Mobile Location Tracking in Metro Areas: Malnets and Others

Nathaniel Husted
School of Informatics and Computing
Indiana University, Bloomington
nhusted@indiana.edu

Steven Myers
School of Informatics and Computing
Indiana University, Bloomington
samyers@indiana.edu

ABSTRACT

Digital wireless radios broadcast identification numbers that uniquely identify them. As has been previously observed, given the ubiquity with which people carry smartphones with their embedded WiFi radios powered on, comes the ability to track individuals' movements. The ability to use wireless radios for positioning has been previously observed and developed in to useful products. In these systems a user willingly geolocates themselves by providing identifiers to infrastructure hardware. In this paper we consider the converse question: what rates of monitoring by smartphones devices in a given metropolitan area are necessary to achieve different levels of *involuntary* geolocation. While previous work has looked at countermeasure that attempt to maintain privacy, no work has attempted to quantify the problem and risks. Using appropriate simulations we give the first quantitative support for the number and conditions of tracking devices necessary to track the locations of non-participant individuals in urban environments. We provide evidence that a small, but not insignificant, number of mobile devices can be used to track a majority of users during a significant fraction of their travel with current devices. We conclude that in the immediate future, malnets would require relatively high infection rates to pose a significant threat, but that voluntary networks, with perceived benefit can probably achieve the usage rates necessary to track individual movements of non-subscribed users to a high-degree of accuracy. Our results also suggest ubiquitous deployment of 802.11n in smartphones would make geolocation feasible by malnets.

Categories and Subject Descriptors

C.2.0 [**General**]: Security and Protection; K.6.5 [**Management of Computing and Information Systems**]: Security and Protection— *Invasive software (e.g., viruses, worms, Trojan horses)*; K.4.1 [**Computers and Society**]: Public Policy Issues—*Privacy*

CCS'10, October 4–8, 2010, Chicago, Illinois, USA.

General Terms

Security, Measurement

Keywords

geolocation, malnets, privacy, WiFi, surveillance, simulation

1. INTRODUCTION

Location privacy has long been a topic of interest among information security researchers. Recently it has become a topic of specific concern [18, 27, 45]. This is largely due to the proliferation of GPS devices, web location services, and the geolocation of network addresses. The ability to locate a wireless device for beneficial reasons has been looked at by general computer scientists as well [8, 46, 58]. These research projects considered either a static, geolocated set of nodes that broadcast their existence, or an area that had been radio-mapped. We consider the question of wide-scale — potentially involuntary— tracking of individuals in dense metropolitan areas by a small and mobile sub-population working in collusion to monitor others' locations.

We show that wireless location tracking is conceivable in a modern metropolitan area as the prevalence of WiFi enabled smartphones increases. In the malicious case, we assume that there will exist *mobile* malnets[1] [51, 52] whose purpose is to geolocate specific individuals. Work has already been done on the dynamics for self spreading router and cellphone malware as well as its potential [28, 52, 55]. Recently we have seen the first spread of worm based malware on jailbroken iPhones. Thus the potential of such malnets is clear. The question we address is in determining how effective they could be at positioning an individual given different detector levels in the population. More legitimate scenarios are the use of a locating system by location based advertisement networks, theft locators, or Amber Alert services. Users who have opted in to such services, perhaps for some financial or karmic incentive, would monitor the positions of users not in the service. The widespread epidemic of the Gator spyware/advertising network suggests that such unscrupulous behavior could present itself in a corporate advertising setting. In the case of an Amber Alert service, users might sign-up out of altruistic reasons: to help locate missing children by searching for carried electronic toys with embedded WiFi such as PSPs and Zunes. A theft recovery network might search for stolen goods with WiFi devices incorporated, for a potential finders fee. In such a case, users would

[1]Malnets are botnets created from routers, cellphones, and other non-traditional computational WiFi devices.

voluntarily add tracking software to their devices. A myriad of potential groups have incentives to try and geolocate individuals for positive or questionable reasons.

There have been papers discussing the potential of 802.11 and Bluetooth tracking and location privacy [6, 8, 13, 29]. Similarly, there have been papers that have proposed countermeasures to protect against privacy loss due to WiFi identifier broadcasts [21, 31]. Our paper is the first to attempt to specifically quantify the degree to which WiFi Identifier leakage presents a potential threat due to malnets or crowdsourcing applications, versus a mere novelty of being able to position individuals on odd occasions, when users happen to pass by a detector. It differentiates itself by providing these contributions: **1)** We provide strong evidence that a small percentage of oblivious detectors in an average population of a dense metropolitan area can track a majority of users a significant fraction of the time. **2)** We quantify the degree to which different variables, such as population density, detection rate and broadcast radius affect tracking. **3)** We show that small changes in the broadcast radius of wireless signals have a significant effect on the ability to track individuals. Thus, when 802.11n overtakes 802.11g (As predicted in [47]), the tracking capacity will increase at a small cost of accuracy. In particular, we find malnet tracking is likely infeasible with current systems, but when 802.11n overtakes 802.11g, geolocating malnets become feasible in dense metropolitan areas.

Road Map of Paper.

Section 2 characterizes how subnets track individuals with WiFi devices. Section 3 describes our simulation methodology. Section 4 provides our results. Section 5 discusses the need for simulation. Section 6 discusses the construction of tracking networks. Section 7 discusses previous mitigating strategies on location privacy attacks, and their benefit in our scenario. We finish with related work in Section 8 and conclusions in Section 9.

2. OVERVIEW & MOTIVATION

Nearly the entire population of the industrialized world now carries cell-phones. Worldwide, 68.2 out of 100 inhabitants have mobile phone subscriptions, up from 12 in 2000 [3]. Similarly, smartphone usage is rising according to Nielsen [44] which predicts that smartphones will amount to 49% of the cellphone market by Q3 of 2011. With this evolution and adoption of mobile technology come new potential risks: Smartphones are much more susceptible to having homogeneous operating systems—due to their complexity—than traditional cellphones, and thus to wide scale attack. By market share, four operating system families (Android, RIM, Windows and iPhone) had 91% of the smartphone market in 2010 [33].

Since, smartphones, routers, tablets, netbooks and laptops have access to a number of sensors that traditional PCs do not, it permits the use of these sensors for new types of attacks with such malnets, or to provide crowd-sourced sensor applications.

Consider the following scenario. Alice has no malicious software on her phone, but she has her WiFi radio on for ease of connection.[2] She wanders through an urban environment throughout the day, with the phone broadcasting its unique WiFi radio identifier. Concurrently, Eve controls a malnet of smartphones, and she wants to geolocate Alice. Eve, has previously established—through some means—the unique identifier of Alice's radio, and sends a request to all of the bots: on detection of Alice's device to report back the time and location. Based on the information that is returned, Eve triangulates Alice's position over time.

In this scenario, the ability for Eve of determine Alice's position at any given time depends on three factors: i) the coverage of infected devices *with respect to the route Alice is taking over time*; ii) the broadcast diameter of Alice's wireless radios; and iii) the frequency with which Alice's device broadcasts its identity.

Eve's ability to track targets is made easier by common behaviors of smartphones and their users. In many phones (e.g., the iPhone and Android phones), the WiFi is enabled while the phone is powered on and user active (i.e., not in a power-saving mode). Many phones will turn off the WiFi after a given period of inactive time to save battery life. However, despite this, some users find it preferable to have the WiFi constantly on, as is evidenced by applications such as the insomnia application for iPhone[3]. As battery life is extended in phones, we can expect more users to have WiFi that is always on. We discuss phone use more in Sec. 3.

The always-on nature of WiFi means that unless the users consciously decide to turn off the WiFi, which many non-technical users will not, they can be tracked whenever their phone is active. Note that WiFi radios constantly send out probe frames even when not connected to a network. Thus the smartphone need not have an active WiFi connection to be detected.

