RC22781 (W0304-163) April 16, 2003 (publicly available September 2003)
Computer Science

IBM Research Report

A Practical Approach to Location Privacy in Public WiFi Networks

M. T. Raghunath, Chandrasekhar Narayanaswami
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 704
Yorktown Heights, NY 10598

A Practical Approach to Location Privacy in Public WiFi Networks

M. T. Raghunath and Chandra Narayanaswami

IBM TJ Watson Research Center
19 Skyline Drive, Hawthorne NY 10532

Abstract. The rapid growth in the number of public hotspots offering wireless connectivity enables tracking the location of a mobile subscriber at a much finer level compared to other widely deployed technologies. While businesses may seek to convert fine-grain location information into valuable services, subscribers may not want their location revealed. We propose a practical location privacy solution, which can be readily understood by a non-technical subscriber. We also believe that our solution is attractive enough to be offered by service providers keen on using privacy as a competitive advantage. Our solution is based on decoupling user identity from device identity by relying on real-world mechanisms that provide anonymity. We believe this is a simple and practical step along the path to realizing Mark Weiser's vision of ubiquitous computing.

1 Introduction

Over the last few years we have witnessed the growth of wireless LANs from availability in university and technology companies to widespread availability in other enterprises and in public places such as trains, busses, airports, coffee shops, fast-food restaurants, etc. Recently, commercial airlines such as Lufthansa and British Airways have begun offering this technology on certain trans-Atlantic flights. It can be argued that high speed wireless LANs are one of the most significant developments in mobile computing in recent times. The relatively low cost of 802.11 hardware has made it attractive for several people to deploy a wireless network even in their homes, especially if they have a broadband connection to the internet. Several companies such as Boingo™, Wayport™, Tmobile™, Cometa™, etc., offer various plans for nationwide (USA) 802.11 wireless access through hotspots distributed across the country. Typical plans include one hour, one day, ten day, unlimited for a month, and other forms of metered access. Current service costs also appear to be affordable. Recent laptop computers offer built-in 802.11 interfaces. Handheld computers may soon follow with built-in 802.11 interfaces as well.

Overall, WiFi networks have changed the way business professionals work. Just as the cell phone helped liberate people from the land line, WiFi access is helping liberate people from wired networks. People are less tied to their desks. Instead of having

to carry network cables and finding seats next to network jacks in meeting rooms, people can sit anywhere they like. Business travelers may synchronize their email or download information from their corporate intranet during lunch at a restaurant. WiFi networks at airports and inside planes are likely to help travelers stay in touch and get more work done during their travel.

The above benefits and flexibility of ubiquitous and affordable wireless access in public spaces also come with certain questions. Questions of privacy and security are two key issues. Who else in the public space can see the data that is being sent to you? Can the service provider constantly track your physical location? Can the service provider build a profile of the web sites you visit? How much of your privacy do you need to give up in order to benefit from these services?

The initial security mechanism for 802.11 networks, called WEP, turned out to have serious problems [1] rendering it largely ineffective as a security mechanism. Vendors have developed several proprietary mechanisms to mitigate the security loopholes of WEP [2]. Security initiatives such as 802.1X [3] are currently underway in the standards bodies [4].

While security is a closely related topic, the main focus of this paper is privacy of the users of public WiFi networks. Specifically, the issue we are primarily concerned with is the protection of location privacy, namely safeguards that enable users of WiFi networks to avoid revealing their current location as they move among different wireless hotspots.

When the portable computer belonging to a user connects to a WiFi network, the network operator can tell which access point the user is associated with. With adequate information about the location of the access points, the user can be located to within a few meters. For instance it may possible to pinpoint the location of a user to a particular floor of a hotel, or a particular section of an airport terminal. It may also be possible to know that an individual is currently enroute on a particular flight from London to New York.

WiFi networks carry the potential for revealing much more precise location, compared to other widely-deployed technologies such as cell phones or pagers [5]. WiFi networks operate with much smaller "cell" sizes because they are required to operate at lower power levels and in environments that have poor signal propagation and interference properties. Small cell sizes help maintain signal quality and higher communication bandwidth.

With the increasing popularity of WiFi networks, comes an increasing user population that is likely to have little or no technical background. These users are unlikely to understand how wireless communication works. It is also improbable that they will understand how their privacy can be compromised. They are even less likely to understand and follow security protocols to help improve their privacy. Price and ease of use are generally the overriding factors that determine success of a mass-market offering targeted to such users. It is imperative that privacy protection be made possible without an increase in price, or additional explicit actions by the user.

Their lack of knowledge not withstanding, users still have several tacit expectations of the technology. As technologists it is our responsibility to deliver [6] on these expectations even though the users may not be able to express their expectations in

terms technologists use [7]. Delivering on these expectations is a fundamental requirement for achieving the vision of truly ubiquitous computing [8].

Safeguarding privacy is like transporting water using a bucket that is riddled with holes. Newer technologies, and their usage modes, tend to create more holes in the bucket. While one may not be able to plug all the holes in the bucket, it is still worthwhile to examine each hole individually and devise means to plug that particular hole. Existence of a hole elsewhere in the bucket is not a justification for creating a new hole, or to avoid plugging one that can be plugged.

We propose a simple and practical solution to plug the hole that leaks fine-grain location information as mobile users take advantage of pervasive wireless internet access services. A *practical* solution must be simple and easy for non-technical users to adopt and believe. Further it should be cost-effective and attractive enough so that service providers find it better than alternatives that lack privacy properties.

