# Managing Base Station Location Privacy

Maria Gorlatova
Disney Research
Zurich, Switzerland
mag2206@columbia.edu

Roberto Aiello
Disney Research
Glendale, CA, USA
aiello@disneyresearch.com

Stefan Mangold
Disney Research
Zurich, Switzerland
stefan@disneyresearch.com

***Abstract*—Many of today's location services map locations of wireless base stations and use them to localize mobile devices. Severe security and privacy risks exist when unauthorized third-party location services are able to localize mobile devices. In this work we examine a software module that helps network operators to prevent third parties from aggregating wireless base station identifiers by making the identifiers dynamic. This software operates in the infrastructure and does not require any changes of handsets nor any modification of air interface standards. We also examine another software module that provides authorized mobile devices with the ability to locate themselves at different accuracy levels depending on their permission levels. This module operates in the infrastructure with any air interface and does not require standardization. We analyze the effect of the proposed modules on malicious third-party location services, examine their performance, and analyze potential location service countermeasures.**

## I. Introduction

Wireless network infrastructures, such as Wi-Fi and cellular network base stations, can be used by independent third parties for providing location services to mobile devices [1]–[3]. Such *third party location services* (*TLS*s) discover and store locations of wireless base stations, and use the aggregated location information to later localize mobile devices that are in the base stations' vicinity. We refer to third parties to emphasize that such services are often offered to users without permission or involvement of network operators, which are the respective first and second parties. Although this method has the potential to provide useful services, it also presents numerous security and privacy risks [4], as highlighted by recent news [5], [6]. Traditional access-control based security mechanisms cannot prevent a malicious third party from using the location information, because neither TLSs nor mobile devices need network authorization to obtain base station location information. Location information can simply be obtained using the broadcasts of persistent wireless base station identifiers such as Wi-Fi base station medium access control addresses and cellular base station Cell IDs, i.e., cellular base station identifier.

In this paper we describe a system that allows network operators to prevent third parties from obtaining the base station information, and hence prevents these parties from localizing mobile devices. The proposed system can be implemented on Wi-Fi, cellular, and other networks [7], [8]. Keeping base station identifiers always constant over time is crucial to the stability and performance of third-party location services [9]. Hence, in this work we examine an Intelligent Station Identity Manager (ISIM) module that makes base station identifiers dynamic. In [7] we briefly presented ISIM and provided an initial analysis of its technological feasibility. In this paper we examine ISIM from system security point of view. We examine the disruption ISIM causes to TLSs and discuss potential TLS's countermeasures.

In conjunction with ISIM, we propose to use a Multiple Resolution Location Generator (MRL) module that provides authorized mobile devices with base station location information of different resolution levels. MRL allows to control granularity of location information available to mobile devices, thus MRL is well-suited for coalition environments. Furthermore, MRL allows mobile devices to obtain their location estimates locally, without requiring a centralized location service, thus making mobile device location information more secure. In this work we examine two practical MRL variants, demonstrating, through simulations, the localization errors MRL introduces for different localization algorithms.

The ISIM and MRL modules are software concepts that are relatively easy to realize in existing networks. They do not require any modification to existing mobile devices. Further, ISIM and MRL enable a global security management of location information for base stations in critical defense or homeland security scenarios.

This paper is organized as follows. In Section II we review the related work. We describe the developed system in Section III. ISIM-TLS interplay is examined in Section IV. Section V provides the results of analytical and simulation-based evaluations of ISIM and MRL. Section VI summarizes and concludes the paper.

## II. Related Work

Current third-party location services (e.g., [1]–[3]) localize mobile devices based on pre-mapped positions of wireless base stations (BSs), such as cellular towers and Wi-Fi base stations. The TLSs maintain a centralized database of BS identifier-to-location mappings. Such mappings can be reported to a TLS by specialized devices [1], [2], GPS-enabled mobile devices [10], [11] (e.g., mobile phone crowdsourcing), or can be manually entered [1]. To obtain an estimate of its location, a mobile device records the base station identifiers it overhears and reports them to a TLS. TLS looks up these base stations' coordinates in the database, and, based on them, estimates the location of the mobile device. Current research on TLSs mostly focuses on improving various aspects of TLS

978-1-4673-0081-0/11/$26.00 ©2011 IEEE 

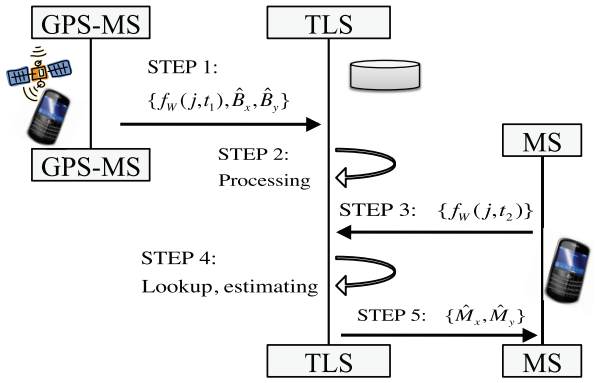



Fig. 1. Timing diagram of TLS operations. GPS-MS devices overhear BS identifiers, estimate BS locations, and report location-identifier mappings to the TLS (St. 1). The TLS aggregates and processes the data (St. 2). To locate itself, an MS records the BS identifiers it overhears, and reports them to the TLS (St. 3). The TLS looks up location-identifier mappings, estimates the location of the reporting MS (St. 4), and returns the location estimate (St. 5).

functionality [10], [11]. Aspects of TLS security are examined in [12], [13].

