

PSEUDONYM CHANGES IN MOBILE ADHOC NETWORK

M.Ganthimathi M.E(cse)

gandhimanmay02@gmail.com

Abstract

Privacy of personal location information is becoming an increasingly important issue. This paper refines a method, called the mix zone, developed to enhance user privacy in location-based services. We improve the mathematical model, examine and minimise computational complexity and develop a method of providing feedback to users. In many envisioned mobile ad hoc networks, nodes are expected to periodically beacon to advertise their presence. In this way, they can receive messages addressed to them or participate in routing operations. Yet, these beacons leak information about the nodes and thus hamper their privacy. A classic remedy consists in each node making use of (certified) pseudonyms and changing its pseudonym in specific locations called mix zones. Of course, privacy is then higher if the pseudonyms are short-lived (i.e., nodes have a short distance to confusion), but pseudonyms can be costly, as they are usually obtained from an external authority. In this paper, we provide detailed analytical evaluation of the age of pseudonyms based on differential equations. We corroborate this model by a set of simulations. This paper thus provides a detailed quantitative framework for selecting the parameters of a pseudonym-based privacy system in peer-to-peer wireless networks.

Index terms: mobile computing, pseudonyms age, mix zone

I. INTRODUCTION

Traditionally, privacy of personal location information has not been a critical issue but, with the development of location tracking systems capable of following user movement twenty-four hours a day and seven days a week, location privacy becomes important: records of everything from the shelves you visit in the library to the clinics you visit in a hospital can represent a very intrusive catalogue of data. Location privacy is an important new issue and several strategies have been suggested to protect personal location information. The strategy is to restrict access. The Geographic Location/Privacy (Geopriv) Working Group [1] have outlined an architecture to allow users to control delivery and accuracy of location information through rule-based policies. Hengartner and Steenkiste [2] describe a method of using digital certificates combined with rule-based policies to protect location information. An alternative approach is to degrade information in a controlled way before releasing it. Gruteser and Grunwald reduce the resolution of location information available to location-aware applications [3]. In previous work [4] we introduced the mix zone model: the model anonymizes user identity by restricting the positions where users can be located. The model provides: a middleware mechanism to

provide anonymised location information to third-party applications, and a quantitative run-time estimate of the level of anonymity provided by the middleware with a particular set of applications. To gain a proper understanding of the privacy properties of mix zones it is important to know how hard it is to break the anonymity the system provides. The mix zone approach for calculating anonymity does this: the degree of success in playing the role of attacker attempting to recover the long-term user identities hidden by the constantly changing pseudonym is an inverse measure of the anonymity offered by the system. In this paper we refine and extend our work on the mix zone model to: show how to deal with irregularly-shaped zone boundaries and how to improve the accuracy of the observations for zones of a given size and shape examine and reduce the computational complexity of the algorithm used by the attacker to break the anonymity, and develop a method of measuring and providing feedback of the level of anonymity the user experiences. Virtually all deployed wireless networks require mobile nodes to communicate in a single hop with the (wired) infrastructure, typically through a base station or an access point. However, the growing popularity of Bluetooth, WiFi in ad hoc mode, and other similar

techniques are likely to fuel the adoption of peer-to-peer wireless communications. In that case, wireless nodes communicate directly with each other over a single hop or over multiple hops. This capability can be used to support a number of applications, ranging from urban sensing to mobile social networks to vehicular ad hoc networks (VANET). Of course, peer-to-peer wireless communications can coexist with the aforementioned classic wireless networks. In this paper, we focus exclusively on the former. In most peer-to-peer wireless communication systems, each node is expected to periodically beacon to advertise its presence. In this way, it can receive messages addressed to it or participate in routing operations. Yet, these beacons leak information about the node and thus hamper its privacy. In particular, external parties can monitor beacons to learn the locations of mobile nodes. A classic remedy to protect the location privacy of mobile nodes consists in relying on multiple pseudonyms: a node uses a pseudonym for a while, then discards it and makes use of a new one. This requires each node to have a repository of pseudonyms that it refills whenever needed. In many cases, these pseudonyms are used by other entities (e.g., other nodes) as trustworthy identifiers for authentication and thus need to be certified by a trusted certification authority. The pseudonym mechanism must thus be designed with great care, because information about the identity of the node can potentially be leaked at various protocol layers, notably by the IP and MAC address [13]. But even with these precautions, changing pseudonyms from time to time might not be enough, because the adversary can track mobile nodes spatially and temporally. As a consequence, nodes should change their pseudonyms in a coordinated fashion with their neighbors in *mix zones*. In other words, location privacy cannot be achieved by itself and requires a collective effort from neighboring mobile nodes. The age of a pseudonym refers to the time period over which a given pseudonym is used. Of course, privacy is higher if the pseudonyms are short-lived. Yet, pseudonyms are costly, as they are usually obtained from an external authority and because a change of pseudonym is a burden for a node: that change can mean remaining unreachable for a short while (typically during the sojourn in the mix zone), entailing the loss of ongoing transactions, or requiring the update of routing tables. Consequently, in many cases a node might consider that its level of privacy is still high enough and might prefer to *not* change its pseudonym, even if it is located in a mix zone. The coordination of pseudonym changes among nodes with different privacy levels is thus a central problem to achieve location privacy with