The devices in the malnet work as triangulators. Smartphones determine their location based on internal GPS systems or other geolocating services they have access to; APs and other WiFi enabled devices' positions are determined through the use of online database reporting schemes, such as Skyhook[4]. Once the location of the detecting device is known, simple trilateration algorithms suffice to position the device to different degrees of confidence. The number of distinct observations, and information correlated with radio signal strength can be used to further improve positional accuracy by the malnet [7, 8].

While we believe the example of the *mobile* malnet is a particularly motivating example, there are many other potential uses of such networks (as previously alluded to), so we will refer to them as tracking networks.

Determining a User's BSSID.

In some scenarios we consider, such as an Amber Alert or theft recovery network, a WiFi device's BSSID (a.k.a.. MAC) address will be given to the tacking network by individuals who have previously collected them. In other scenarios, such as malnets and ad networks, the BSSID will have to be discovered. We believe that determination of an individual's BSSID is not, on average, a difficult problem, and there are many ways in which it can be uncovered. It is, after all, given away to every device one connects to, and broadcast freely. A tracker could find a device's BSSID

[2]This is a common default setting on many phones with WiFi, despite the power consumption implications.

[3]See `http://code.google.com/p/iphone-insomnia/`

[4]Stenvold [43] showed how to geolocate a WiFi router using its BSSID address and the Skyhook interface.

in a number of ways. First, if the tracker can determine the target's work or home, they can ask any detector nodes in either area to collect information about anyone traveling within range and then compare the information to find common and frequently visiting device identifiers. In many cases simple data-filtering could uniquely finger the identifier in question. Second, if a tracker controlled a diverse number of APs, they could attempt to trick a user into connecting to one. Some third-party AP firmwares have the ability to create multiple wireless interfaces. If the tracker modified the access point and created an interface to mimic a common hotspot or an AP the target accesses frequently, the target could easily be tricked into unknowingly and automatically connecting to the duplicated AP interface, revealing its BSSID.

An ad network might be unconcerned about matching a particular BSSID to an individual, but rather be happy to keep profiles on all BSSIDs histories. Alternately, an ad network wanting to match BSSIDs to individuals would not have to resort to access point trickery, but rather could team up with a hot spot provider such as Boingo, T-Mobile, or AT&T, and request the information of customers who connect to the hotspots. In essence, they would willingly purchase BSSIDs that had been associated with individuals from organizations that could easily make the association, due to the nature of their services.

3. METHODOLOGY

Our overall methodology for simulating geolocation using a tracking network follows the following high-level structure:

1. Simulate an appropriate number of traces, $\mathcal{T}$, of individuals' geolocations as they move through a small subsection (indoors and outdoors) of the downtown core of a metropolitan city, during a given time period.
2. Choose a set $S \subset \mathcal{T}$ of locators.
3. Choose a set $T \subseteq \mathcal{T} \setminus S$ of tracked individuals.
4. Simulate the traces over a given time period, transition times of one second.
5. At each time period, for each $x \in T$ record each $y \in S$ that is within transmission diameter d.
6. For each maximal set $\{y\} \subseteq S$ that observes x in a given time period, use a trilateration to minimize the area within which x is expected to be.
7. Determine the frequency with which each tracked individual $x \in T$ is observed and to what area of accuracy (in m^2) his or her position is learnt.

In Step 1 the simulation is done in three dimensions using the UDelModels simulator (See Sec. 5.2 for details). In Step 2, a certain fraction of the population is chosen to act as locators. In the case of a tracking malnet this is the number of infected individuals. In the case of user-installed software, it represents the individuals who installed positioning software. In Step 5, given the location of all individuals being tracked in set T, we determine if any of them are within a close proximity to any of the phones or routers acting as locators in set S. Due to absorption, refraction, diffraction and reflection of radio-waves, which depends both on a number of fixed and intermittent features (e.g., geography and landmarks; or weather, traffic, and radio interference, respectively), we cannot accurately determine exactly where a traced phone's wireless beacons might have emanated. However, this is frequently approximated by using a sphere of a given radius. We considered a number of radii from 15m to 45m to represent lower and upper-bounds on 802.11g wireless transmissions. Additionally, 45m represents a conservative lower-bound for 802.11n. In order to significantly speed up the calculation of proximity detection, and since the sphere is already an approximation, we detect locations within the tightest bounding cubes of the corresponding sphere. See Fig. 1 for a 2-D visualization of this process.[5] If multiple locators are able to locate a device then in Step 6 we use trilateration to narrow down its location. Specifically, we determine the area of overlaps from all of the detecting locators.

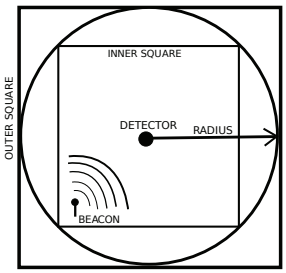

Figure 1: For computational savings we approximate a sphere or radius r (vol. $\frac{4}{3}\pi r^3$) by an outer and inner (vol. $8r^3$ and $\frac{2r^3}{3^{1/2}}$ resp.) bounding cube. This is depicted in 2d to simplify.

Population Density and Simulations.

We derived the number of people for our simulation areas by using LandScan [2] data to derive an accurate estimate for the census population of the area. In particular, we simulate a 9 block region of Chicago with 9056 people, and a 16 block region of Dallas with 2988. This was included in our simulator as the *maximum* population for the given areas at any time of day. We see that the actual populations simulated during the day are often below that (e.g., our 7:30am simulations average 42% of that population in Chicago, with 11:00am having 90% of the population). The lower density earlier in the day can also be explained by the fact most simulated agents do not start in the simulated area but travel into and out of the area throughout the day. We suspect, although could find no data, that the census populated data (of 9056 and 2988) is a conservative estimate on the population of both areas, as they are both relatively absent of housing, and mostly businesses. Thus the census maximums probably correspond to evening maximums, and during the busy periods of the day when the workforce arrives the population presumably grows. Thus we believe our simulations are conservative.

3D vs. 2D.

The experiments are performed in three dimensions, with people going up and down buildings, working in offices on specific floors, etc... One might be tempted to assume that all individuals are on a plane, and perform the simulations

[5]We note that using both ray-tracing or radio-mapping to get a more realistic simulation of wireless propagation is inconceivable due to the computational effort required.

in two dimensions. However, we preformed some early experiments in two-dimensions, for comparison purposes, and results are substantially improved in these models. Therefore, it is more cautious and realistic to maintain three dimensional results. All simulation results stated in this paper are for three-dimensional simulations.

Mobile vs. stationary devices.

We consider only mobile device detectors. This focus is for two reasons: i) we could find no dataset of wireless routers that contained altitude coordinates, and as stated we need three dimensional simulations; and ii) in the case of 2D data, preliminary *2D-simulations* using wardriving placement datasets (e.g., the WiGLE war driving DB [5]) showed limited effect of router surveillance with low detector rates. Manual inspection showed that this was due to the sparse coverage, but inspection of the collected data leads us to believe that it does not include most of the APs and stationary WiFi devices that exists in the simulated area. Given an accurate and full data set of all stationary devices in the area, we believe that they would provide a far stronger role in metropolitan geolocation. Nonetheless, the tracking networks get robust results without the participation of APs and other stationary devices.

Metropolis vs. Other Environments.

The results are specifically for dense metropolises. Our results do not clearly generalize to sparser populations such as suburban environments for several reasons. First, detection requires a certain amount of density, as targets need to be in proximity to detectors. Second, detection of mobile detectors and targets requires social aspects to be modeled. The UDel simulator is designed only to emulate and match statistics for metropolitan areas.

Pedestrians vs. Vehicles.