In this work we focus on preserving base station location privacy and on providing multi-resolution base station location information. Current research in preserving location privacy focuses on location privacy of mobile devices (e.g., Wi-Fi mobile devices [14], or vehicular network devices [15], [16]), rather than base stations' location privacy considered in this work. Similarly, multi-resolution location determination and reporting have been examined for mobile devices [17], [18], rather than base stations we examine in this work.

## III. System Architecture and Preliminaries

In this section we describe TLS operations and the proposed system architecture, and introduce notation, models, and localization algorithms used in this paper. The notation is summarized in Table I.

TABLE I
Nomenclature.

| | |
|---|---|
| $f(j)$ | Permanent (wired) ID of BS $j$ |
| $f_W(j,t)$ | Wireless ID (WID) of BS $j$ at time $t$ |
| $f_{\text{TLS}}(j,t)$ | TLS database entry for BS $j$ at time $t$ |
| $F$ | Total number of BSs |
| $\{B_x(j),B_y(j)\}$ | Location of BS $j$ |
| $\{M_x(n,t),M_y(n,t)\}$ | Location of MS $n$ at time $t$ |
| $\{\widehat{M_x}(n,t),\widehat{M_y}(n,t)\}$ | Location estimate for MS $n$ at time $t$ |
| $\lambda_{\text{ch}}(j,t)$ | Rate of WID changes of BS $j$ at time $t$ $[1/h]$ |
| $r$ | BS coverage radius $[m]$ |
| $k$ | Number of WID reports required by the TLS |
| $\lambda_{\text{lz}}(j,t)$ | Arrival rate of GPS-enabled devices (GPS-MS) to coverage area of BS $j$ at time $t$ $[1/h]$ |
| $p$ | MRL resolution level |
| $C_p$ | Cell of a $p$-defined MRL grid |
| $S_p$ | Side length of a $p$-defined MRL grid cell $[m]$ |
| $CC_p$ | Center of a $p$-defined MRL grid cell |
| $\{B_x^p(j),B_y^p(j)\}$ | Location of BS $j$ provided for MRL resolution level $p$ |
| $d_F$ | Distance between adjacent BSs $[m]$ |
| $l(j)$ | Distance from an MS to a BS $j$ $[m]$ |
| $L(n,t)$ | Number of BSs an MS $n$ overhears at time $t$ |
| $E_x$ | Location estimation error, $x$-axis $[m]$ |

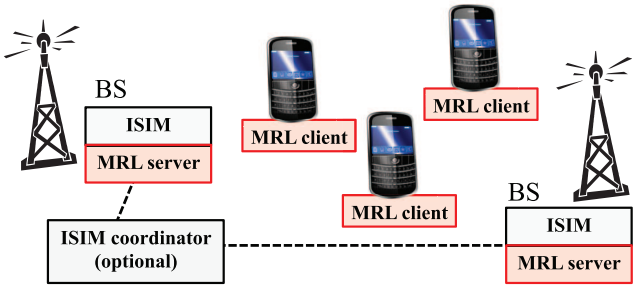



Fig. 2. Schematic diagram of the developed system. Each BS runs an ISIM module and an MRL module. Each MS runs an MRL client module. Some ISIM schemes may additionally use a centralized ISIM coordinator.

### A. TLS Operations

We denote the wireless identifier of a base station (BS) $j$ at time $t$ by $f_W(j,t)$. We denote $j$'s location by $\{B_x(j),B_y(j)\}$, and $j$'s location estimate by $\{\widehat{B_x}(j),\widehat{B_y}(j)\}$, and use $F$ to denote the total number of BSs deployed.