2 Common WiFi Privacy concerns

A simple approach to providing WiFi access involves a subscriber establishing an account with a service provider. To establish the account, the subscriber will typically provide her name, address, and a credit card number. In addition the service provider may collect other personal information such as phone numbers, and an email address. The subscriber will establish a login id and password as part of the service set up. Subsequently, she will sign on using the login id and password to obtain WiFi access. The service provider will use the login id to measure her usage and bill her for the service. The service provider may also have roaming agreements with several other providers to enable subscribers to obtain WiFi service at various locations.

The service provider will prepare a service agreement which states what information they gather about subscribers, how long they retain the information, how they use the information, and who they share that information with. The service agreement will typically run to several pages of legal language that most subscribers will not fully comprehend, or even bother to read. Nevertheless, the service provider insists that the subscriber sign a statement accepting their terms. Most subscribers will assume that the agreement is benign, and sign it without fully understanding the implications. The subscriber's signature gives the service provider a license to use the information gathered about the subscriber.

Most subscribers will generally be unaware of the amounts of information that the service provider can potentially obtain and link with them. The individual pieces of information may just be minor privacy leaks. However, when someone can build a bigger picture by correlating different bits of information and associating all of these bits of information with a particular subscriber, the privacy invasion becomes much more worrisome. If the details of the correlated information gathered about a subscriber, is subsequently revealed to her, the reaction will generally be one of shock and disbelief. For instance the service provider may be able to tell which cities a subscriber visited. Depending on the extent of WiFi coverage, the service provider may have knowledge of which restaurants or other public places the subscriber vis-

ited and at what times. The service provider may also know which web sites the subscriber normally visits and what kinds of information she reads.

Kotz and Essien [9] have shown that it is possible to collect several pieces of information about WiFi users and also correlate pieces of information that are gathered at different geographic locations at different points in time. They collected data at a university WiFi network, using simple low cost instrumentation. The analysis and correlation was also done using relatively inexpensive hardware. Even so, they were able to build a detailed and rich picture of the users of the WiFi network. A service provider with a profit motive and access to additional resources, could be easily tempted to collect, correlate, and hoard much more information.

Once such information is available, it may be used in ways that may surprise most subscribers. An employee of the service provider might notice that the mobile computers belonging to top executives of company A are frequently seen at the same hotspots as the mobile computers of the top executives of company B. This might lead the employee to speculate on an impending deal between the companies even if all the communication between the two companies was both oral and private. A business may want to buy the email addresses of people who travel on a particular route and send them targeted email solicitations.

Note that the security mechanisms that are being proposed to replace WEP will do nothing to prevent the service provider from gathering and using information. While 802.11 security schemes may prevent malicious bystanders from snooping the subscriber's internet traffic or modifying the traffic in nasty ways, it is unlikely that the proposed security schemes will impose any form of deterrent on the service provider from obtaining and logging information about a subscriber.

Many subscribers will typically establish an IPSec/VPN connection to the intranet at their place of employment because of corporate requirements. The VPN tunnels hide Intranet traffic from the service provider and everyone else. However any traffic to Internet sites are typically sent directly, and can be observed by the service provider unless protected by SSL. And more importantly, establishment of a VPN does not prevent the leak of location information to the service provider.

3 Business dynamics affecting privacy

Safeguarding personal privacy is a fundamentally difficult problem because businesses inherently seek more information about people they serve. In general, the more information a business has about its customers, the better its chances of catering to the needs of its customers, and better its chances of improving profits. All things being equal, a business that has more information is likely to outperform its competition. Any privacy mechanism designed to safeguard user privacy has to fight this fundamental proclivity of businesses seeking more information about their customers.

Businesses sometimes cannot function without obtaining certain pieces of personal information about its customers. For instance, laws require some businesses to obtain private information about their customers. US financial institutions are required to obtain social security numbers in order to report income to the government.

Ignorance of the ways in which private information is collected and used enables businesses to develop technologies and business models that continue to punch holes in the privacy bucket. Businesses often develop innovative and useful services that leverage such information. Once such a service has been deployed, it may be hard to justify technologies that plug the privacy leak which enabled the service. It may also be difficult to lobby for laws that plug the leak because privacy advocates would be pitted against customers and businesses who benefit from the service.

Ignorance and apathy among users, helps businesses avoid compensating the users for the usage of information. As businesses exploit some private information successfully, they are encouraged to collect even more. Effectively, a vicious cycle gets established, resulting in a continuous and progressive erosion of privacy.

Sometimes users are offered a benefit for giving up some private information, sometimes the information is stolen from them without their knowledge. At other times giving up information is made a precondition to obtaining a service. For instance, many US-based mobile phone companies collect customer social security numbers to run credit checks.

3.1 Examples

A real-life example of businesses exploiting and benefiting from the lack of customer awareness of privacy issues is that of telephone companies and caller-id. The underlying caller-id technology is relatively simple. When a call is initiated, the technology merely makes the phone number of the initiator available to the person being called.

From a privacy perspective, the initiator's phone number ought to be considered private information that belongs to the initiator. For example, one may want to call several car dealerships to check if a particular model is available for a test drive, without revealing the phone number to the salespeople. When caller-id services were first offered, telephone subscribers were largely unaware that this information could be obtained and used. The telephone companies took advantage of their access to this information to design a service that they made available to their customers, in the form of a caller-id box that displayed the number of the person who was calling. Currently, caller-id is a popular service that several customers pay for. The telephone companies effectively created a revenue source by delivering private information that belonged to call initiators to some call recipients.

As some customers became aware of the invasion of their privacy due to caller-id, the telephone companies offered another service, namely caller-id blocking which they again sold at a price. The telephone companies successfully made their customers pay just to ensure that the telephone company honored the customer's right to privacy.