We only simulated pedestrians. We could not find figures that accurately represented the traffic densities in the areas in question at a fine grained time scale. These are parameters that the UDel simulator needs to properly simulate vehicles. However, the addition of vehicles should only increase the detection rates, as it would only add to the number of potential observers.

The Usage of Phones.

While we can assume that devices in the tracking network activate their WiFi radios whenever it is necessary, the same cannot be said for those being tracked. A question of key importance is how frequently people are actively using their devices, and thus the WiFi radio is sending out probes. Karlson et al. [32] studied the work habits of professionals and their Windows Mobile smartphone habits. Their results show that individuals used their smartphones to stay connected to their work life throughout the day using email and other applications on their windows mobile smart phones. In terms of the rate of use, Karlson et al. showed that some users have near ubiquitous use, while others intermittently, but consistently, use their phones throughout the day. This shows that individuals will be susceptible to being detected frequently by a tracking network over a large range of their travel. A report from ABI Research in 2009 [49] states that 74% of smartphone owners use their phone's WiFi features. Quantcast has stated iPhone is now the predominant device for mobile Web consumption and has a 63.7% share of mobile browser hits on websites [39]. Given the market share of the iPhone, we can conclude that the use of extended features of smartphones is increasing significantly with the improvements to networking and ease-of-use of the phones. Our results should be interpreted as saying that individuals can be tracked when their devices are active. The frequency has some range amongst users. Note for tracking the device does not need to be actively used (which Karlson et al. [32] required), and so their results are a lower-bound.

4. RESULTS

We simulated scenarios in which the tracking network comprises 1%, 5%, 10%, 25%, 50%, 75%, and 100% of the entire population in the designated areas of Chicago and Dallas (as shown in Figs. 9 and 8 in App. **??**). The 802.11g standard has a maximum broadcast radius of $\approx 50m$ under ideal circumstances. We considered simulations with broadcast radii of 15m, 30m, and 45m similar to [28]. We perform simulations throughout the day to show the effects of human circadian activity.

In Fig.2, 10% of the population is actively tracking, and the broadcast radii is15m (outer bounding box). This is a large percentage of users for a malnet, but we believe reasonable for an installed piece of software. We plot which fraction of the tracked population can be observed with a given frequency as time progresses through the day in both Dallas and Chicago. Observe that the ability to track people depends significantly on the time of day, with low-points being when people are most dispersed during rush-hour entry and exit from the workplace (recall, we simulate only pedestrians and not vehicles). Note that with 10% of the population being detectors, they are always detectable. Thus 10% is the minimal detection rate in Fig.2.

As can be seen, the 7-7:30am time period correspond to a low-point in the detection rate, and therefore we will use this time period as a conservative window upon which we can modify alternate variables of the experiments.

Frequency of Detection.

It is evident that to effectively track an individual one does not require 100% detection rates over time, but instead a tracker is presumably content to position someone a significant fraction of the time, and interpolate the position between observations. We calculated the average duration tracked individuals went without detection, assuming different rates of observation. Fig. 3 shows the tracked users' mean undetected walk duration for our experiment in Chicago from 7-7:30am, with a 15, 30, and 45m detection radius. We consider only the individuals who were detected greater than or equal to 10% of the time. The standard deviations for mean undetected times (not shown in the figure for clarity) range in the worst case from 3.23 seconds at 15m to 2.92 at 45m. Thus for a 3σ deviation, an individual will go undetected for at most 18 seconds with 10% of the population as detectors. This number also improves as we increase the size of the tracking network. When considering the whole population, we have a sparser set of detections with higher means and standard deviations. In this sparse case, low percentages might provide scattered detections, but the low entropy of human movements and the walking speed of individuals (recall, only pedestrians are simulated)

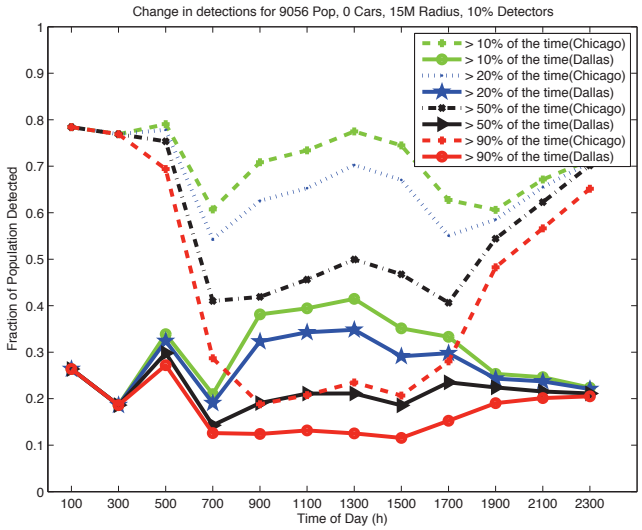

Figure 2: A simulation of 9056 people in Chicago and 2988 in Dallas. Tracking network comprises 10% of the pop. Broadcast radius is 15m.

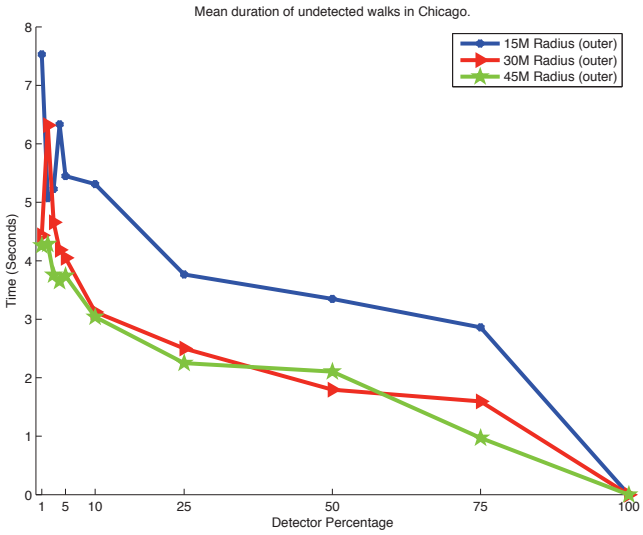

Figure 3: Mean undetected walk durations for Chicago at 7-7:30am for those detected at least 10% of the time.

allow us to interpolate the position of individuals as detections are aggregated over time [48].

Comparison of Dallas to Chicago.

We performed the simulations on regions of two cities, Dallas and Chicago, as they have different properties of density and building use. For example, the height of the different buildings in Chicago is more uniform whereas Dallas has some exceedingly high buildings surrounded by relatively short ones. Both have 22 buildings and an average height again of 27 stories (minus rounding error). Similarly, the distribution of buildings dedicated to housing vs. industry and commerce changes. The result is quite immediate. Although the number of people in Dallas is about a third of Chicago and the area simulated greater, the rates of detection are only about half that of Chicago. Due to the different make-up of the Dallas area, there are significant differences in detections at the beginning and end of the day as compared to Chicago. We conclude that the detection rate can be quite sensitive to the geometry and purpose of the buildings. The appendix has a map of the areas simulated in both cities, and denotes the use of buildings and the number of floors.

Density of Populations.

In Fig. 4 we see that as the number of individuals increases in an area as the fraction of the population in the tracking network remains constant, the ability to track the non-participants also increases. This is expected. The sensors are increasing coverage in a finite area, as there are more detectors, and increasing potential trackees, as there are more people. However, due to the clustering behavior of humans, the increase is not as pronounced as might be expected, as the increase in detectors can result in the same area being covered multiple times. However, this does lead to more accurate trilateration.

The Effect of Network Prevalence on Detection.

As the tracking network comprises a larger percentage of the population, detection rates go up. In Fig.5, we simulate from 7-7:30am using a 15m and 30m broadcast radius to provide an upper and lower bound for 802.11g performance. The simulation is done in Chicago. We observe the effect on the percentage of individuals that can be tracked a given frequency of the time. We see that rates as low as 10% are sufficient to begin tracking people's whereabouts significant fractions of the time. For example, we see that 41% of the population can be positioned to within 15m during 50% of the time steps up to 65% at 30m. While 10% of the population dedicated to the tracking network may seem high, this percentage of detectors should be feasible if the tracking software came bundled with a popular smartphone application.