TLS operations are shown schematically in Fig. 1. In step 1 GPS-enabled mobile devices (GPS-MS) arrive to the coverage area of BS $j$. We denote the rate of arrival of these localizing devices by $\lambda_{\text{lz}}(j)$. For analytical tractability we assume arrivals of individual GPS-MS devices and arrivals of GPS-MS devices to different BS coverage areas to be independent. The GPS-MS devices note the BS identifiers they overhear, estimate the locations of the BSs, and report this information to the TLS. That is, a GPS-MS device in the coverage are of a BS $j$ at time $t_1$ will report to the TLS a triple $[f_W(j,t_1),\{\widehat{B_x}(f_W(j,t_1)),\widehat{B_y}(f_W(j,t_1))\}]$. In step 2 the TLS database aggregates and sanitizes the information received from GPS-MS devices. We denote the TLS database entry corresponding to a BS $j$ at time $t$ by $f_{\text{TLS}}(j,t)$. TLS databases use various, typically not disclosed, methods to ensure that the information they store is correct [9]. In this paper we assume that a TLS needs to receive reports from $k$ different GPS-MS devices containing the BS identifier $f_W(j,t)$ before committing $f_W(j,t)$ to the database (that is, before setting $f_{\text{TLS}}(j,t) \leftarrow f_W(j,t)$). Next, in step 3 an MS that wants to localize itself reports the BS identifiers it overhears to the TLS. We denote the number of BSs the MS $n$ observes at time $t$ by $L(n,t)$, and let $\{M_x(n,t),M_y(n,t)\}$ to denote the location of the MS. In step 4 the TLS database looks up the location information it has stored for the BS identifiers reported by the MS, and, using the BS identifiers, generates an MS location estimate $\{\widehat{M_x}(n,t_2),\widehat{M_y}(n,t_2)\}$. Finally, in step 5 this estimate is returned back to the MS.

### B. Proposed Architecture

The scheme we present in this paper consists of an Intelligent Station Identity Manager (ISIM) module and a Multi-Resolution Location Generator (MRL) module. A schematic diagram of the proposed system is provided in Fig. 2: An ISIM module, running on each BS, makes the broadcasted BS identifiers $f_W(j,t)$ dynamic. The identifiers $f_W(j,t)$ differ at

different time instances.[1] ISIM prevents TLSs from providing localization services by ensuring that $f_W(j, t_1)$ reported to the TLS at time $t_1$ does not match $f_W(j, t_2)$ observed by an MS at time $t_2$. In [7] we preliminarily demonstrated the feasibility of using dynamic BS identifiers. With ISIM, a BS $j$ changes its $f_W(j, t)$ at a rate $\lambda_{\text{ch}}(j, t)$. WIDs $f_W(j, t)$ can be chosen randomly independently by each BS, or can use a coordinated assignment scheme. One possible simple coordinated WIDs assignment scheme is briefly examined in Section V-C.

An MRL module running on each BS $j$ broadcasts the location of the BS specified at a set of different resolution levels $\{p\}$. The location information corresponding to each resolution level $p$ is separately encrypted. MSs, running an MRL software client that queries the MRL central server, decrypt the location information corresponding to their permission level. Thus, MRL allows an MS to directly obtain the locations of the BSs it overhears (thus replacing TLS steps 3-5). In this work we focus on MRL location information specifications and the MRL-introduced intentional MS localization errors. Group management, encryption, and key distribution associated with MRL are out of scope for this paper.

We denote the coordinates the BS $j$ provides for a resolution level $p$ by $\{B_x^p(j), B_y^p(j)\}$.

**MRL-CC**: the reported location of the BS $j$ is the center of the $p$-grid cell the BS $j$ is located in, that is,

$$\{B_x^p(j), B_y^p(j)\} \leftarrow \{CC_p : \{B_x(j), B_y(j)\} \in C_p\}.$$

**MRL-RND**: the reported BS $j$ location is chosen randomly from all coordinates in the $p$-grid cell the BS $j$ is located in, that is,

$$\{B_x^p(j), B_y^p(j)\} \leftarrow \text{rand}(\{x, y\} \in C_p : \{B_x(j), B_y(j)\} \in C_p).$$

Demonstrative examples of MRL-CC and MRL-RND location specifications are shown in Fig. 3, where the BS coordinates $\{B_x(j), B_y(j)\}$ are indicated by dots, and the BS coordinates provided for a particular $p$, $\{B_x^p(j), B_y^p(j)\}$, are indicated by squares.

### C. Localization Algorithms

In this work we examine ISIM and MRL using the following set of localization algorithms.

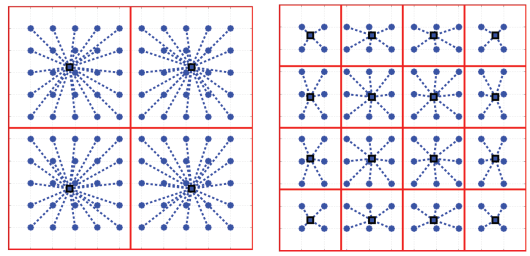
**CO-LOC**: an MS is considered to be colocated with the closest BS [19]. This algorithm is used in RFID-based tracking systems and in Cell ID-based cellular localization.