As stated earlier, the underlying basis for all of these offerings is the lack of awareness amongst customers that enabled the telephone companies to offer caller-id services without opt-in authorization. If caller-id had been based on an opt-in system where each customer had to explicitly authorize the telephone companies to reveal

their number, the caller-id business model would have never taken off. Only a small fraction of the subscribers would have responded to the opt-in request. This would have rendered the caller-id boxes mostly worthless, because only a small fraction of incoming calls would result in any information showing up on the caller-id box.

Another example of a privacy hole that has been created recently is the RFID based toll collection system on many highways across the USA. When drivers sign up for this toll collection service, they are sent a tag that has a unique serial number. The driver installs this tag on his/her vehicle. When the vehicle enters a section of the highway through a gate, the gate records the serial number. Subsequently when the vehicle exits the highway through another gate, the second gate records the serial number again. By matching serial numbers at entry and exit, the highway authority can determine usage and appropriately bill the driver for the relevant tolls.

For billing purposes the highway system needs to maintain an association between the owner (or the vehicle) and the serial number on the tag. Therefore the organization that operates the toll collection system has the ability to track the location of vehicles, and driving patterns. A timestamp on toll gate records can generate additional information that can be exploited.

3.2 Turning the tide

Given the motivating factors described above, safeguarding privacy seems to be a losing battle, perhaps even a lost cause. However there a few factors working in favor of privacy. The following are some factors that prevent businesses from gathering and using more information than they rightfully require:

1. Laws that place limits on the businesses.
2. The cost of acquiring, retaining and processing huge volumes of information.
3. The tendency of businesses to protect information they hold.
4. The bad publicity that might arise if customers were to learn about the information that is being gathered about them, and how the information was being used.
5. Competitive pressures.

Laws: Fundamentally businesses exist to generate revenues and profits. However, they need to obey the laws that govern their behavior. Nations often pass laws that intend to safeguard user privacy. However a majority of the people are unaware that their privacy is being violated. As a result, law makers seldom hear requests or demands for stringent privacy protection laws. As technologists we help people become more aware of privacy issues [10]. An aware public is likely to pressure their law makers to make laws that protect privacy. An aware public will also pressure lawmakers into *avoiding* laws that mandate the collection of excessive amounts of data.

One recent example is Health Insurance Portability and Accounting Act (HIPAA) [11], where the US federal governing body has specified the privacy requirements for medical records in great detail. The European Union has also passed several laws aimed at protecting privacy.

Data Acquisition and Management Costs: In several cases, the high cost of acquiring and managing the data works in favor of privacy. If businesses cannot per-

ceive a near term return on their investment in data gathering and management costs, it is unlikely that they will bother.

For instance, due to the recent regulations, cell phone providers are required to deploy technologies capable of precisely locating subscribers who call to report an emergency situation. However, there is usually a significant cost involved in obtaining precise location. This high cost generally prevents cell phone providers from tracking all of their subscribers at the same level of precision at all times.

Nevertheless, better technologies are rapidly reducing the cost of collecting, managing and correlating information. As costs reduce, the return-on-investment equation becomes easier to satisfy.

Information hoarding: One business may acquire some information about a particular user and another business may acquire some other information about the same user. If the two businesses were able to share and cross-correlate their databases, they may be able to build a user profile that is much more complete. However, businesses tend to be protective of the data they control and tend not to share. Nevertheless, mergers and acquisitions amongst businesses can eliminate such barriers (e.g. In 1999 online advertising company DoubleClick merged with an offline consumer database Abacus Direct [12]. The merged organization intent to correlate their databases was the subject of several complaints and lawsuits.)

Brand Image: Businesses place a high value on their image in the public view and are wary of publicity that can impact this image negatively. A business that receives public attention as a result of their privacy violations (or even potential privacy violations) often suffers a significant blow to their brand image. There are several well-known examples such as the recent release of many credit card numbers, unique serial numbers on CPU chips [13], etc. As a result, publicity concerning the misuse or leakage of private information, is a powerful deterrent aiding privacy protection.

Competition: Another powerful factor motivating businesses to honor privacy is marketplace competition. If one business develops a technology and business model that can offer better privacy protection to its customers, its competitors may be pressured into adopting similar models. If a business can advertise its privacy advantages in the popular media, its competition will be under greater pressure. Effectively competition can build a *virtuous* cycle that encourages businesses to outdo each other on the privacy front.

Note that for privacy to be a selling point, the technology must be simple and obvious enough that a short 30 second TV commercial or a half page of printed advertising can explain the advantages to the customer. The privacy advantages of the solution should be self-evident to most non-technical customers. Privacy enhancing mechanisms must be easily adopted by non-technical users. Solutions that meet these requirements are candidates capable of creating virtuous cycles.

Solutions capable of creating virtuous cycles may already exist. Lack of awareness may be the only issue preventing the cycle from taking hold. A virtuous cycle leading to the eventual demise of caller-id, can be initiated by one phone company offering caller-id blocking as the default and free option, actively advertising the privacy benefits of their service, and successfully stealing customers from their competitors.

Our objectives, then, as technologists, are two-fold. The first is to help improve awareness of privacy issues, and second is to develop privacy enhancing solutions that are simple to understand and easy to deploy.

4 Our solution

Service providers have several objectives. At a minimum, the service provider must be able to show a profit. They must be able to sign up a large number of customers to cover their large infrastructure costs.

Other desirable objectives include the ability to offer their customers a choice of service plans and a choice of payment options. It is also important for service providers to prevent the theft of service by non-paying "customers". (Generally, most service providers settle for limiting the amount of theft rather than outright prevention, since cost of outright prevention may outweigh the cost of tolerating limited theft). Service providers also want to design some affinity into their services so that customers incur a cost to switch to a competitor. Affinity and superior service can help retain customers and improve long term profitability.