The Effect of Broadcast Diameter.

In Fig. 6, we simulate over the same 7-7:30am period with the tracking network now only comprising a conservative 1% of the individuals. We now consider the effect of extending the broadcasting range from 15m to 45m. We see a significant increase in tracking rates for range increases that are well within conservative limits for 802.11n radios. Also, the difference in performance between the inner and outer detection squares is minor compared to the increase in broadcast radius. The maximal indoor range of 802.11n connections is often quoted as 90m, so a conservative estimate of a 45m broadcast range for detecting frames is, in our opinion, still reasonable, if not conservative. With 802.11g, 30m and 45m can be considered maximum ranges under optimal conditions thus providing a low, mid, and high boundary for 802.11g communications.

Note that since the 7-7:30am period is quite conservative, a 1% infection rate for a botnet is feasible, and a 45m broadcast radius for 802.11n is also conservative, it implies that as we move to ubiquitous 802.11n equipped smartphones, the potential for malnet tracking is quite real! Based on this, we re-performed our full day experiments using 1% of the popu-

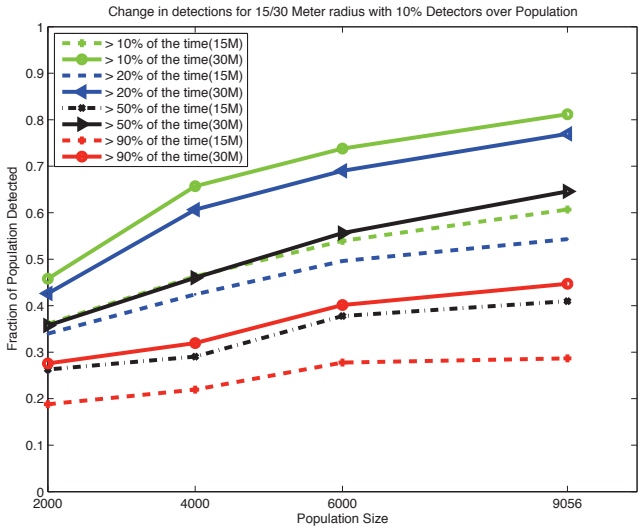

Figure 4: Simulations of Chicago from 7-7:30am for 15m and 30m broadcast radii. The tracking network comprises 10% of the pop.

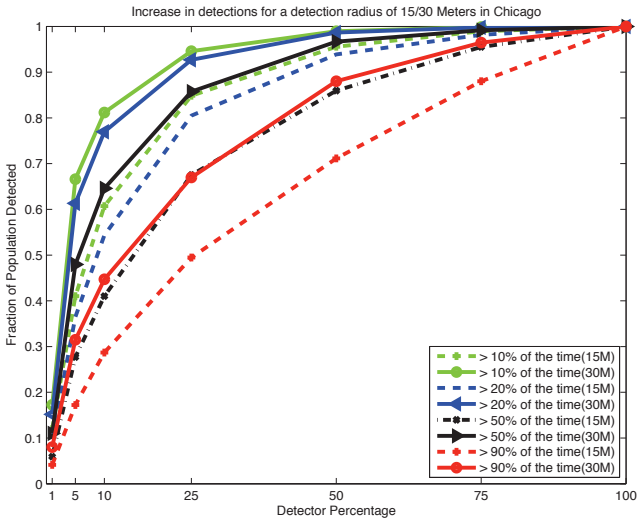

Figure 5: Simulation of 9056 people in Chicago from 7-7:30am with a 15m broadcast radius.

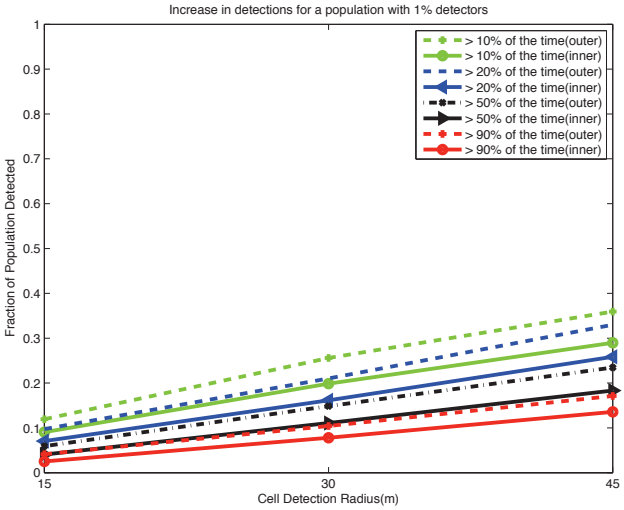

Figure 6: A simulation of 9056 people in Chicago from 7-7:30am. The tracking network is comprises 1% of the pop.

lation as detectors. Results are shown in Fig. 7. We observe that a 1% infection rate is sufficient to monitor nearly 50% of the population of Chicago on 10% of the observations throughout the day. Combined with the small standard deviations on lengths of time between observations, and we suggest that such a botnet would be effective at tracking in Chicago. Noticeably, it would seem that Dallas, would not be susceptible.

Detection occurs in three dimensions, but we measure triangulation in two dimensions: The altitude location of individuals is much less interesting then their latitude and longitude when trying to determine what buildings they are in or near by. While the fraction of the population that can be continuously monitored increases with an increase in broadcasting range, a side-effect is to decrease the accuracy to which an individual can be placed. For example, an individual detected by the tracking network at only one location assuming a 15m broadcast diameter specifies their location to approximately $30^2 = 900m^2$ area ($450m^2$ for our inner-square approx.), whereas the detection by just one location at a 45m diameter specifies their location to approximately a $8100m^2$ region ($4050m^2$ for our inner-square approximation). However, since the increase in broadcasting region allows for multiple detections, and multiple detections lead to the ability for trilateration of position, which gives a more precise position, the trade-off in accuracy is not immediately clear. In Table 1 we show the calculation for the performance of the tracking network using trilateration of position by finding the minimum area that is contained in the overlap of all the bounding detection squares. Thus increasing range results in some loss of accuracy, but due to increased observations the effect is somewhat mitigated. For example, there is a similar percentage of observations by four detectors with a 30m radius, as with two detectors at a 15m radius, yet the mean trilateration only slightly more than doubles from $559.39m^2$ to $1335.73m^2$.

5. SIMULATING HUMAN MOBILITY

We use the UDelModels simulator [4,35] to produce somewhat realistic simulations of people moving through several city blocks of downtown cities. The simulator was constructed to test the effectiveness of ad-hoc wireless routing schemes, but it suits our needs, as we are interested in the same population and its mobility dynamics. We briefly describe the simulator's key benefits in Subsection 5.2. First we justify the need for a simulator based on the lack of any feasible dataset and the unlikelihood and costliness of being able to acquire an appropriate one.

5.1 The Need for Simulation

To determine the frequency with which people can be tracked *we would ultimately like a data-set that contains the precise movements* of all *individuals with smartphones in a given metropolitan area over* a significant *period of time.* This would allow one to simulate the detection and tracking

Average Detection Area (m^2) for 10% Detectors and 9056 Pop., 7am									
Device	15 Meter Radius			30 Meter Radius			45 Meter Radius		
Obs.	μ	σ	%	μ	σ	%	μ	σ	%
≥ 1	734.94	239.66	73	2323.12	1107.00	84	4655.80	2377.01	88
≥ 2	559.39	219.04	18	1709.50	849.53	41	3815.77	1864.70	61
≥ 3	498.16	172.10	7	1437.08	742.69	24	3509.13	1736.61	44
≥ 4	481.63	184.17	3	1335.73	706.85	14	3097.04	1619.86	28

Table 1: The number of m^2 a user can be triangulated to given the number of sightings in the Obs. column. μ is mean, σ is std. dev., and % indicates the percentage of observations with the appropriate radius and minimal number of observations.