**Centroid**: the position of an MS is estimated as the geometrical center of the overheard BSs' position estimates:

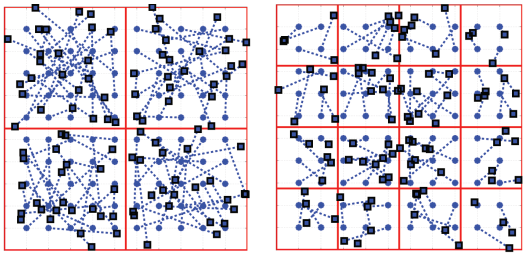
$$\{\widehat{M_x}(n,t), \widehat{M_y}(n,t)\} \leftarrow \frac{1}{L(n)} \left\{ \sum_{j=1}^{L(n)} \widehat{B_x}(j), \sum_{j=1}^{L(n)} \widehat{B_y}(j) \right\}.$$

Centroid-based localization has been examined in Wi-Fi, cellular, and sensor networking contexts, e.g., [20].

[1]While each BS $j$ has a unique ID $f(j)$, this unique ID is only visible towards other infrastructure nodes and does not play a role for the wireless air interface. It is never revealed to MSs.



(a) MRL-CC, left: $S_p = d_F \cdot 5.5$, right: $S_p = d_F. \cdot 2.5$



(b) MRL-RND, left: $S_p = d_F \cdot 5.5$, right: $S_p = d_F. \cdot 2.5$

Fig. 3. Sample BS coordinate specifications with MRL-CC (a), and MRL-RND (b). Dots indicate actual BS locations $\{B_x(j), B_y(j)\}$, while squares correspond to $\{B_x^p(j), B_y^p(j)\}$.

**LAT-LOC**: an MS measures the distance to the BSs it overhears, denoted by $\{l(1), l(2), \ldots, l(L)\}$, and runs a lateration algorithm [19] to obtain its position estimate. In this paper we use a simple algorithm [21] that estimates MS coordinates $\{\widehat{M_x}(n,t), \widehat{M_y}(n,t)]\}$ by solving $H_1 \cdot [\widehat{M_x}(n,t), \widehat{M_y}(n,t)]^T = H_2$, where

$$H_1 = 2 \cdot \begin{bmatrix} \widehat{B_x}(1) - \widehat{B_x}(2), \widehat{B_y}(1) - \widehat{B_y}(2) \\ \widehat{B_x}(1) - \widehat{B_x}(3), \widehat{B_y}(1) - \widehat{B_y}(3) \\ \ldots \\ \widehat{B_x}(1) - \widehat{B_x}(L), \widehat{B_y}(1) - \widehat{B_y}(L) \end{bmatrix}, \; H_2 =$$

$$\begin{bmatrix} \widehat{B_x}(1)^2 + \widehat{B_y}(1)^2 - l(1)^2 - \widehat{B_x}(2)^2 - \widehat{B_y}(2)^2 + l(2)^2 \\ \widehat{B_x}(1)^2 + \widehat{B_y}(1)^2 - l(1)^2 - \widehat{B_x}(3)^2 - \widehat{B_y}(3)^2 + l(3)^2 \\ \ldots \\ \widehat{B_x}(1)^2 + \widehat{B_y}(1)^2 - l(1)^2 - \widehat{B_x}(L)^2 - \widehat{B_y}(L)^2 + l(L)^2 \end{bmatrix}.$$

In analyzing localization error, without loss of generality we examine the error along the $x$ axis, denoting the localization error by $E_x$.

## IV. ISIM Intentions and TLS Countermeasures

Current TLSs employ relatively high-latency BS ID reporting techniques (for example with crowdsourcing BS ID collection with GPS-enabled phones, wardriving), and can be disabled with relatively infrequent WID changes [7]. In this section we examine more capable and more dynamic TLS than the current ones. We consider approaches a TLS may select to maintain its database integrity and to provide localization service in the presence of ISIM. A stationary GPS-MS device positioned next to a BS is able to report all WIDs used by the BS [7]. However, such widespread deployments of GPS-MS devices are costly. Furthermore, in many applications (i.e., a Wi-Fi network deployed in a protected area [8]), this attack can be dealt with using physical site security measures.

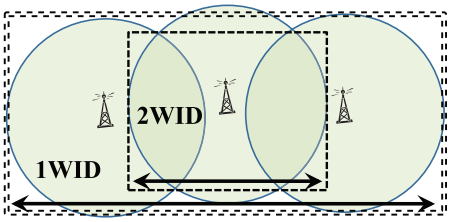



Fig. 4. Schematic view of MS refining its location estimates based on the knowledge of BS locations (but not BS identities) in an area.

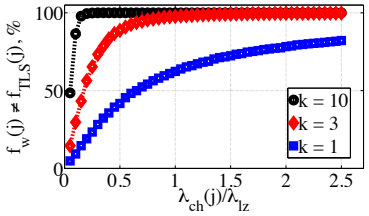



Fig. 5. The percentage of the time BS $j$ TLS database entry is incorrect, for three different values of the number of GPS-MS updates required $k$.