Our solution addresses both user privacy concerns and provider requirements. It works by examining each unique identifier that can be used to associate location information with a particular subscriber, and making each such identifier useless. In WiFi networks, there are two identifiers that can be used to compromise location privacy. The first is the user id that is typically assigned as part of the sign on process. Once this user id is rendered useless, the next identifier of concern is the globally unique network interface identifier (MAC address). We first explain our solution from the two perspectives of the subscriber and the service provider and then provide a detailed analysis of how the solution succeeds in achieving its goals.

4.1 The solution from a subscriber's perspective

Our solution to the WiFi access problem is based on a USB dongle which subscribers can purchase from stores or vending machines by paying cash if they choose to. The dongle enables the subscriber to access WiFi services provided by a particular provider. When the subscriber plugs the dongle into her mobile device at a wireless hotspot, the dongle authenticates itself automatically and connects to the network. The explicit sign-on step required by traditional schemes is eliminated. Once connected, the subscriber has access to the Internet.

The price of the dongle consists of two components, a deposit amount which is refunded when the subscriber returns the dongle, and a pre-paid service fee for a certain amount of service. The deposit amount may be lower for long-term service plans. As the subscriber spends time using the service, the pre-paid amount left on the dongle reduces at a rate depending on the service plan she picked when she purchased the dongle. The dongle has a small digital display showing both the deposit amount and the amount of money (or the amount of service) left on it.

When the pre-paid service amount on the dongle gets close to zero, the subscriber takes the dongle to a vending machine, plugs it in and makes a payment to add value to the dongle, and optionally change her service plan. Cash is one of the acceptable forms of payment. The new balance shows up on the dongle.

If the subscriber chooses, she may return the dongle to a store or vending machine and get a refund of the deposit amount plus the pre-paid balance left on the dongle. (Service providers may prefer to refund only a fraction of the service charge to encourage affinity) She may also purchase a brand new dongle from the same vending machine. A subscriber may own multiple dongles and switch between dongles several times a day. Subscribers may also swap or trade dongles with other subscribers.

If the service provider requires the subscriber to install any custom software on her machine, the service provider makes this software available in source code form. The service provider may also make this code available in a compiled and packaged form to aid subscribers. Other businesses or universities may offer this compiling and packaging service. Alternatively, the subscriber may request a trusted party (such as a systems administrator at her place of employment) to install the software for her.

If a subscriber loses her dongle, the subscriber has lost the cash equivalent of the deposit amount plus the unused balance on the dongle. This is the risk the subscriber has to undertake as the price for increased privacy.

All dongles appear identical to the eye of the subscriber with the exception of the balances shown on them.

4.2 The solution from the service provider's perspective

Though outwardly identical, each dongle has a unique MAC address that is burned in and cannot be modified by software. Dongles do not really maintain any running balances, they merely display balance information that they obtain from the access point. Dongles may maintain statistics to help the infrastructure compute running balances.

In order to meter service, the service provider maintains a database indexed by MAC addresses indicating the balance left on the corresponding dongle. Theft due to MAC address spoofing (on hardware not controlled by the service provider) is limited by security mechanisms (discussed below) used by standard-issue dongles as part of the automatic sign-on. Since dongles need to be returned to vending machines or stores for refill, these refill opportunities permit the service provider to upgrade the security mechanism on the dongles. Other than a security update, the refill operation is merely a database update changing the remaining balance associated with a MAC address.

Dongles carry a tamper-evident seal with a stern warning threatening prosecution if the dongle is tampered with. Since the subscribers do not sign any explicit license agreements with the service provider, the seal gives the service provider the authority to legally prosecute anyone who tries to break into the dongle and compromise the security mechanism. The intent of the seal is not to prevent reverse engineering but to deter it. Prevention of reverse engineering is also possible by using ideas described by Dyer et al [14].

Obviously there is a trade-off between the cost of security versus the cost of losses on account of stolen service. Careful monitoring of suspicious usage patterns, periodic security updates, and legal deterrents are used to limit large-scale theft of service by a thief selling spurious dongles. Hackers who succeed in small amounts of service theft are written off, or maybe targeted as potential hiring candidates!

4.3 Possible Security Mechanisms

The well publicized failure of WEP was caused by an attempt to solve all of the problems of confidentiality, integrity as well as access control using a very simple shared secret [1]. Not only was the same secret shared across all communicating entities, the basic security protocol also had flaws which resulted in the shared secret being revealed without much effort on the part of an attacker. Follow-on efforts to address WiFi security are underway and when an acceptable mechanism is defined, service providers are likely to adopt it.

In the interim, the security mechanism that a service provider is concerned about has only one primary goal, namely limitation of theft. Data confidentiality and integrity are not the primary concern of the service provider. Subscribers can achieve these goals using IPSec or SSL. Service providers would like outright theft prevention, but are usually willing to live with less bullet-proof solutions if the cost is lower.

The theft limitation problem is simpler to solve because it can be solved at a higher level of the protocol stack. For instance, we can permit the client device to associate with the access point, establish an IP address for itself via DHCP, and then authenticate over a TCP connection. Until the client device is authenticated, its packets can be prevented from going out to the Internet by outbound packet filtering. While low level (WEP style) data confidentiality and integrity may be useful, they do not play a pivotal role in sign-on authentication accomplished at a higher level of the protocol stack.

One simple mechanism is to just use the balance associated with a MAC address. When the client device attempts to obtain WiFi service, the MAC address can be looked up to verify whether its balance is non zero. Unsold dongles report a zero balance. The balance lookup can happen either at the WiFi association request level or higher up the stack. Admittedly this is an extremely simple scheme that is susceptible to theft. A thief can passively observe valid MAC addresses, and spoof MAC addresses to not only gain access to service but also deplete the account balance of the victim. Though, similar schemes of recording numbers at the point of sale are used by many scratch-off pre-paid phone cards, this may be inadequate since stealing MAC addresses is easy, while stealing phone card pins is difficult.