multiple pseudonyms. However, such solutions require the help of the infrastructure or that mobile nodes learn prior to entering the network the location of mix zones. where mobile nodes coordinate pseudonym changes to dynamically obtain mix zones. This solution is particularly appealing to mobile ad hoc networks because it does not require the help of the infrastructure. In a distributed setting, it remains unclear how successful nodes will be in coordinating their pseudonym changes and how it will affect the age of their pseudonyms. Most existing evaluations do not model the dynamics of the system and consequently do not provide critical conditions for the success of the multiple pseudonym approach. In this paper, we push this distributed approach further and provide a framework for analytically evaluating the privacy obtained with mix zones. That framework captures the mobility of the nodes and the evolution of their privacy level over time. It provides designers

II. SYSTEM MODEL

In this section, we introduce the assumptions made throughout the paper.

A. Network Model

We study a network where mobile nodes are autonomous entities equipped with WiFi or Bluetooth-enabled devices that communicate with each other upon coming into radio range. In other words, we consider a mobile wireless system such as a vehicular network or a network of directly communicating hand-held devices. Without loss of generality, we assume that each user in the system has a single mobile device and thus corresponds to a single node in the network. As commonly assumed in such networks, we consider an offline central authority (CA) run by an independent trusted third party that manages, among other things, the security and privacy of the network. In line with the multiple pseudonym approach, we assume that prior to joining the network, every mobile node in $i=1, \dots, N$, where N is the total number of mobile nodes in the system, registers with the CA that preloads a finite set of *pseudonyms* (e.g., certified public/private key pairs, MAC addresses). Mobile nodes change pseudonyms in mix zones in order to achieve location privacy. Upon changing pseudonyms, we consider for simplicity that the old pseudonym expires and is removed from the node's memory. Once a mobile node has used all its pseudonyms, it contacts the CA to obtain a new set of pseudonyms. We assume that mobile nodes automatically exchange information (unbeknownst to their users) as soon as they are in communication

range. Note that our evaluation is independent of the communication protocol. Without loss of generality, we assume that mobile nodes advertise their presence by periodically broadcasting proximity beacons containing the node's identifying information (i.e., the sender attaches its pseudonym to its messages). Due to the broadcast nature of wireless communications, beacons enable mobile nodes to discover their neighbors. For example, when a node receives an authenticated beacon, it controls the legitimacy of the sender by checking the certificate of the public key of the sender. After that, the node verifies the signature of the beacon message. We consider a discrete time system with initial time $t = 0$. At each time step t , each mobile node can move independently of others on a plane in the considered area. We consider a random-trip mobility model characterized by the rate of encounter and the average number of nodes met in an encounter 1N . The rate determines the number of encounters with nearby nodes that occur on average. The average 1N establishes the average number of nodes that participate in each encounter. The meeting rate λ and the average 1N depend on nodes' speed and the topology of the underlying road network. In our simulations, satisfying predetermined λ and 1N values.