A TLS might potentially deploy powerful GPS-MS devices that can provide WID-to-location mappings for many BSs. Similarly to the case of less powerful GPS-MS devices, such deployments will ensure the immediate correctness of the observed base stations' TLS entries at a particular time, but the TLS identifiers will become outdated as the BSs change their WIDs.

Since each BS ID change produces an additional WID $f_W(j,t)$, the TLS may attempt to remove older $f_{\text{TLS}}(j)$ entries from its database. Varying $\lambda_{\text{ch}}(j,t)$ between different BSs (that is, ensuring that the validity intervals of different $f_{\text{TLS}}(j)$ are different) ensures that a TLS using a timeout concept will not be able to remove older identifiers without disrupting the information related to the current ones. Moreover, if GPS-MS devices provide reasonably precise base station coordinate estimates $\{\widehat{B_x}(j), \widehat{B_y}(j)\}$, from the history of WID updates corresponding to a particular location $(x,y)$ a TLS may deduce the WID change interval $1/\lambda_{\text{ch}}(j,t)$ used by a particular BS. Hence, varying $\lambda_{\text{ch}}(j,t)$ between different times $t$ for a single BSs ensures that the TLS will not be able to learn the validity interval of the WIDs corresponding to the BS, and will not be able to remove outdated database entries $f_{\text{TLS}}(j)$.

Through site surveying or through leakage of information from devices with high permission levels, a TLS may learn BS locations. As permanent BS identifiers $f(j)$ are never wirelessly transmitted, a TLS may thus know the BS locations, but does not know which location corresponds to which BS identifier overheard by a mobile station. An MS could restrict its location to some geographic area (using, for example, geographic cues, navigation history, or MRL-provided coarse-grained location information). Knowing BS locations, an MS can further localize itself within this area. Consider Fig. 4, where shaded parts indicate BS coverage areas. Receiving $L(n) = 1$ BS WIDs, the MS can deduce that its $(x,y)$ coordinates are bound by the border of area $1WID$, receiving $L(n) = 2$ BS WIDs – by the border of area $2WID$. The MS can thus refine its location estimate $\{\widehat{M_x}(n,t), \widehat{M_y}(n,t)\}$ without any knowledge of BS identifiers. In Section V-D we examine such refinements through simulations. The precision of the MS localization results obtained using this method is highly dependent on the BSs deployment. In general, when an MS can restrict itself to an area where only a few BSs are deployed, knowledge of BS locations can substantially help the MS to refine its location estimate. To alleviate this attack we may, for example, employ various cloaking techniques. For example, we might introduce a set of time-varying fake BS beacons to makes the $L(n)$ values of mobile devices incorrect. Furthermore, when providing the MRL $p$-grid defined information, we need to make sure the $p$-grid cells are specified such that that the number of base stations in a grid cell is relatively large.

## V. Analysis and Performance Evaluation

In this section we examine various aspects of the developed system using mathematical analysis and simulations. In the simulations we deploy 100 evenly spaced base stations on a 1100x1100 grid. We denote the distance between BSs by $d_F$; here, $d_F = 100$.

### A. ISIM Impacting TLS Performance

When a BS station WID $f_W(j,t)$ changes, the TLS database entry for station $j$, $f_{\text{TLS}}(j,t)$, becomes incorrect. As result, the TLS will provide incorrect coordinates to mobile devices. In this section we investigate the interplay between $\lambda_{\text{ch}}(j)$, $\lambda_{\text{lz}}(j)$, and the percentage of time the TLS entry corresponding to $j$ is incorrect. After a WID change, the TLS database entry will be incorrect until the $k^{\text{th}}$ arrival of a GPS-MS to the BS coverage area. Since we assume that GPS-MS arrivals are independent, the time intervals from a WID change until the $k$th arrival, $T_g(j)$, is k-Erlang distributed [22], with the probability density function

$$g(T_g(j)) = \frac{(\lambda_{\text{lz}}(j))^k T_g(j)^{k-1} e^{-\lambda_{\text{lz}}(j)\cdot T_g(j)}}{(k-1)!}. \quad (1)$$

If $k$ GPS-MSs do not arrive until the subsequent WID change, the $f_W(j,t)$ is never entered into the TLS database. Hence, assuming that BS WIDs change at set intervals $1/\lambda_{\text{ch}}(j,t)$, the overall expected time that a WID $f_W(j,t)$ is not correct, $\mathbb{E}(T_W)$, can be calculated as

$$\begin{aligned}\mathbb{E}(T_W) &= \int_0^{1/\lambda_{\text{ch}}(j,t)} T_W g(T_W) dT_W \\ &+ \frac{1}{\lambda_{\text{ch}}(j,t)} \int_{1/\lambda_{\text{ch}}(j,t)}^{\infty} g(T_W) dT_W, \end{aligned} \quad (2)$$

where the first term corresponds to the case where the $k^{\text{th}}$ GPS-MS arrives before the subsequent WID change, and the second term accounts for the case where it does not.