Going one step further, the service provider may place one secret key on *all* the dongles. As part of the sign on procedure, the access point may challenge the dongle to prove that it holds *the* secret key. The challenge may be in the form of a random number generated by the access point that is sent to the dongle encrypted using the secret key. The dongle hardware and software on the client decrypt the number using the secret key, transform the number in a manner that is agreed upon, re-encrypt the number and send it back. The client side operations must be designed in a manner

that avoids compromising the secrecy of the key. Since the number of sign on requests is likely to be few and far apart, it is unlikely that attackers can compromise this scheme without breaking into the dongle. The secret may be changed periodically, so long as it is done as a rolling upgrade. For service providers this scheme may offer better security at an added cost of the dongle. The dongle may need non volatile storage and some crypto processing capabilities. If the secret on the dongle is leaked, a thief may manufacture and sell spurious dongles resulting in large scale theft.

Instead of having a single shared key amongst all dongles, the service provider may choose to have a per-dongle secret key, and a database of secret keys indexed by MAC address. During the sign on process, the access point locates the key and challenges the dongle to prove that it has the key. An attacker would need to obtain a MAC address and the corresponding key to attack this scheme. The added security of this scheme comes with the additional cost of managing a large number of keys.

As an extreme case, each dongle may be assigned a certificate with a corresponding private key that is stored in the dongle. The dongle and the access point may set up the equivalent of an SSL connection with both server and client side certificates as part of the sign on procedure. Gupta and Gupta [15] have demonstrated the viability of SSL on small devices. The dongle can also be much more powerful than the small PIC controllers and limited memory in smart cards.

In all of the methods, there is a trade-off between cost of deploying the solution and the protection that the service provider gets. The decision is entirely up to the service provider, and invisible to the subscriber. The most important aspect in the eyes of the subscriber is the ease of use that comes with the elimination of the sign-on step.

4.4 An analysis of the solution

From a layman's perspective, the solution achieves privacy by breaking the connection between the user's personal information (such as name, address or email address) from the information used to sign on and use the WiFi services. The user never overtly reveals personal information to the service provider.

However, from a technical perspective the solution has many more aspects.

Sign on user-id or MAC address: The dongles have a unique MAC address and may also carry other unique identifiers. While such identifiers enable the service provider to track the precise location of the dongles, the rest of the solution attempts to break the connection between the dongle and the particular user, making such unique identifiers useless from the perspective of tracking a particular user's location. All of the different aspects of the solution are schemes that aid in breaking this connection.

Client-side software: All efforts to unlink the identity of the subscriber from the identity of dongle can be easily compromised if the service provider requires the installation of a piece of opaque software on the subscriber's device. Software that executes on the subscriber's device may be privy to information that can readily identify the user. Malicious client side software may enable location tracking even with

all of these mechanisms in place. The only protection against this is to make the client side software transparent and open to public scrutiny. While most non-technical users are unlikely to examine the source, experts and privacy advocates will probably complain if a problem exists.

Pre-paid versus post-paid: In a pre-paid service model, the metering of the delivered service does not require the identification of the consumer of the service. It is conceivable that one could design a post-paid service model that shields the identity of the user from the service provider using complex anonymous credit schemes [16]. However, we chose a pre-paid model since it is simpler and easier to understand.

A pre-paid service model, which can be purchased using cash, is a simple model that can be easily explained to a non-technical user. Most users readily understand and appreciate the anonymity of cash purchases. Subscribers who desire a greater assurance of their privacy can pay for the dongles using cash.

Pre-paid models can also be attractive to service providers since they get paid in advance for service they will deliver in the future. In addition, there may be cases where some customers who pre-pay may not actually consume all of the service they are entitled to, which again works in favor of the service provider.

One argument against general pre-paid service models is that pre-paid models deter impulse peaks in service consumption which can result in significant revenues to the service providers. While prevention of such peaks may indeed be a feature that is valued by some customers, service providers may not like it. Most users tend to prefer flat rate plans in any case, where impulse peaks are not an issue.

Obvious value of the dongle: While cash offers anonymity to the subscriber, many subscribers may not like carrying large amounts of cash. In some countries where anonymous cash cards such as Visa Cash [17] are available, these may be used instead with the same anonymity properties of real cash.

However, some customers may wish to purchase the dongles using a non-anonymous payment mechanism such as a credit card or a personal check. When a subscriber uses such a payment mechanism, there is the potential of making an association between the subscriber's identity and the MAC address of the dongle, especially if no corporate boundaries exist between the merchant accepting the payment and the service provider.

The displayed balance on the dongle helps customers, even those who use non-anonymous payment schemes, achieve privacy. The balance displayed on the dongle makes barter of dongles possible. One subscriber has the opportunity to exchange her dongle with a friend, perhaps paying her friend in cash for difference in balances left on the respective dongles. Note that for enhanced privacy, we only need to *allow* for the possibility of such exchanges. Even if no barter actually takes place, the mere suggestion of a non-mediated and an unrecorded barter breaks the association between the dongle purchaser and the dongle user. The possibility of barter makes it impossible for the service provider to reliably match dongles with users.

A subscriber may lend her dongle to a friend. The friend gets a chance to try out public WiFi access and may pay the subscriber for usage based on the difference in the balance shown. The possibility of lending also breaks the association between the purchaser and user. For the service provider the lending of dongles is a valuable form of free word-of-mouth advertising.

