner, and T. Rappaport. An overview of the challenges and progress in meeting the e-911 requirement for location service. *IEEE Personal Communications Magazine*, 5(3):30–37, April 1998.
16. Webraska Mobile Technologies. Webraska website. http://www.webraska.com/.
17. Keith Cheverst, Nigel Davies, Keith Mitchell, and Adrian Friday. Experiences of developing and deploying a context-aware tourist guide: the guide project. In *Proceedings of MOBICOM*, pages 20–31. ACM Press, 2000.
18. The Economist. The end of privacy, 29th Apr 1999.
19. The Economist. The coming backlash in privacy, 9th Dec 2000.
20. Michael Froomkin. The death of privacy? *Stanford Law Review*, 52:1461–1543, May 2000.
21. Donna L. Hoffman, Thomas P. Novak, and Marcos Peralta. Building consumer trust online. *Communications of the ACM*, 42(4):80–85, 1999.

22. Location privacy protection act of 2001. US Congress, Sponsor: Sen. John Edwards (D-NC), Contact: Maureen Mahon, Legislative Assistant, Sen. Edwards / 202.224.3154 / fax 202.228.1374, http://www.techlawjournal.com/cong107/privacy/location/s1164is.asp, 2001.
23. P. A. Karger and Y. Frankel. Security and privacy threats to ITS. In *Proceedings of the Second World Congress on Intelligent Transport Systems*, volume 5, Yokohama, Japan, Nov 1995.
24. Phil E. Agre. Red rocks eater news service — notes and recommendations. http://commons.somewhere.com/rre/1999/RRE.notes.and.recommenda14.html, Dec 1999.
25. Tele Atlas North America, Inc. Geocode website. http://www.geocode.com/.
26. Venkata N. Padmanabhan and Lakshminarayanan Subramanian. An investigation of geographic mapping techniques for internet hosts. *Proceedings of SIGCOMM'2001*, page 13, 2001.
27. Ram Periakaruppan and Evi Nemeth. Gtrace —a graphical traceroute tool. In *13th Usenix Systems Administration Conference — LISA*, Seattle, WA, Nov 1999. Nov 7-12.
28. Paramvir Bahl and Venkata N. Padmanabhan. RADAR: An in-building RF-based user location and tracking system. In *INFOCOM (2)*, pages 775–784, 2000.
29. Paul Castro, Patrick Chiu, Ted Kremenek, and Richard Muntz. A probabilistic room location service for wireless networked environments. In *Proceedings of Ubicomp*, Atlanta, GA, Sep 2001.
30. Andrew M. Ladd, Kostas E. Bekris, Algis Rudys, Lydia E. Kavraki, Dan S. Wallach, and Guillaume Marceau. Robotics-based location sensing using wireless ethernet. In *Proceedings of MOBICOM*, pages 227–238. ACM Press, 2002.
31. Wireless geographic logging engine. http://wigle.net/gpsopen/gps/GPSDB/, Oct 2002.
32. Netstumbler software. http://www.netstumbler.com, Oct 2002.
33. J. Cuellar, J. Morris, and D. Mulligan. IETF Geopriv requirements. http://www.ietf.org/html.charters/geopriv-charter.html, 2002.
34. Sastry Duri, Marco Gruteser, Xuan Liu, Paul Moskowitz, Ronald Perez, Moninder Singh, and Jung-Mu Tang. Framework for security and privacy in automotive telematics. In *Proceedings of the second international workshop on Mobile commerce*, pages 25–32. ACM Press, 2002.
35. Einar Snekkenes. Concepts for personal location privacy policies. In *Proceedings of the 3rd ACM conference on Electronic Commerce*, pages 48–57. ACM Press, 2001.
36. Jalal Al-Muhtadi, Roy Campbell, Apu Kapadia, M. Dennis Mickunas, and Seung Yi. Routing through the mist: Privacy preserving communication in ubiquitous computing environments. In *International Conference of Distributed Computing Systems*, 2002.
37. A. Fasbender, D. Kesdogan, and O. Kubitz. Analysis of security and privacy in mobile IP. In *4th International Conference on Telecommunication Systems Modeling and Analysis*, Nashville, TN, Mar 1996.
38. T. Narten and R. Draves. RFC3041—privacy extensions for stateless address autoconfiguration in ipv6. http://www.faqs.org/ftp/rfc/rfc3041.txt.
39. Nissanka B. Priyantha, Anit Chakraborty, and Hari Balakrishnan. The cricket location-support system. In *Proceedings of the sixth annual international conference on Mobile computing and networking*, pages 32–43. ACM Press, 2000.
40. Asim Smailagic and David Kogan. Location sensing and privacy in a context-aware computing environment. *IEEE Wireless Communications*, 9:10–17, oct 2002.
41. Rakesh Agrawal and Ramakrishnan Srikant. Privacy-preserving data mining. In *Proc. of the ACM SIGMOD Conference on Management of Data*, pages 439–450. ACM Press, May 2000.