Evaluation Framework of Location Privacy of Wireless Mobile Systems with Arbitrary Beam Pattern

Ford Long Wong, Min Lin, Shishir Nagaraja, Ian Wassell and Frank Stajano
University of Cambridge
{fw242, ml406, sn275, ijw24, fms27}@cam.ac.uk

Abstract

Position localization of transmitters can be carried out by an adversary owning a network of pervasive receivers, which can pinpoint the victim mobile nodes' locations with high temporal and spatial accuracy, such that pseudonym changing and higher-layer obfuscation are insufficient to protect their location privacy. Our contribution is to consider covert beam patterns, generated using multiple antennas to do adaptive beamforming, and so reducing radio signature. We architect such a privacy-enhancing system, give an informal security analysis, and develop an evaluation framework to analyze its location privacy. We performed simulations using wireless LAN parameters, and found that signal-to-noise-ratios for successful direction-finding are more stringent than those required for mere communications. We composed an end-to-end integrated radio and mobility simulation, and compared location privacy performance of omnidirectional versus adaptive beamforming antennas. Our proposal is shown to perform better. In addition, our evaluation framework is flexible and extensible.

1 Introduction

The proliferation of always-on wireless personal devices has been accompanied by an attendant loss of privacy: not only is it easier for radio communications to be overheard; more subtly, our movements can be silently traced by an adversary who observes the location of our transmissions. The latter is an attack on our location privacy. It is very hard to guard against such an adversary if it is the wireless connectivity provider, and we are not attempting to do so: we assume him to be trustworthy. Protecting location privacy is difficult because information can leak at many levels. A mobile user can be recognized at different locations by a global observer looking for a variety of identifiers in the intercepted packets, at different levels in the protocol stack.

A number of other papers [2, 3, 10, 11, 12, 24] have discussed location privacy for upper layers and have proposed countermeasures such as degrading the location information or changing pseudonyms at appropriate intervals. These are however insufficient in the presence of a passive observer at the lowest (physical) layer. Such an adversary could detect the location of the victim by bypassing the degradation and could observe the changeover from one pseudonym to the other, thereby re-linking the two pseudonyms of the victim.

The essence of our proposal is to change the transmission mode from omnidirectional to a shaped beam, for transmission from mobile station to base station. Thus, the attacker needs to be within the coverage of the beam pattern, which reduces the chances of a successful attack. The capability to form a shaped beam is not common in currently deployed hardware, but is expected to become widespread in the next generation of products owing to advantages in terms of lower interference and greater power efficiency. Thus, one of our contributions is the idea that the multiple element antennas in upcoming digital radios originally intended to do MIMO (Multiple Input Multiple Output) for better communication performance can be exploited to do adaptive beamforming to make mobile terminals harder to detect by an adversary.

2 Threat Model

You carry a wireless-enabled laptop. Your laptop maintains connectivity to the Internet through base stations connected to a network backbone, and they may operate, by 802.11 wireless LAN for example, as shown in Fig. 1. To maintain your location privacy, the wireless MAC of your laptop periodically changes disposable pseudonyms [11]. The adversary does not control the public network of base stations, but he controls a network of radio receivers. We assume the adversary uses only radio and no visual surveillance. His objective is to track your movement over pseudonym changes. As long as he can maintain a high enough resolution location fix on your emitting mobile node, even if the anonymity set of the perceived mix

zone [2, 3] is large, he can compromise your linkability significantly. High-level anonymizing mechanisms, such as pseudonyms, would fail under the above threat model. The effect of using changing pseudonyms is: firstly, to obscure the real identity of the user, and secondly, to confer unlinkability between different pseudonyms of the same user. Pseudonyms fail if the attacker can continue to pinpoint the nodes even as these change pseudonyms within mix zones.