Since dongles can be purchased using different payment schemes (cash, credit card, check), it is useful to have a visual indication of which method was used to purchase the dongle. Vendors generally prefer to refund the deposit amount using a payment scheme similar to the one used at purchase time. Having an indication of the payment method can help subscribers understand what form of payment they can expect when they eventually claim their refund on the dongle. The indication may or may not limit barter to dongles that were originally purchased using identical payment schemes depending on the preferences of the individual subscribers. Service providers should be willing to refund the deposit amount to a credit card owned by the current holder of the dongle, regardless of who originally purchased it using a credit card.

Display of the value associated with a dongle may be difficult when the dongle is not plugged in or away from a hotspot, since the information is in the infrastructure and not in the dongle itself. It is acceptable for a service provider to limit the activation of the value display to hotspot locations. Such a limitation will only serve to limit the physical locations where barter can take place. However, note that a permanent display on the dongle only requires a small battery in the dongle. The dongle displays the most recent balance information that was sent to it. Since the power load on this battery is expected to be quite low, it is conceivable that the battery can last until the dongle hardware becomes obsolete.

Most barters are likely to take place only amongst subscribers who know each other, so while it is desirable to make hacking the displayed value difficult, it is not a fundamental security exposure. The actual values are maintained in the back-end anyway. While hacking the displayed value can help one subscriber cheat another, it does not result in any theft of service.

In addition to enabling lending and barter, having an obvious cash value on the dongle is important to subscribers, since it helps them monitor usage.

Barriers to entry: One of the general arguments against pre-paid service models is that such models pose a barrier to entry. Some customers may not want to pay a significant amount of money upfront for service that they do not have any experience with. Similar barriers in the form of long-term service contracts exist in the post-paid model as well. Nevertheless, there is a need to reduce the deposit amount to lower the barrier to entry. The deposit amount protects the service provider from bearing the cost of the dongles that are not returned. Given the current costs of network interface hardware in relation to the price of the service, there may not be a way of completely eliminating the deposit in the near term.

Service providers may want to work around this issue using innovative pricing schemes that enable customers to try out the service for a nominal monetary cost, and maybe even a nominal privacy cost during the trial period. In other words, the subscriber needs to reveal her identity during the trial period so that the service provider can be protected from significant monetary losses.

Service providers also dislike imposing the difficulty of periodic refills on their customers. The client side software may offer to take a credit card number and do automatic refills. Customers concerned about their privacy can choose the cash refill option. If provided, the credit card number should be kept on the client machine and not be permanently associated with the dongle since the dongle can be traded.

Location-based services: WiFi service providers are hoping to use location-based services to generate additional revenues. An interesting feature of our scheme is that it solves both problems. It enables service providers to offer several location-based services to their subscribers without compromising subscriber location privacy. For instance, the client side software can pop up a coupon for a local restaurant on the screen of a WiFi subscriber at lunch time. The subscriber may also indicate her dietary preferences to the client side software to filter such offers.

Built in wireless interfaces: Mobile computer manufacturers are beginning to offer computers with integrated WiFi network interfaces. While built in interfaces offer a lot more convenience compared to external attachments, built-in WiFi interfaces often come with pre-assigned MAC addresses. As mentioned earlier, this MAC address has a strong association with the subscriber who owns the computer. Even without an explicit sign on process, the subscriber's identity may be revealed and can be logged as being associated with that particular MAC address. From that point on, the subscriber's location privacy is compromised. Essentially the subscriber needs to be extremely careful to *never* enable the service provider to associate the MAC address with her identity. Subscribers are unlikely to be careful enough. A simple step such as filling out a non-SSL protected web form with an email address or a listed phone number is enough to permanently leak location privacy, and also link past MAC address logs with a user identity.

Even in our scheme the service provider may be able to associate a leaked identity with the subscriber's current MAC address. However the service provider cannot be reliably assured that this association will persist at the next sign-on due to the possibility of dongle barter. As a result, the ROI equation on the cost of acquiring this information is unlikely to be satisfied.

There is discussion underway in the standards bodies towards temporary MAC addresses that get assigned dynamically [18]. Dynamic MAC addresses may help address some privacy leaks but it may take a long time before dynamic MAC addresses become the default. We discuss dynamic MAC addresses in detail in Section 7.

Initial roll out costs: The scheme proposed in this paper requires the service providers to incur the cost of deploying and operating dongle vending machines at hotspot locations. While this cost is likely to be significant, it does not have to be all incurred at once. This scheme can be incrementally rolled out, and can potentially coexist with traditional solutions that do not offer the same level of subscriber privacy. Initially dongles may be sold and refilled on the net. Vending machines and sales at regular stores may come after the provider has developed a customer base.

WiFi service providers also face a fundamental practical difficulty of showing their subscribers where exactly the coverage hotspots are. Readily identifiable vending machines can also be used as indicators of WiFi hotspots. In addition, the presence of other subscribers working on their mobile computers, with their dongles attached will be a visible indicator of the presence of a hotspot. The dongles themselves can act as advertising vehicles by triggering the curiosity of non-subscribers.

The store or the vending machine will need access to the network in order to communicate with the WiFi service provider as part of the activation or deactivation of the dongles. For vending machines at WiFi hotspots, this problem is easily solved by providing the vending machine with a WiFi interface of its own.

USB: We have confined our discussion above to USB dongles, since USB is a popular and widely-supported interface. One can easily consider supporting other interfaces such as PCMCIA, CompactFlash, etc. Self-service vending machines that print out hard copies of digital photos are widely deployed and accept different formats such as Smart media, memory stick, compact flash, etc. It should be easy to design a vending machine that works with dongles conforming to different interfaces.