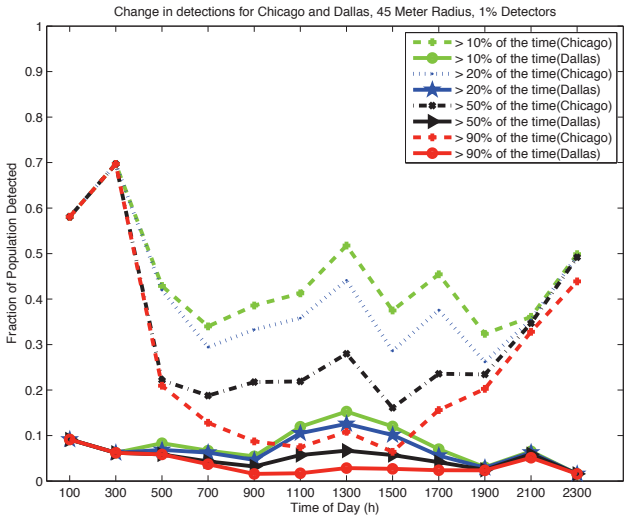

Figure 7: A simulation of 9056 people in Chicago and 2988 in Dallas. Tracking network comprises only 1% of the pop. Broadcast radius is 45m.

rates by following our methodology and simply replacing the simulated traces of individuals' positions with actual traces.

Unfortunately, creating such a dataset is extremely difficult at best and unrealistic or exorbitantly expensive at worst. A useful dataset needs to include the full number of individuals in a given region for simulation results to give us some meaningful indication of the ability to track phones over a given area. A dataset of the geolocated traces of a tiny sample of the population cannot be used in simulations to determine tracking rates for larger populations, as it is not known how to extrapolate from the area covered by the sampled individuals, their clustering, and interactions to that of a large population size. Further, peoples' behavior and thus movements are dependent on social and natural processes (e.g., rush and lunch hours), which means the dynamics of the individuals changes during the day, making any likely extrapolation process dependent on the time of day; it is clear that the relationship is far from linear.

Thus, to acquire a dataset for our modeling, we would need to get a substantial fraction of a large number of people in a metropolitan downtown core to participate in a tracking study. The task of creating this study seems unimaginable and those costs of giving them all sensors is exorbitant. Privacy concerns, mean we would be unlikely to get voluntary subject participation in the rates required. The ability to get data without their permissions (say from a mobile phone company) to the resolution necessary in the US, even if ethically handled, appears to be illegal. Since the feasibility of such an attack is most likely to be encountered in a dense metropolis, small scale experiments on suburban or campus environments are unlikely to provide useful information, as we are unlikely to get even minimal detection with small scale experiments, and we could not generalize to larger and denser populations.

Current Data Sets.

There are current data sets that have mobility and geolocation data for large groupings of people over large metropolitan and campus environments. In particular, the cellphone data set used by P.Wang et al. [55] and Gonzalez et al. [20] has the cellular phone positions of a 100,000 users in a large European city over a 6-month period. Seemingly perfect for our application, unfortunately the dataset's positional accuracy is only to the nearest cellular tower, measuring in km^2 as opposed to the m^2 accuracy we need. It also does not contain altitudinal data, something early simulations showed necessary. Finally, the locations in the dataset are only revealed when the phone placed calls or text-messages, making the timeline of the trace not continuous enough to use in simulations. The Reality Mining [15, 16] data set also has detailed logs of individual activity, but they lack the high fidelity geographic information that is required by our simulation: positioning is only to the nearest cellular towers. The data set lacks the population size is also too small at only 100 participants.

In contrast, the CRAWDAD [1] data repository has several high-resolution geolocated mappings of individuals and detections in its dataset. However, these datasets are either too sparse in population samples, not of high-enough fidelity in their positioning or not effective continuous traces through time. Further, because some of the datasets represent students or volunteer subjects performing specified tasks, many datasets do not represent natural movement patterns of a population in a metropolitan area. For example, no individuals go in to office buildings and reside there all day while they perform work.

Human Social Patterns and their Effect on Tracking.

The movement of people carrying phones is not easily statistically/deterministically modeled. While there has been work at developing a coarse granularity statistical modeling based on long time periods, such as those based on Lévy flight [9], and that of Truncated Lévy flights based on radial gyrations [20], there is little that is known over short periods of time at a high-resolution of positioning. Specifically, the authors are aware of none. However, it is essential to capture, to the best of our ability, the movements and behaviors of humans, as this greatly effects the expected coverage of

the sensor network. Specifically, assuming brownian-motion over the plane is not a reasonable simulation strategy, as has been show in other contexts such as the physical placement of routers for wireless worms [28].

5.2 The UDelModels Simulator

The UdelModels simulator is part of the UDel Models project [4]. This mobility simulator attempts to recreate accurate human mobility, both micro and macro, by producing traces that match a number of key observed statistics from a number of data sources. A comprehensive review is found in [34, 35], but we provide a very high-level overview. The UDel Mobility simulator simulates the movement of both cars and pedestrians. To accurately represent the locations of movement, the simulator uses accurate Geographic Information Systems (GIS) data to create a graph that models locations as nodes and edges as paths between locations. Human activity behavior is recreated based on information gathered from sources on topics such as urban planning, meeting analysis, and the 2003 Bureau of Labor Statistics (BLS) use of time study. Pedestrian mobility is based on research discussing worker meetings. This research determines how nodes move within buildings and between building clusters. To simulate how individuals traverse pathways between their destinations, urban planning research is used.

6. PHYSICAL REALIZATION OF A TRACKING NETWORK

In order to physically enact a tracking network, one needs to control a large number of smartphones and have them scan for appropriate BSSIDs of the targets being tracked. In order to realize such a system one needs: i) tracking software installed on a large number of smartphones in a given region; ii) a control mechanism for smartphones to monitor BSSIDs; and iii) a central system to accumulate the data. We assume the reader has a certain familiarity with 802.11 networks.

One example of a system that meets these criteria is a mobile botnet (malnet) [52]. Each device in the malnet would have to be set to a monitoring mode so that it could promiscuously detect all the WiFi users in the vicinity. This is feasible. It is in principle possible for most smartphones to promiscuously sniff traffic while in normal operation, with many hackers looking to create drivers for popular phones. For many windows mobile phones promiscuous mode software has been written.[6] While promiscuously monitoring wireless traffic, the malnet will look for probe request frames (and potentially other wireless frames), record the BSSID address of those frames, the Received Signal Strength Indication (RSSI) value associated with that frame, and the time the frame arrived. RSSI measures the strength of the radio signal detected and has historically been used as a proxy for distance in wireless positioning. This record would then be sent to a central data collection server that would process the records and triangulate the location of an individual. The amount of strain the data transfer from mobile nodes would put on the cellular data network should be negligible because the only time information would be sent to the central server is when a targeted BSSID is seen. In the case of blanket surveillance that seeks to record everyone's position, the smartphones in the tracking network could collect the data and send it in batches to ensure that network congestion is kept to a minimum. This makes an attack more viable then one that could be limited by the bandwidth in the voice portion of the cellular network such as the attack in [17, 52]. The proliferation of 3G and 4G cellular networks in major metropolitan areas makes this concern even less significant.