The overall percentage of time that a BS WID in the TLS database is incorrect, obtained by numerically evaluating Eq. (2), is shown in Fig. 5. This figure demonstrates the percentage of time the WID is wrong as a function of the ratio of $\lambda_{\text{ch}}(j)/\lambda_{\text{lz}}$ for different values of the number of GPS-MS updates required $k$. It can be seen that even for $k$ as

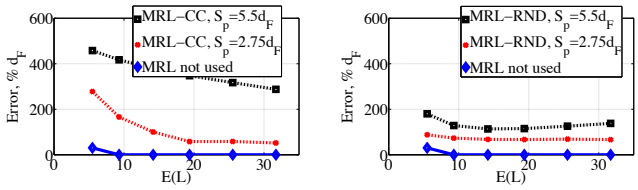



Fig. 6. The errors for the LAT-LOC localization method with MRL-CC (a) and MRL-RND (b).

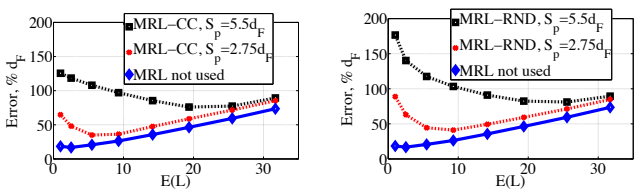



Fig. 7. The errors for the centroid localization method, for (a) MRL-CC, and (b) MRL-RND.

low as 1, when the WID change frequency $\lambda_{\text{ch}}(j)$ matches or exceeds the frequency of the arrival of localizing devices $\lambda_{\text{lz}}$, the TLS database entry for a station is incorrect more than 50% of the time. Hence, relatively infrequent WID changes will be sufficient to greatly disrupt the operations of TLSs that rely on WID updates generated via regular periodic updates or manual entry of WID coordinates. Further, for the case of $k$ relatively large (a more practical case), the TLS database entry for a station is incorrect the majority of the time for the changeover frequency $\lambda_{\text{ch}}(j)$ much smaller than the frequency of localizing devices' arrivals.

### B. MRL-induced Location Errors

In this section we examine the intended MS localization errors introduced by MRL-CC and MRL-RND described in Section III-B. We examine the error under the localization algorithms introduced in Section III-C. Similarly to other sections, our analysis focuses on the error along the $x$-axis, $E_x$.[2]

*1) CO-LOC:* When MRL is not used, $E_x$ with CO-LOC, for an MS $n$ located next to BS $j$, is $E_x(n) = \parallel B_x(j) - M_x(n) \parallel$, and $\max(E_x(n)) = r$ (since the BS coverage radius is $r$). Assuming MSs to be equally likely to be positioned anywhere in the BS coverage area, the average error $\mathop{\mathbb{E}}_{j,n}(E_x(n))$ is $r/2$.

For CO-LOC with MRL-CC at a particular resolution level $p$, $E_x(n) = \parallel B_x^p(j) - M_x(n) \parallel$, and $\max_{j,n}(E_x(n)) = S_p/2 + r$. For a given base station position $B_x(j)$, $\max_n(E_x(n)) = \parallel B_x(j) - B_x^p(j) \parallel + r$. Assuming uniformly distributed BSs $j$, the average distance between the BS $j$ and its declared coordinate $B_x^p(j)$ is $\mathop{\mathbb{E}}_j = \parallel B_x(j) - B_x^p(j) \parallel = S_p/4$. Assuming that the MRL grid cell size $S_p$ is much larger than the BS coverage radius $r$, we may ignore the contribution of the distance from the MS $n$ to the BS $j$ to the localization error. Thus, the average MS localization error of CO-LOC with MRL-CC is $\mathop{\mathbb{E}}_{j,n}(E_x(n)) = \mathop{\mathbb{E}}_j \parallel B_x(j) - B_x^p(j) \parallel = S_p/4$.

For CO-LOC with MRL-RND, $\max_{j,n}(E_x(n)) = S_p + r$. Assuming uniformly distributed BSs, we can demonstrate that the expected distance between the BS $j$ and its stated coordinate $B_x^p(j)$ is $\mathop{\mathbb{E}}_j(E_x(n)) = \parallel B_x(j) - B_x^p(j) \parallel = S_p/3$ [22]. Thus, assuming, similarly to the above, that $S_p$ is much larger than $r$, the expected MS localization error is $\mathop{\mathbb{E}}_{j,n}(E_x(n)) = \mathop{\mathbb{E}}_j \parallel B_x(j) - B_x^p(j) \parallel = S_p/3$.

[2]In these sections we omit index $t$ from the notation.