Other privacy leaks: While the scheme proposed here can help plug one privacy leak, there are several other ways in which a subscriber can leak their location information. If the confidentiality of the WiFi communication traffic can be compromised, the subscriber risks the loss of both their location privacy as well data that may be much more valuable. Therefore we believe that better over-the-air security protocols will certainly be designed in the near term. In addition, even non-technical subscribers are likely to use VPN software to tunnel into their corporate intranets, offering better confidentiality and integrity, at least to their corporate communications. They are also likely to use SSL to protect some of their communication with outside sites.

Even if the confidentiality of some of the communication is preserved, the user may still be leaking coarse location information. For instance, when the user browses a web site on the network, the IP address that was assigned to the user may be traceable to the particular city or a particular service provider.

While several solutions to this problem exist (e.g. Anonymizer.com), the one that may be simplest and easiest for non-technical users is configuring their browser to use a proxy within their intranet. This way all traffic originating at the mobile computer will pass through the VPN tunnel into the intranet first, and then out to the Internet without the subscriber's IP address being visible to all sites visited.

7 Related work

Research in the areas of security and privacy in electronic communications, pre-dates WiFi networks by several decades. Cryptography theory [19, 20] is a well established field that underlies several protocols and schemes devised to support privacy in electronic communications. Several fundamental results in this area have arisen out of the quest for creating an electronic version of cash. Chaum's MIX networks [21], which rely heavily on asymmetric key cryptography form the basis for several protocols such as Onion routing [22] that provide anonymity in electronic communications. Reed et al [24], discuss a scheme for hiding cell phone location using caller anonymity obtained using Onion routing. All of these schemes rely on interposing a collection of proxies between the communicating parties. If at least one of the proxies guarantees secrecy the privacy properties are preserved. In our case, it is difficult to interpose a proxy between a WiFi card and the public WiFi access point.

A recent paper [25] proposed tackling the privacy issue by balancing the outflow of private data with feedback to users about the gathering of data. While feedback is useful to raise awareness, we believe that it is important to plug all leaks that can be plugged easily.

Another recent paper [26] proposed a method for sending anonymous email to a known recipient from a wireless hotspot using the concept of a dynamically generated

MAC address coupled with payment using anonymous e-cash. The problem addressed in this paper is more complex than the one we are trying to address, and accordingly the solution relies on asymmetric key cryptography along with the broadcast of public keys. An implementation of this scheme would require substantial changes to the WiFi infrastructure, which require a strong and practical business case.

The potential for fixed MAC addresses leaking privacy information is a problem that is being discussed in the 802.11 standards bodies [18]. One of the proposals to combat this problem is to make the mobile computers request temporary MAC addresses from the access point, in a manner similar to the one used by mobile computers to request temporary IP addresses from a DHCP server.

In the case of dynamic IP addresses, the mobile computers rely on a unique MAC address to communicate with the DHCP server. Assigning a MAC address dynamically is harder because there is a boot-strapping problem. In other words, there is no underlying unique address to rely on for sending the first request. The proposed solution relies on mobile computers choosing random addresses just to send out the initial request. The access point assigns a MAC address to the client, which is used for future communications. Not all access points are expected to support dynamic MAC addresses. Eventually, if and when dynamic MAC addresses become the default in WiFi networks, the privacy leak due to fixed MAC addresses will get plugged.

For true location privacy in subscription-based public networks we also need to eliminate unique user ids for signing on. Once dynamic MAC addresses are default, our scheme can be simplified to one along the lines of a pre-paid phone card. Subscribers can buy a scratch-off card with a pre-printed user-id and password and use these pieces of information to obtain network access.

In fact, a scratch-off card scheme already exists but without dynamic MAC addresses. Current users of the Cometa™ wireless service available at McDonalds™ restaurants, can purchase a pre-paid with a temporary user id and password. However the pre-paid card does not plug the leak of location privacy as discussed above.

Systems based on scratch-off pre-paid cards are also more susceptible to service theft since the service provider cannot design as strong a security mechanism as they could have with the dongle mechanism proposed above, since the dongle is a piece of hardware the service provider controls and can periodically update as well.

The potential for MAC addresses leaking location privacy has also been recognized by the IPv6 community. One of the proposals for IPv6 address assignment is for each computer to assign itself an IPv6 address formed by concatenating a router advertised prefix with its own MAC address. The privacy problems associated with this approach are more acute than the ones we have discussed. The MAC address is normally seen only in the immediate vicinity of the mobile computer. However, if the MAC address is part of the IPv6 address, it is observable by everyone. From a privacy perspective this is similar to making a person's cell phone number by prefixing their social security number with an area code. Not only does the user reveal her (coarse-grain) location, but also her complete identity, when communicating. A standard's track RFC [27] is under discussion in the IETF to address privacy concerns.

Some of the complexities associated with our solution, namely the need for deploying vending machines where dongles can be purchased using cash can be ameliorated if truly anonymous e-cash were available. Several proposals for anonymous e-

cash have been investigated [28], but to the best of our knowledge none of these have had the widespread success. Therefore our solution relies on the real cash that users are familiar with. Once e-cash becomes available our solution may become even simpler since it is an ideal choice for refilling dongles.

8 Conclusions

Technology is constantly improving our ability to track the location of people and things to much finer granularities. While location tracking of things such as parcels, shipping containers, livestock, wildlife, etc., is useful, privacy concerns must be addressed when similar technologies can be applied to tracking people. When the location of a tracked device reveals the location of a person a privacy hole is created.

Owners of mobile phones or pagers can be continually tracked at a coarse granularity. WiFi network users can be tracked to much finer granularities. We believe that users should not have to give up their location privacy in order to benefit from the convenience of public WiFi networks, nor should they be required to take complicated steps to safeguard their privacy. WiFi service providers also need solutions that enable them to operate profitably while respecting the privacy of their subscribers.