B. Threat Model

An adversary A aims at tracking the location of some mobile nodes. In practice, the adversary can be a rogue individual, a set of malicious mobile nodes, or might even deploy its own infrastructure (e.g., by placing eavesdropping devices in a given area). We assume that the adversary is *passive* and simply eavesdrops on communications. In the worst case, A obtains complete coverage and tracks mobile nodes throughout the entire area. We characterize the latter type of adversary as *global*. A collects identifying information (e.g., the MAC address or the public keys used to sign messages) from the entire network and obtains *location traces* that allow him to track the location of mobile nodes. The problem we tackle in this paper consists in protecting the *location privacy* of mobile nodes, that is, in preventing other parties from learning a node's past and current location [4]. It must be noted that, at the physical layer, the wireless transceiver has a wireless fingerprint that the adversary could use to identify it. However, this requires a costly installation for the adversary and stringent conditions on the wireless medium; in addition, countermeasures could be developed. Hence, it remains unclear how much identifying information can be extracted in practice from the physical layer and we do not consider this threat. Finally, note that higher layer defenses such as mix zones can be useful whether or not physical layer

attacks are in place. For example, some applications may need to store location data to do congestion analysis.

C. Location Privacy Model

There are several techniques to mitigate the tracking of mobile nodes. In this paper, we consider the use of *multiple pseudonyms*: mobile nodes change over time their pseudonym to reduce their long term linkability.

1) *Mix Zones*: Mobile nodes in proximity of each other coordinate pseudonym changes in regions called mix zones in order to avoid temporal correlation of their locations. Mix zones can also conceal the trajectory of mobile nodes in order to protect against the spatial correlation of location traces, e.g., by using (i) silent mix zones (ii) a mobile proxy (iii) regions where the adversary has no coverage or (iv) encrypted communications. Without loss of generality, we assume silent mix zones: mobile nodes turn off their transceivers and stop sending messages for a certain period of time. If at least two nodes change pseudonyms in a silent mix zone, a mixing of their whereabouts occurs and the mix zone becomes a *confusion point* for the adversary. *Distance to Confusion or the Age of Pseudonyms*: As observed in [19], the degree of location privacy not only depends on the location privacy achieved in mix zones by the nodes traversing it, but also on how long an adversary can successfully track mobile nodes between mix zones. A longer tracking period increases the likelihood that the adversary identifies the mobile nodes. Hence, mobile nodes should evaluate the distance over which they are potentially tracked by an adversary (i.e., the *distance to confusion* [17]) and act upon it by deciding to change pseudonyms accordingly. To capture the notion of distance to confusion, we define the *age of a pseudonym* as the time period over which a given pseudonym is used. In this work, we model the evolution of the age of pseudonyms over time $Z_i(t)$ for each mobile node ui as a linearly increasing function of time with an *aging rate*

$$Z_i(t) = \lambda_i \cdot (t - T_i^l)$$

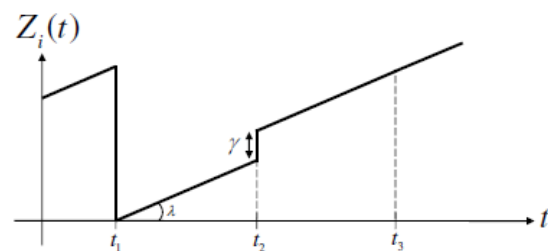


Fig. 1. Example of evolution of the age of pseudonyms. At t_1 , node ui successfully changes pseudonym with another node and the age of its.

The age of pseudonym of node u_i then increases with rate λ . At t_2 , the pseudonym change fails and node u_i pays the cost γ of changing a pseudonym. At t_3 , the node refuses to change its pseudonym.

To gain a proper understanding of the privacy properties of mix zones it is important to find out how hard it is to break the anonymity the system provides. The mix zone approach for calculating anonymity does this: the degree of success in playing the role of attacker attempting to recover the long-term user identities hidden by the constantly changing pseudonyms is an inverse measure of the anonymity offered by the system. In this paper we refine and extend our work on the mix zone model to show how to deal with irregularly-shaped zone boundaries and how to improve the accuracy of the observations for zones of a given size and shape; examine and reduce the computational complexity of the algorithm used by the attacker to break the anonymity, and develop a method of measuring and providing feed back of the level of anonymity the user experiences.