Hence, for CO-LOC the average MS localization error with MRL-RND is higher than the localization error with MRL-CC, and for both MRL-RND and MRL-CC the introduced error is much higher than the error without the MRL.

*2) Centroid:* The MS localization error for centroid with MRL-CC is shown as a percentage of $d_F$ in Fig. 7(a), for $S_p = 5.5 \cdot d_F$ and $S_p = 2.75 \cdot d_F$. It can be observed that, as expected, use of a coarser grid (larger $S_p$) results in larger errors than use of a finer grid (grid with smaller cells). It can also be observed that as the number of base stations visible to a mobile device increases, the error in the centroid method and the error with the MRL become close to each other, as the errors introduced by the MRL essentially get canceled out. The MS localization errors for MRL-RND (averaged over 20 instances) are shown in Fig. 7(b). It can be seen that for the centroid method the error with MRL-RND is generally higher than the error with MRL-CC.

*3) LAT-LOC:* The MS localization error for LAT-LOC with MRL-CC, as determined by our simulations, is shown in Fig. 6(a), and the error for LAT-LOC with MRL-RND (averaged over 20 instances) is shown in Fig. 6(b). It can be seen that the localization error with both versions of the MRL is much higher than the error of LAT-LOC without the MRL. Further, it can be observed that with LAT-LOC the MS localization error is higher for MRL-CC than for MRL-RND.

### C. ISIM Impacting TLS Localization

Since the correctness of the TLS database entry corresponding to a station $j$ depends on the interplay between $\lambda_{\text{ch}}(j)$ and $\lambda_{\text{lz}}(j)$ as we have demonstrated in Section V-A, at any given time some of the entries in a TLS database will be correct and some will be wrong. Hence we examine how the percentage of WIDs that are incorrect in the TLS database affects the TLS ability to provide MSs with precise location estimates.

We consider a WID assignment scheme where WIDs are permuted (randomly interchanged), and examine the error of centroid localization method. Fig. 8 demonstrates the average MS localization error, displayed as a function of the percentage of WIDs permuted. It can be observed that even if a substantial number of WIDs in the TLS database are correct, ISIM's disruption of the other WIDs results in substantial errors in the MS localization. Additionally, note that these simulation results correspond to WIDs randomly interchanged in a relatively small deployment area. In larger deployments the MS localization errors will be substantially larger than the errors indicated in Fig. 8.

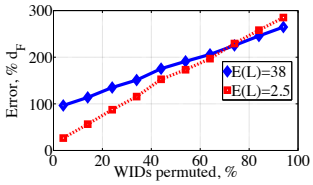



Fig. 8. The localization error introduced by permuting BS WIDs, shown as a function of the percentage of WIDs permuted.

### D. TLS Performance with Anonymous BS Signals

In this section we examine MS positioning precision achieved by a system that has knowledge of base station locations, but not their identifiers. This scenario was introduced in Section IV. We focus on the interplay of such system with MRL-provided grid cell information. We assume that using the MRL-provided BS coordinates, an MS $n$ can narrow its location down to a particular grid cell $C_p$. An MS can easily estimate $L(n)$, the number of different BS IDs it is overhearing. Having the side knowledge of BS locations, the MS estimates how many BS broadcasts $L_{(x,y)}$ are received in each location $(x, y) \in C_p$. Using this information, the MS determines the maximal and the minimal $x$ and $y$ coordinates that receive $L(n)$ broadcasts, and estimated its own location as their geometric mean: $\{\widehat{M_x}(n) \leftarrow (\max x_{C_p} + \min x_{C_p})/2$, $\widehat{M_y}(n) \leftarrow (\max y_{C_p} + \min y_{C_p})/2\}$, where $x_{C_p} \doteq \{x : (x, y) \in C_p, L_{x,y} = L(n)\}$ and $y_{C_p} \doteq \{y : (x, y) \in C_p, L_{x,y} = L(n)\}$. The average error of such estimates for different $p$ values are provided in Table II for three different values of the expected number of BSs seen by an MS, $\mathbb{E}(L)$. It can be observed that when the MS can restrict its location to a relatively small area based on the MRL information (e.g., $S_p = 2.75d_F$), the side knowledge of base station locations allows the MS to obtain its location with substantial precision. These results indicate the need for additional measures to alleviate this attack, as described in Section IV, and demonstrate the need for a careful selection of grid cell sizes.

## VI. Conclusions

Current third-party location services use the locations of wireless networks' base stations to localize mobile devices. However, security and privacy risks exist when such services are provided by unauthorized and uncontrolled third parties. In this paper we analyzed a proposed Intelligent Station Identity Manager (ISIM) module that preserves wireless base station location privacy by making broadcasted base station identifiers dynamic. Further, we evaluated a proposed Multiple Resolution Location Generator (MRL) module that provides base station locations to authorized users at accuracy levels corresponding to user permission levels. We examined analytically and through simulations the effect of ISIM on third-party location services, discussed potential third-party location service countermeasures, and examined localization errors introduced by different versions of MRL.