Another example of a system that meets the realization criteria is a geolocating advertisement network. This advertisement hosting network would provide incentives for its members to install software so that the network can track non-members. This ad hosting network now would have the ability to send geo-specific ads via various means to non-members of the network, determining their position without their consent. An Amber Alert application could be used to search for known BSSIDs of devices a missing child is known to posses and frequently carry with them, such as video-game devices, or music players. A theft retrieval network could search for the BSSIDs of stolen laptops. Apple and Google already have infrastructure in place in their smartphone operating systems to collect personal location data. Apple has even reserved the right to share location data with partners [53] while Google reserves the right to use the data internally. It would be very simple for them to modify their phone software to detect nearby probe frames as well.

Power Saving Mode Effects.

Our assumptions on 802.11n tracking performance still hold even when taking into consideration power saving modes such as PSMP, U-APSD, and S-APSD. These technologies work by buffering packets to batch send, but they are only useful when stationary. Stationary people need to be observed relatively infrequently by trackers. Spatial Multiplexing Power Saving turns off all but the 802.11a/g radios and thus are less likely to be used while in transit.

7. MITIGATING PRIVACY ATTACKS

While there may be beneficial aspects of such tracking networks, overall we think the privacy concerns make mitigation a positive strategy. The easiest ways to minimize such tracking is to ensure that the radios do not broadcast except in those cases in which they are actually being used. This has usability drawbacks, as the user will no longer be able to automatically connect to devices such as APs without explicitly initiating the connection. Yet, the less time the phone is transmitting its location, the less the potential for detection.

For more involved mitigation strategies we look towards previous solutions to other wireless location systems. In particular, Jiang et al. [31] mentions three specific defenses. The first defense is a BSSID pseudonym that changes every time a client connects to a mobile access point. The second defense is an opportunistic silent period. An opportunistic silent period is a period in time in which the client does not send any information, and after that period of time is reached, the client asks the access point for a different BSSID pseudonym. The final solution posed in the paper is decreasing the transmit power of the wireless device dynamically so that it only transmits the distance to the AP it is currently

[6]Further, new phones, such as the Droid X, can function as 802.11 APs. Therefore, they can be detected by phones not in promiscuous mode, as APs broadcast a beacon visible to all 802.11 nodes

connected to, preventing APs that are further away from detecting it.

Dynamically modifying the transmit power of the client may help in minimizing the effectiveness of a tracking network as proposed herein, as it will reduce the tracked's broadcast radius. However, we have considered scenarios where the broadcast radius is 15m, and still show high potential for tracking. It is not frequent that one is consistently less than 15m away from a known AP. Further, in cases where no known friendly APs are known, the broadcast power will be maximized. Thus we suspect there is a significant possibility that for the vast majority of cases this will have little to no effect on the ability of the tracking network.

Opportunistic silent periods and pseudonyms appear as if they will be an effective countermeasure against small numbers of probe attacks. However, if the user is under constant surveillance, and travels between known endpoints, then it is not clear if the intermittent locations cannot be reestablished through statistical analysis. This requires auxiliary information, that a tracker may learn or have, but which is not considered in Jiang [31]. Therefore, we are unclear as to how effective this countermeasure is in many practical cases.

The most in-depth solution is the SlyFi protocol developed by Greenstein et al [21]. It develops a network link layer that is identifier free. While SlyFi would be the best suited to prevent positioning attacks, the protocol was meant for actively managed networks (i.e. corporate networks). SlyFi, as discussed in [21], states that modern smartphones would have to support the traditional 802.11 link-layer protocol if the smartphones were to connect to personal routers, coffee shop routers, and other devices that are not actively managed. This support of the traditional link-layer protocol opens up SlyFi enabled smartphones to positioning attacks.

8. RELATED WORK

There is a large body of research discussing the location tracking of wireless nodes. Many systems [7, 8, 13, 19, 23, 50, 57, 58] require a radio map of the area in which the researchers were trying to track users.

Traditional methods for location tracking in wireless networks include the use of Centroids and Kalman Filters. Kim et al. [37] modified traditional techniques for positioning in cellular systems [12,24,25] to track users in the 802.11 wireless network at Dartmouth. Their system is able to replicate the paths of a human being through the campus with a high degree of accuracy. Hightower and Borriello outline a number of these techniques in "Location Sensing Techniques" [26].

A number of other novel positioning schemes have been developed. Vu and Li [54] use Delaunay triangulation to find node locations in a wireless network. Savarese et. al. [46] provide an overview of the triangulation problem of wireless networks and provide an overview of a triangulation method. They propose an algorithm called TERRAIN to use information from neighboring nodes in an ad-hoc network to triangulate the node positions. Kim et al. [38] use round-trip travel time and angle-of-arrival information to generate a location estimation algorithm.

With the increase in research on wireless location finding techniques, there has been an increase in research on location privacy and attacks on location privacy. Hu and Wang [29] designed a framework for location privacy. Tippenhauer et al. [43] actually devise a number of spoofing and false-information attacks on current publicly available wireless location positioning systems. Gruteser and Grunwald [22] as well as Jiang et al. [31] provide techniques for link layer privacy in wireless networks. Gruteser and Grunwald focus on a quantitative analysis of the effects of BSSID pseudonyms. Jiang et al. discuss BSSID pseudonyms, opportunistic silent periods, and a modification of wireless transmission power. Wong and Stajano [56] discuss the use of pseudonyms when maintaining location privacy in Bluetooth networks. Akritidis et al. bring up the idea of a metrpolitan tracking network as well, but only consider access points and the coverage problem [6].

Sensor Networks.

The surveillance we propose is similar to research done in sensor network coverage. However there are considerable differences. The coverage problem focuses on maximizing the detection area of each sensor given deterministic or nondeterministic placement, *independent of the subjects being detected*. Early sensor research dealt with the coverage problem [11, 30, 40]. Our network need only cover the locations where the vast majority of people are at any given time, and not locations where individuals are unlikely to be. A second set of research in sensor coverage looked at how directed mobile sensors would be able to improve the coverage in a traditional network [41, 42]. In our work we do not have deterministic control over any of our "sensors" (i.e. individuals with smartphones). Our "sensors" are oblivious individuals that move in their own complex patterns following their daily routine.

The latest research in sensor networks has focused on "People-centric urban sensing", or sensor networks in an urban environment using individuals as sensors [10]. One such urban sensing network is the MetroSense network constructed and discussed in [10]. Another example is the AnonySense network discussed in [14] which attempts to add privacy elements at the application level to a MetroSense style network. MetroSense creates an architecture for an urban sensor network that uses individuals as sensors. It's goal is to maximize the sensor coverage of the area the sensor network is targeted at. In this case, a large metropolitan area. The goal of the sensor network would be to measure traffic conditions, temperature, or perhaps various social conditions that might require human feedback. While these new designs for sensor networks have included humans as sensors, the goals are exactly the same as the old paradigm: maximizing the sensor coverage of a given area. This is significantly different then attempting to maximize coverage of people. Concerns over privacy were directed at individuals being identified when posting sensor results, not being targeted for surveillance. Additionally, privacy concerns were focused on application level, not at the link-layer device level.

Inaccuracies in Geolocation.

One issue, we have not addressed is error in the geolocation of the sensors. Kim et al. [36] discuss the problem of war driving router databases having inaccurate geolocations for routers, and the resulting inaccuracies this produces when the position of these routers are then used to establish the locations of other individuals through the use of WiFi. Our triangulation makes use of smartphone positions to trian-

gulate, and similar issues could apply. However, modern smartphones have enhanced GPS positioning abilities, that use a combination of GPS, cell-tower information and available 802.11 APs to appropriate locate themselves to high-degrees of accuracy. We believe it is reasonable to assume far less error than that recorded by Kim et al., and that positioning of these devices will only become more accurate over time, as alternate applications develop that require high degrees of positional accuracy.