TABLE I
LIST OF SYMBOLS.

| Symbol | Definition |
|-----------------|--|
| t | Time |
| λ | Pseudonym aging rate |
| γ | Cost of changing pseudonym |
| N | Total number of nodes in the system |
| $f(z, t)$ | Probability distribution function of age of pseudonyms at time t |
| $F(z, t)$ | Cumulative distribution function of age of pseudonyms at time t |
| $c(z)$ | Probability of cooperation of mobile nodes with age z |
| $q(t)$ | Probability that at least one node in a meeting at time t cooperates |
| η | Rate of meetings |
| $\bar{c}(t)$ | Probability of cooperation for a randomly selected node at time t |
| h_n | Probability that a meeting involves $n + 1$ nodes |
| $H(\mathbf{z})$ | Z-transform of h_n |
| T_i^t | Time of last successful pseudonym change for node u_i |
| $Z_i(t)$ | Age of pseudonym of node u_i at time t |
| \bar{N} | Average number of nodes in a meeting |
| z | Age of pseudonym |
| θ | Threshold for cooperation |

is thus an incentive to carefully manage pseudonyms. Finally, if a node defects, its pseudonym age is unchanged. Figure 1 illustrates how the age of pseudonyms evolves with time in the case of meetings between several nodes. With this model, nodes control the distance over which they can be tracked. Mobile nodes decide when to change pseudonyms next based on the time of their last successful pseudonym change T_i . We define $ci(z)$ the probability distribution over the age Z_i that gives the

probability of cooperation of each node u_i . For simplicity, we assume that the distribution is the same for all nodes and we write $ci(z) = c(z)$. Hence, when several nodes meet, each node decides whether to change its pseudonym with probability $c(z)$.

III. BLIND CERTIFICATE SCHEME

3.1 Construction of blind certificate scheme

In this section, we present a Pseudonym-Based Signature (PBS) scheme that can be used to create blind certificates for AUs. The PBS scheme is designed based on the BLS scheme proposed by Boneh et al. (2001) and blind signature scheme proposed by Boldyreva (2003). It is desirable that an anonymous communication organiser who publishes the system parameter $params$ is able to grant the admission to AUs. One way to do this is to generate certificates for the pseudonyms that are self-generated by the AUs. In this way, during the anonymous communications, a pseudonym can be validated through verifying its certificate, that is, only the pseudonym with valid certificate will be used as an encryption key during an anonymous communication session. We define the requirements of our blind certificate scheme as follows:

1. Each AU can self-generate his/her pseudonym and corresponding private key
2. The blind certificate generator (i.e. the organiser) is responsible for the certificate generation and both the pseudonym and the corresponding private key are blind to the organiser.
3. The AUs cannot generate new and valid certificates based on the existing valid certificates
4. The AUs can validate a pseudonym and its corresponding certificate by using the publicly known params. To fulfil the above requirements, we propose a four-step scheme: *KeyGen*, *Sign*, *Recover*, and *Verify*. The PBS scheme is presented as follows:
KeyGen: $params = G1, G3, \hat{e}, n, P, Q0, \delta P, H, H2, H3, H4$ is published. The description of params is given in *Verify*: to verify the signature, the AU performs the following test: $e(H(PDA), Q0) = e(\sigma, P)$. In PBS scheme, the AU generates the hash value of PDA. This will allow the AU to mask the point QB . Thus, the organiser will not know the real pseudonym of the AU and the corresponding signature σ . Comparing with the blind signature scheme proposed by Boldyreva (2003), PBS introduces two additional operations: the hash operation H on the pseudonym and the multiplicative masking operation of Next, we will discuss the security and privacy analysis on these two operations.

3.2 Security and anonymity analysis of the PBS

To verify a signature, the PBS scheme tests two pairings operations. To see how it works, we demonstrate the correctness of the testing operations in the *Verify* algorithm as follows:

Boneh et al. (2001). The authors have proved that the BLS scheme is secure against existential forgery under adaptive chosen message attack in the random oracle model assuming *CDH* problem (presented in Appendix) is hard in G_1 . *The PBS scheme is secure and can be reduced to the BLS scheme. Proof:* To prove the our PBS scheme is secure, we first assume that the BLS scheme is secure and then we present how to securely reduce the PBS scheme to the BLS scheme. In BLS scheme, the user's identity IDA is mapped to the point $QA = H_1(IDA) = [k]P$ in G_1 by using a random function $H_1 : \{0, 1\}^n \rightarrow G_1$. H_1 prevents the adversary from determining k by knowing the point QA . The adversary cannot solve k which is equivalent to solving *ECDLP* problem, which is a hard problem. Thus the adversary cannot forge the signature $[k]sP$ for arbitrarily selected user identity IDA by knowing the public key $[s]P$. The PBS scheme introduces the random mapping (one-way) function $H : G_1 \times G_3 \rightarrow G_1$ in the *Verify* algorithm, in which the function H serves the same purpose of H_1 in BLS scheme. Although the adversary knows the public key $[s]P$ and a point QA in G_1 , cannot forge the signature $[s]H(QA, c) = [sk]P$ without knowing the k . Thus, the PBS scheme is another form of BLS scheme with different parameter setting and assumptions. Based on the above analysis, the PBS scheme is a modified version of BLS scheme and the PBS scheme does not change the security strength of BLS scheme.