ISIM and MRL are helpful approaches to enable network operators to control the usage and commercial exploitation of their infrastructure. The required software mainly operates in the infrastructure and does not require any changes of handsets nor any modification of air interface standards. Undesirable and unauthorized usage of location information based on network infrastructure can be inhibited. The proposed MRL enables a differentiation of service classes in location based services.

TABLE II
BS LOCATIONS KNOWLEDGE: AVERAGE MS LOCALIZATION ERRORS.

| | E(L) = 2.6 | E(L) = 9.5 | E(L) = 30 |
|---|---|---|---|
| $S_p = 11d_F$ | $2.75d_F$ | $2.75d_F$ | $2.75d_F$ |
| $S_p = 5.5d_F$ | $1.34d_F$ | $1.26d_F$ | $1.06d_F$ |
| $S_p = 2.75d_F$ | $0.55d_F$ | $0.22d_F$ | $0.25d_F$ |

## References


[1] "Skyhook Wireless," www.skyhookwireless.com.
[2] "Google My Location," googlemobile.blogspot.com/2007/11/new-magical-blue-circle-on-your-map.html.
[3] "Apple Inc. responce to request for information regarding its privacy policy and location-based services," markey.house.gov/docs/applemarkeybarton7-12-10.pdf, accessed May 4, 2011.
[4] I. Bilogrevic, M. Jadliwala, and J. Hubaux, "Security issues in next generation mobile networks: LTE and femtocells," in *2nd International Workshop on Femtocells*, June 2010.
[5] "Inquiries grow over Apple's data collection practices," http://www.nytimes.com/2011/04/22/technology/22data.html, accessed May 2, 2011.
[6] "Got an iPhone or 3G iPad? Apple is recording your moves." http://radar.oreilly.com/2011/04/apple-location-tracking.html, accessed May 2, 2011.
[7] M. Gorlatova, R. Aiello, and S. Mangold, "Managing location privacy in cellular networks with femtocell deployments," in *Proc. IEEE IOFC'11*, May 2011.
[8] K. Collins, S. Mangold, and G.-M. Muntean, "Supporting mobile devices with wireless LAN/MAN in large controlled environments," *IEEE Commun. Mag.*, vol. 48, no. 12, pp. 36–43, Dec. 2010.
[9] K. Jones and L. Liu, "What where Wi: An analysis of millions of Wi-Fi access points," in *Proc. IEEE PORTABLE'07*, May 2007.
[10] J. Yang, A. Varshavsky, H. Liu, Y. Chen, and M. Gruteser, "Accuracy characterization of cell tower localization," in *Proc. ACM Ubicomp'10*, Sept. 2010.
[11] A. Subramanian, P. Deshpande, J. Gaojgao, and S. Das, "Drive-by localization of roadside Wi-Fi networks," in *Proc. IEEE INFOCOM'08*, Apr. 2008.
[12] K. P. N. Puttaswamy and B. Y. Zhao, "Preserving privacy in location-based mobile social applications," in *Proc. HotMobile'10*, Feb. 2010.
[13] N. Tippenhauer, K. Rasmussen, C. Popper, and S. Čapkun, "Attacks on public WLAN-based positioning systems," in *Proc. ACM MobiSys'09*, June 2009.
[14] T. Jiang, H. J. Wang, and Y.-C. Hu, "Preserving location privacy in wireless LANs," in *Proc. ACM MobiSys'07*, June 2007.
[15] M. Gerlach and F. Guttler, "Privacy in VANETs using changing pseudonyms - ideal and real," in *Proc. IEEE VTC'07*, Apr. 2007.
[16] J. Freudiger, M. Manshaei, J. Le Boudec, and J. Hubaux, "On the age of pseudonyms in mobile ad hoc networks," in *Proc. IEEE INFOCOM'10*, Mar. 2010.
[17] W. Wang and C. Cui, "Achieving configural location privacy in location based routing for MANET," in *Proc. IEEE MILCOM'08*, Nov. 2008.
[18] Y. Chen, J. Yang, and F. He, "A trusted infrastructure for facilitating access control of location information," in *Proc. IEEE MILCOM'08*, Nov. 2008.
[19] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 37, no. 6, pp. 1067–1080, 2007.
[20] F. Akgul and K. Pahlavan, "Location awareness for everyday smart computing," in *Proc. IEEE ICT'09*, 2009.
[21] A. Savvides, C.-C. Han, and M. B. Strivastava, "Dynamic fine-grained localization in ad-hoc networks of sensors," in *Proc. ACM MobiCom'01*, Jan 2001.
[22] H. Tijms, *A First Course in Stochastic Models*. Wiley, 2003.