9. CONCLUSION

Wireless mobile devices are becoming a ubiquitous aspect of modern society. The large expectation is that in the near future smartphones will be as or more ubiquitous than cellphones are today, and most likely the primary access device to the Internet for a large percentage of the world's population. Most individuals are unaware of the privacy implications that these devices bring. While previous work has been concerned about providing countermeasures to ensure privacy, little has been done to quantify the potential for pervasive monitoring. We have shown that with realistic transmission distances of 802.11g wifi, and small but not insignificant fractions of the population being part of a sensor network(malicious or otherwise), one should be able to track significant portions of the population of metropolitan areas with a high-granularity over long periods of time. We've shown that population density has an effect, but that specific geometry and building usage of the area being surveilled can have a stronger effect. Most importantly, changes in radius have significant effects on tracking capability, without huge losses in accuracy. Our results are based on a series of conservative estimates, making the upside potential for loss of privacy higher. The next generation 802.11n WiFi standard will quickly be moving to wireless phones, making tracking even easier to perform at levels that might be compatible with a malnet compromise.

For the first time the technology for creating a ubiquitous tracking network is well within the reach of digital criminals and small organizations. Thus the ability to ubiquitously physically track individuals is moving from the hands of large nation states and multi-national telecommunications corporations to small groups of organized criminals and small companies which often lack oversight. All that is required is a way of gathering up mobile nodes in a sensor network via legitimate software or malware, and a system to process the sensor data. *The potential outcomes are disconcerting, and thus we suggest that the need to strongly consider and implement current and new methods to mitigate wireless detection.*

In future work we would like to determine how effective different countermeasure are that do not analyze the potential damage that widely known auxiliary information can do in deanonymizing subjects. Further, we would like to determine if we can determine some ability to normalize simulations so they are independent of the simulated areas geometry.

As smartphones continue become the norm, we must address the privacy implications involved in these privacy-leaking devices. This is especially true as computing power becomes more accessible and pervasive, and digital criminals find new ways of profiting from insecurity.

Acknowledgements.

We thank Alessandro Vespignani and Hao Hu for discussions. We thank Dr. Stephan Bohacek, Jonghyun Kim, and Vinay Sridhara for help with their simulator. This research was supported by an Indiana University Faculty Research Support Program grant.

10. REFERENCES

[1] CRAWDAD: A Community Resource for Archiving Wireless Data At Dartmouth. *http://crawdad.cs.dartmouth.edu/*, July 2010.

[2] LandScan. *http://www.ornl.gov/sci/landscan/*, July 2010.

[3] Mobile Telephone subscribers per 100 inhabitants, 1997-2007. *http://www.itu.int/ITU-D/ict/statistics/material/graphs/Global_ICT_Dev_98-09.jpg*, July 2010.

[4] UDel Models. *http://www.udelmodels.eecis.udel.edu/*, July 2010.

[5] WIGLE.NET. *http://wigle.net/*, July 2010.

[6] P. Akritidis, W. Chin, V. Lam, S. Sidiroglou, and K. Anagnostakis. Proximity breeds danger: emerging threats in metro-area wireless networks. In *Proceedings of 16th USENIX Security Symposium on USENIX Security Symposium*, pages 1–16. USENIX Association, 2007.

[7] P. Bahl and V. Padmanabhan. RADAR: An in-building RF-based user location and tracking system. In *IEEE infocom*, volume 2, pages 775–784. Citeseer, 2000.

[8] P. Bahl, V. Padmanabhan, and A. Balachandran. Enhancements to the RADAR user location and tracking system. *Microsoft Research*, 2000.

[9] D. Brockmann, L. Hufnagel, and T. Geisel. The scaling laws of human travel. *Nature*, 439(7075):462–465, 2006.

[10] A. Campbell, S. Eisenman, N. Lane, E. Miluzzo, and R. Peterson. People-centric urban sensing. In *Proceedings of the 2nd annual international workshop on Wireless internet*, page 18. ACM, 2006.

[11] M. Cardei and J. Wu. Coverage in wireless sensor networks. *Handbook of Sensor Networks*, 2004.

[12] D. Catrein, M. Hellebrandt, R. Mathar, and M. Serrano. Location tracking of mobiles: a smart filtering method and its use in practice. In *2004 IEEE 59th Vehicular Technology Conference, 2004. VTC 2004-Spring*, volume 5, 2004.

[13] Y. Cheng, Y. Chawathe, A. LaMarca, and J. Krumm. Accuracy characterization for metropolitan-scale Wi-Fi localization. In *Proceedings of the 3rd international conference on Mobile systems, applications, and services*, page 245. ACM, 2005.

[14] C. Cornelius, A. Kapadia, D. Kotz, D. Peebles, M. Shin, and N. Triandopoulos. AnonySense: Privacy-aware people-centric sensing. In *Proceeding of the 6th international conference on Mobile systems, applications, and services*, pages 211–224. ACM, 2008.

[15] N. Eagle and A. Pentland. Reality mining: sensing complex social systems. *Personal and Ubiquitous Computing*, 10(4):268, 2006.

[16] N. Eagle, A. Pentland, and D. Lazer. Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences*, 106(36):15274, 2009.

[17] W. Enck, P. Traynor, P. McDaniel, and T. La Porta. Exploiting open functionality in SMS-capable cellular networks. In *Proceedings of the 12th ACM conference on Computer and communications security*, page 404. ACM, 2005.

[18] J. Freudiger, M. H. Manshaei, J.-P. Hubaux, and D. C. Parkes. On Non-cooperative Location Privacy: A Game-theoretic Analysis. In *ACM Conference on Computer and Communications Security (CCS)*, 2009.

[19] S. Ganu, A. Krishnakumar, and P. Krishnan. Infrastructure-based location estimation in WLAN networks. In *IEEE Wireless Communications and Networking Conference (WCNC 2004)*, volume 1, pages 465–470. Citeseer, 2004.

[20] M. González, C. Hidalgo, and A. Barabási. Understanding individual human mobility patterns. *Nature*, 453(7196):779–782, 2008.

[21] B. Greenstein, D. McCoy, J. Pang, T. Kohno, S. Seshan, and D. Wetherall. Improving wireless privacy with an identifier-free link layer protocol. In *Proceeding of the 6th international conference on Mobile systems, applications, and services, June*, pages 17–20. Citeseer, 2008.

[22] M. Gruteser and D. Grunwald. Enhancing location privacy in wireless LAN through disposable interface identifiers: a quantitative analysis. *Mobile Networks and Applications*, 10(3):315–325, 2005.

[23] Y. Gwon, R. Jain, and T. Kawahara. Robust indoor location estimation of stationary and mobile users. In *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, volume 2, 2004.

[24] M. Hata and T. Nagatsu. Mobile location using signal strength measurements in a cellular system. *IEEE Transactions on Vehicular Technology*, 29(2):245–252, 1980.

[25] M. Hellebrandt and R. Mathar. Location tracking of mobiles in cellular radio networks. *IEEE Transactions on Vehicular Technology*, 48(5):1558–1562, 1999.

[26] J. Hightower and G. Borriello. Location sensing techniques. *IEEE Computer*, 2001.

[27] B. Hoh, M. Gruteser, H. Xiong, and A. Alrabady. Preserving privacy in gps traces via uncertainty-aware path cloaking. In *Proceedings of the 14th ACM conference on Computer and communications security*, page 171. ACM, 2007.

[28] H. Hu, S. Myers, V. Colizza, and A. Vespignani. WiFi networks and malware epidemiology. *Proceedings of the National Academy of Sciences*, 106(5):1318, 2009.

[29] Y. Hu and H. Wang. A framework for location privacy in wireless networks. In *ACM SIGCOMM Asia Workshop*. Citeseer, 2005.

[30] C. Huang and Y. Tseng. The coverage problem in a wireless sensor network. *Mobile Networks and Applications*, 10(4):519–528, 2005.