Lemma : PBS scheme is blind to the blind certificate issuer; the pseudonym holder cannot derive new certificates without getting the Sign procedure from the certificate issuer.

Proof: The PBS scheme introduces the random oracle function $H : G_1 \times G_3 \rightarrow G_1$ in the *KeyGen* algorithm and *Verify* algorithm. The function H serves the similar purpose as of $H_1 : \{0, 1\}^n \rightarrow G_1$ in BLS scheme. Instead of mapping an identity to a point in G_1 , the H first combines a point in G_1 and a value in G_3 and then maps them to a random point in G_1 . This feature will prevent the pseudonym holder from generating new valid certificates based on a known certificate. It might be noted that any bogus AU can impersonate other AUs by submitting others pseudonyms to the organiser in order to derive a valid certificate. However, the bogus AU will get no benefit from the derived certificates since he does not have the private keys of the corresponding pseudonyms. Thus, he cannot decrypt the ciphertext and derive the shared keys as we have discussed in

the PBE and ZKE schemes. We note that in BLS scheme, the $H_1 : \{0, 1\}^n \rightarrow G_1$ is performed at the signer side and it prevents a user from generating an identity from a known point in G_1 . In PBS scheme, we use the random function H to replace H_1 and move the operation of H to the pseudonym holder side. In summary, using the blind certificate scheme, the organizer has the capability in controlling the population of the AUs in the anonymous communication system. In the next section, we will present the pseudonym revocation scheme based on our PBE, ZKE and PBS schemes.

IV. PSEUDONYM REVOCATION

We have presented a PBE scheme. It may be noted that a simpler solution of the PBE scheme for AUs is to randomly select a number $k \in \mathbb{Z}\delta P$ as the private key and uses $[k]P$ as the AU's pseudonym. In this way, the traditional ECC encryption/decryption algorithms and key exchange algorithm can be applied. However, using this approach, it is difficult to achieve revocation in anonymous communication system since the organiser has no control on the self-generated pseudonyms. In previously presented PBE and PBS schemes, the research goal is to utilise pseudonyms to achieve anonymity in order to prevent the adversaries (including the organiser) from linking a pseudonym to an acting subject. In contrast, the proposed revocation scheme is to grant the revocation abilities to the organiser to revoke one or a set of pseudonyms from the anonymous communication system without knowing the revoked pseudonyms (however, the organiser should know the masked pseudonyms). In addition, the revocation can be issued based on the type of anonymous services or the participants' roles. In all our following discussions, we assume the information originated from the organiser is signed by his/her private key. All AUs can verify the signature by using the his/her public key.

4.1 Service revocation

The revocation can be deployed by changing the public known params. For example, an organiser is in charge of the anonymous services within an anonymous communication system. He/she can publish (via periodically broadcasting or accessible publicly known websites) the system parameters $\text{params} = _G1, G3, \hat{e}, n, P, Q0[i], H, H2, H3, H4_$. All AUs trust the organiser except disclosing their pseudonyms to him/her. The organiser maintains a set of system public keys $Q0[i] = [s_i]P$, where $i = 1, \dots, m$ (m is the total number of services). If the organiser wants to revoke the service number 5 (as an illustrative example, the anonymous file *Pseudonym-*

based cryptography for anonymous communications 281 downloading service is defined as the service number 5), he/she just simply excludes the $Q0[5]$ in the params. Thus, all pseudonyms and their corresponding private keys derived from $Q0[5] = [s5]$ will be revoked (see Section 3.2 on how to generate pseudonyms and corresponding private keys). As the results, none of the anonymous service providers can use their pseudonyms to provide anonymous downloading services within the anonymous communication domain controlled by the organiser, since the service requesters will consider the system public key $Q0[5]$ is revoked.