We have presented a simple and practical solution that achieves the above goals. Our method uses an externally attached WiFi interface dongle that can be purchased and bartered in transactions that break the association between the subscriber and the dongle. Our method also enables service providers to offer location-based services even while subscribers retain their right to location privacy. Our solution enables non-technical users safeguard their location privacy by continuing to rely on well understood and familiar technologies such as SSL and VPN. Our solution may also be used along with existing anonymity technologies for greater privacy protection.

The inherent complexity associated with the different amounts of private information that can be acquired, retained and correlated, has resulted in a major effort on the part of technologists [29] to simplify and present privacy related information in a manner that non-technical users can comprehend it. Comprehending privacy and the loss of privacy are fundamental steps before users can take action to protect it. We hope that more businesses can be co-opted into this effort of educating consumers about privacy by enabling more businesses to advertise and sell solutions that respect customer privacy, and effectively making privacy a competitive advantage.

References

1. Borisov, N., Goldberg, I., David Wagner, D.: Intercepting mobile communications: The insecurity of 802.11. In Proceedings of MOBICOM 2001, (2001) 180-189.
2. Convery, S., Miller, D., Sundaralingam, S: Cisco SAFE: Wireless LAN Security in Depth http://www.cisco.com/warp/public/cc/so/cuso/epso/sqfr/safwl_wp.htm
3. Mishra, A., Arbaugh, W.A.: "An Initial Security Analysis of the IEEE 802.1X Standard, Dept of Computer Science, Univ. of Maryland at College Park, CS-TR-4328, (Feb 2002)

4. IEEE 802.11 Wireless LAN Standards. IEEE 802.11 Working Group (http://grouper.ieee.org/groups/802/11/).
5. Wireless Location Privacy, http://www.cdt.org/privacy/issues/location/
6. Cranor, L. F.: The Role of Privacy Enhancing Technologies, In Considering Consumer Privacy: A Resource for Policymakers and Practitioners, Center for Democracy and Technology, March 2003, pp. 80-83.
7. Langheinrich, M., Privacy by design: Principles of Privacy-aware Ubiquitous Systems, In Proceedings of Ubicomp (2001) 273-291
8. Weiser, M.: The Computer for the 21st Century, Scientific American 1991, 265(3), 94-104.
9. Kotz, D., and Essien, K.: "Analysis of a Campus-Wide Wireless Network", Proc. of the 8th Annual Intl. Conf. on Mobile Computing and Networking, ACM Press, (2002), 107-118.
10. Nguyen, David H., and Elizabeth D. Mynatt. Privacy Mirrors: Understanding and Shaping Socio-technical Ubiquitous Computing Systems. Georgia Institute of Technology Technical Report GIT-GVU-02-16. June 2002
11. Health Insurance Reform: Security Standards Final Rule, Federal Register Vol 68, No 34., http://aspe.hhs.gov/admnsimp/FINAL/FR03-8334.pdf
12. Privacy Groups See Danger in Merger, New York Times, June 22, 1999. Section C, Page 6.
13. Intel Pentium III Processor Serial Number, http://www.cdt.org/privacy/issues/pentium3/
14. Dyer, J.G., Lindemann, M, Perez, R., Seiler R, van Doorn, L., Sean W. Smith, Weingart, S.: Building the IBM 4758 Secure Coprocessor, IEEE Computer, Vol 34 No 10 (2001) 57-66.
15. Gupta, V. and S. Gupta, S.:Experiments in Wireless Internet Security," in Proc. Wireless Communications and Networking Conference, (Mar. 2002), 860-864.
16. Low, S. H, Maxemchuk N. F., Paul S., Anonymous Credit Cards. In Proceedings of second ACM Conference on Computer and Communication Security. (1994) 108-117.
17. Visa Cash. http://international.visa.com/products/vcash.
18. Orava P, H. Haveniren, J-P. Honkanen, Edney, J,: Temporary MAC Addresses for Anonymity. IEEE Document doc.:802.11-02/261r2.
19. Schneier, B.: Applied cryptography (2nd ed.): protocols, algorithms, and source code in C, John Wiley & Sons, Inc., New York, NY, (1995)
20. Stallings, W.: Cryptography and network security (2nd ed.): principles and practice, Prentice-Hall, Inc., Upper Saddle River, NJ, (1998)
21. Chaum, D.: Untraceable electronic mail, return addresses, and digital pseudonyms. Communications of the ACM, Vol 24 No. 2 (1981) 84-88
22. Goldschlag, D.M., Reed, M.G, Syverson, P.F.:Onion Routing for Anonymous and Private Internet Connections, Communications of the ACM, vol. 42, num. 2, (1999)
23. Reiter M. K. and Aviel D. Rubin. Crowds: Anonymity for Web Transactions. ACM transactions on Information and Systems Security, (1)1, 66-92, June 1998.
24. Reed M., P. Syverson and D. Goldschlag, Protocols using Anonymous Connections: Mobile Applications, 1997 Security Protocols Workshop.
25. Jiang X., Hong, J. I., Landay, J.A., Approximate Information Flows: Socially-Based Modeling of Privacy in Ubiquitous Computing. In proceedings of Ubicomp (2002) 176-193.
26. Molina-Jiménez, C, Marshall, L,: Anonymity without Mixes. In: Second IEEE Workshop on Internet Applications (WIAPP '01), San Jose, CA (2001), 32-40.
27. Privacy Extensions for Stateless Address Autoconfiguration in IPv6. IETF RFC 3041.
28. Asokan, N., et al., The State of the Art in Electronic Payment Systems. IEEE Computer 30(9), Sept 1997, 28-25.
29. Platform for Privacy Preferences. http://www.w3.org/P3P/