[31] T. Jiang, H. Wang, and Y. Hu. Preserving location privacy in wireless LANs. In *Proceedings of the 5th international conference on Mobile systems, applications and services*, page 257. ACM, 2007.

[32] A. Karlson, B. Meyers, A. Jacobs, P. Johns, and S. Kane. Working Overtime: Patterns of Smartphone and PC Usage in the Day of an Information Worker. *Pervasive Computing*, pages 398–405.

[33] D. Kellogg. iPhone vs. Android. *nielsenwire*, June 2010.

[34] J. Kim. Realistic mobility modeling and simulation for mobile wireless network in urban environments. 2005.

[35] J. Kim, V. Sridhara, and S. Bohacek. Realistic mobility simulation of urban mesh networks. *Ad Hoc Networks*, 7(2):411–430, 2009.

[36] M. Kim, J. Fielding, and D. Kotz. Risks of using AP locations discovered through war driving. *Pervasive Computing*, pages 67–82.

[37] M. Kim, D. Kotz, and S. Kim. Extracting a mobility model from real user traces. In *Proceedings of IEEE Infocom*, volume 6. Citeseer, 2006.

[38] S. Kim, A. Brown, T. Pals, R. Iltis, and H. Lee. Geolocation in ad hoc networks using DS-CDMA and generalized successive interference cancellation. *IEEE Journal on Selected Areas in Communications*, 23(5):984–998, 2005.

[39] L. Latif. Iphone loses market share. *The Inquirer*, Mar. 2010.

[40] X. Li, P. Wan, and O. Frieder. Coverage in wireless ad hoc sensor networks. *IEEE Transactions on Computers*, 52(6):753–763, 2003.

[41] B. Liu, P. Brass, O. Dousse, P. Nain, and D. Towsley. Mobility improves coverage of sensor networks. In *Proceedings of the 6th ACM international symposium on Mobile ad hoc networking and computing*, page 308. ACM, 2005.

[42] J. Luo, D. Wang, and Q. Zhang. Double mobility: coverage of the sea surface with mobile sensor networks. *ACM SIGMOBILE Mobile Computing and Communications Review*, 13(1):52–55, 2009.

[43] N.O. Tippenhauer and K.B. Rasmussen and C. Popper and S. Capkun. Attacks on public wlan-based positioning systems. In *Proceedings of the 7th international conference on Mobile systems, applications, and services*, pages 29–40. ACM, 2009.

[44] L. Privat. Nielsen: US Smartphone Penetration to Be over 50% in 2011. Mar. 2010.

[45] K. Rasmussen and S. Čapkun. Location privacy of distance bounding protocols. In *Proceedings of the 15th ACM conference on Computer and communications security*, pages 149–160. ACM, 2008.

[46] C. Savarese, J. Rabaey, and J. Beutel. Location in distributed ad-hoc wireless sensor networks. In *Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01). 2001 IEEE International Conference on*, volume 4, pages 2037–2040 vol.4, 2001.

[47] L. Snol. 802.11n in 87% of Wi-Fi smartphones by 2014. *Computer World*, Jan. 2010.

[48] C. Song, Z. Qu, N. Blumm, and A. Barabasi. Limits of predictability in human mobility. *Science*, 327(5968):1018, 2010.

[49] A. Spong. Report: Wifi is a must for smartphones. *phonemag*, Apr. 2009.

[50] N. Swangmuang and P. Krishnamurthy. Location fingerprint analyses toward efficient indoor positioning. In *Proceedings of the 2008 Sixth Annual IEEE International Conference on Pervasive Computing and Communications*, pages 100–109. IEEE Computer Society, 2008.

[51] P. Traynor, K. Butler, W. Enck, P. McDaniel, and K. Borders. malnets: large-scale malicious networks via compromised wireless access points. *Security and Communication Networks*, 3(2-3):102–113, 2010.

[52] P. Traynor, M. Lin, M. Ongtang, V. Rao, T. Jaeger, P. McDaniel, and T. La Porta. On cellular botnets: Measuring the impact of malicious devices on a cellular network core. In *Proceedings of the 16th ACM conference on Computer and communications security*, pages 223–234. ACM, 2009.

[53] J. Valentino-DeVries. Apple changes privacy policy to collect location data. *Wall Street Journal*, June 2010.

[54] C. Vu and Y. Li. Delaunay-triangulation based complete coverage in wireless sensor networks. *IQ2S2009 in conjunction with PERCOM*, pages 9–13, 2009.

[55] P. Wang, M. Gonzalez, C. Hidalgo, and A. Barabasi. Understanding the spreading patterns of mobile phone viruses. *Science*, 324(5930):1071, 2009.

[56] F. Wong and F. Stajano. Location privacy in bluetooth. *Security and privacy in ad-hoc and sensor networks*, pages 176–188, 2005.

[57] Z. Xiang, S. Song, J. Chen, H. Wang, J. Huang, and X. Gao. A wireless LAN-based indoor positioning technology. *IBM Journal of Research and Development*, 48(5-6):617–626, 2004.

[58] J. Yin, Q. Yang, and L. Ni. Adaptive temporal radio maps for indoor location estimation. 2005.

APPENDIX

In Figs. 11 and 9 we see the area modeled in Chicago. Specifically, it is the 3×3 block area bound on the north by W. Wacker Dr., the south by W. Washington St., the east by N. Clark St and the west by N. Franklin St. In Figs. 10 and 8 we see the area modeled in Dallas. It is the 4×4 block area bound on the north by Pacific Ave., the south by Jackson St., the west by N Field St., and the east by North St. Paul St. In both cases, the figure on the right gives the simulator's parameters, with the number of floors of the building inscribed on the top, and its designation as an office (OF), residence (RE), or service building (SR). Some buildings have multiple designations, as services are provided on ground and lower levels, and other uses for higher levels. The UDel simulator makes use of this information in its path and destination planning during simulations.

Figure 8: The 16 Block area in Dallas (USGS Urban Area Ortho).

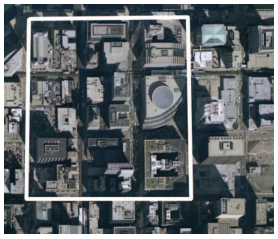

Figure 9: The 9 Block area in Chicago (USGS Urban Area Ortho).

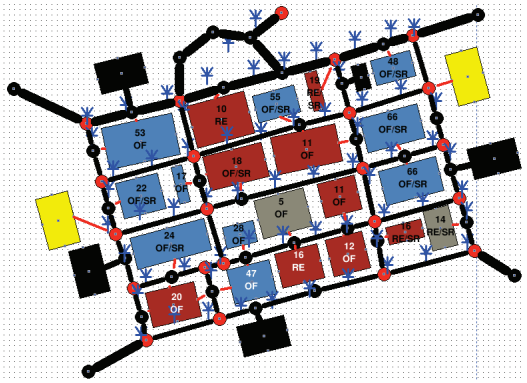

Figure 10: Map of the Dallas Area created by the UDel Models group [4].

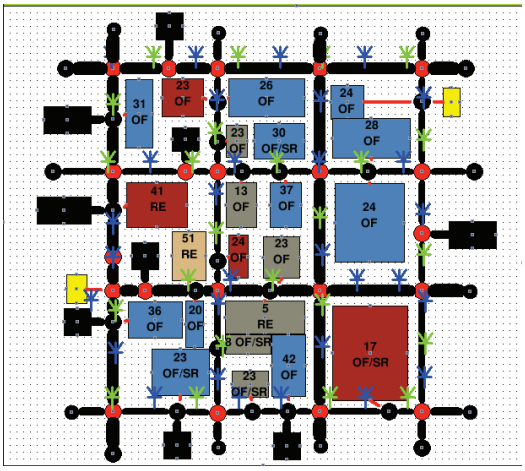

Figure 11: Map of the Chicago Area created by the UDel Models group [4].