4.2 Pseudonym revocation

Each pseudonym has a unique signature, for example, $e(QA, Q0)cA$ for the pseudonym $PDA = QA, cA$, where $e(QA, Q0)cA = e(P, Q0)kA$. The value $e(QA, Q0)cA$ will not be changed regardless if the AU A changes the masker cA . In order to revoke a pseudonym, the organiser can maintain a revocation list of values $e(Qx, Q0) \cdot cx$, where x represents the revoked pseudonyms. The revocation list is publicly accessible to all AUs.

5. CONCLUSION

In this paper, we propose a PBE scheme, a zero-round key exchange scheme, a PBS scheme, and revocation schemes for anonymous communications. Using PBE scheme, an AU can self-generate his/her pseudonym and corresponding private key based on a set of publicly known system parameters. During the anonymous communication, a pseudonym uniquely identifies an AU and serves as his/her public key. The PBE scheme is an anonymous version of IBE scheme. For anonymous communication, the PBE scheme is more flexible and scalable since no PKG is required and it is more secure due to the self-generated private key. The proposed ZKE, PBS and pseudonym revocation schemes ensure both the accountability and the admissibility within an anonymous communication system, that is, only the pseudonym with a valid certificate is admissible to the anonymous communication system; in addition, a pseudonym can be revoked by the system organiser. Our proposed schemes ensure both security and anonymity for AUs. Here, we present several research directions based on existing anonymous solutions. For anonymous communications, the AU will change his/her pseudonym frequently to prevent the adversaries from identifying his/her involved anonymous sessions. Thus, it is highly desired:

1. a certificate can be used for multiple pseudonyms

2. multiple non-identifiable pseudonyms map to the same private key by using the same set of system parameters

3. the changes of pseudonyms is traceable, that is, only the anonymous communication peers can trace the changes of the peer's pseudonyms (in this way, the anonymous communication peers will not lose the tracks of established anonymous communication sessions).

Acknowledgements

The author would like to thank Professor Partha Dasgupta for his valuable discussion of this work. The author also thanks anonymous reviewers for their valuable comments to improve the quality of this paper.

References

- [1] M. Benaïm and J.-Y. Le Boudec. A class of mean field interaction models for computer and communication systems. *Performance Evaluation*, 65(11-12):823–838, 2008.
- [2] A. R. Beresford. *Location Privacy in Ubiquitous Computing*. PhD thesis, University of Cambridge, 2005.
- [3] A. R. Beresford and F. Stajano. Location privacy in pervasive computing. *Pervasive Computing, IEEE*, 2(1):46–55, 2003.
- [4] A. R. Beresford and F. Stajano. Mix zones: user privacy in location-aware services. In *Pervasive Computing and Communications Workshops*, pages 127–131, 2004.
- [5] V. Brik, S. Banerjee, M. Gruteser, and S. Oh. Wireless device identification with radiometric signatures. In *MobiCom*, 2008.
- [6] L. Buttyan, T. Holczer, and I. Vajda. On the effectiveness of changing pseudonyms to provide location privacy in VANETs. In *ESAS*, 2007.
- [7] T. Camp, J. Boleng, and V. Davies. A survey of mobility models for ad hoc network research. *Wireless Communication & Mobile Computing (WCMC)*, 2(5):483–502, 2002.
- [8] A. Chaintreau, J.-Y. Le Boudec, and N. Ristanovic. The age of gossip: Spatial mean-field regime. In *ACM Sigmetrics*, 2009.
- [9] B. Danev and S. Capkun. Transient-based identification of wireless sensor nodes. In *IPSN*, 2009.
- [10] J. Freudiger, M. H. Manshaei, J.-P. Hubaux, and D. Parkes. On noncooperative location privacy: A game-theoretic analysis. In *CCS*, 2009.
- [11] J. Freudiger, M. Raya, M. Felegyhazi, P. Papadimitratos, and J.-P. Hubaux. Mix zones for location privacy in vehicular networks. In *WiN-ITS*, 2007.
- [12] J. Freudiger, R. Shokri, and J.-P. Hubaux. On the optimal placement of mix zones. In *PETS*, 2009.
- [13] B. Greenstein, D. McCoy, J. Pang, T. Kohno, S. Seshan, and D. Wetherall. Improving wireless privacy with an identifier-free link layer protocol. In *MobiSys*, 2008.
- [14] M. Gruteser and D. Grunwald. Enhancing location privacy in wireless LAN through disposable interface identifiers: a quantitative analysis. *Mob. Netw. Appl.*, 2005.