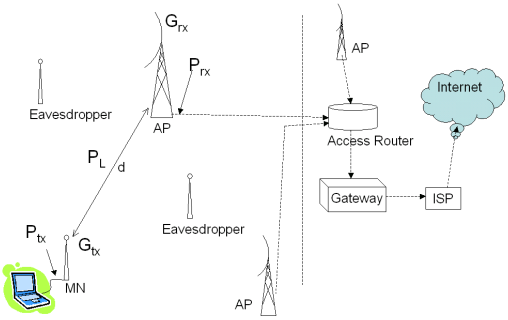


Fig 1. Adversary Model

3 Proposal: Low-Level Countermeasure through Adaptive Beamforming

Advances in RF frontends and digital electronics have made it possible to produce highly directional antenna patterns using adaptive beamforming. Smart antennas are increasingly being used in wireless LANs and cellular communications in their base station antennas. Smart antennas direct radio energy towards the intended users by beamforming, and reduce energy away from unintended users. The use of multiple antennas at both ends capitalizes on spatial multiplexing and increases data throughput, a concept epitomized by the MIMO model. Smart antennas also mitigate any negative effects of multipath fading. For location privacy, we are more concerned with multiple antennas' beamforming property – we propose for mobile nodes to use smart antenna to produce beam patterns with narrow main lobes and low sidelobes. It is worth noting that what we are proposing is actually a different use of multiple antenna elements, compared to MIMO. In MIMO, generally *multiple* beams are created by an antenna array, and can be used to take advantage of transmit diversity, whereas in our proposal, we advocate a single beam, to achieve a *covert* mode.

4 System Level Strategy

4.1 Assumptions

We make the following assumptions:

(1) Mobile nodes have a robust method of estimating their own location (new generation GPS receivers themselves are able to use technologies such as adaptive beamforming [4] or analog electronics, to suppress the effect of jammers and spoofers). Mobile nodes are also assumed to have a means of accurately estimating their azimuthal bearing, such as by means of a compass.

(2) The intended physical environment ranges from (a) flat unobstructed terrain to (b) moderately urban environment. If the environment is composed of many RF scatterers and reflectors, then the attacker's job of finding the direct and dominant signal path is made very difficult too.

4.2 System

The main features of the system are the following:

(1) Mobile nodes have a secure way of learning the true coordinates of the base stations. These coordinates could be pre-distributed by a trusted authority, signed and attested.

(2) A mobile node will shape a radio beam in the direction of the base station it is associated and communicating with, based on knowledge of their positions.

(3) Mobile nodes will carry out mutual authentication with base stations, as a further protection against an attacker which attempts to trivially spoof a base station.

(4) The mobile node will not emit much power outside its main lobe; this is dependent on the antenna array geometry.

5 Security Considerations

We analyze security issues raised by our proposal, without elaboration here due to space constraints:

(1) Base Station Spoofing: This is countered by pre-distributing trusted base stations' coordinates. Mobile nodes will ignore beacon frames as an indicator of position.

(2) Self-Positioning Inaccuracies: Radio-navigation receivers with shaped antenna having deep nulls [4], resistant to spoofing and jamming attacks are becoming available. Tolerances may be built into nodes' beam-pointing to account for slight self-positioning imprecision. (For indoor use, inertial navigation systems can be cued from GPS.)

(3) Inter-Layer Linkability: De-linking between different layers of the protocol stack is important to guarantee overall anonymity. We will assume these have been done.

(4) Samples for Tracking Missing: We can relate our work to previous work in multi-target tracking [12] (MTT). Our method causes *missing samples* for the attacker's MTT.

6 Radio Environment

We consider a communication system similar to the WLAN IEEE 802.11b system.

6.1 Link Budget

The radio link budget equation is given as: $P_{tx} + G_{tx} - P_L + G_{rx} = P_{rx}$, where P_{tx} is the transmitter output power (of the mobile node), G_{tx} is the transmitter antenna gain, P_L is the path loss, P_{rx} is the power received at the receiver of the base station, and G_{rx} is the antenna gain of the receiver. It refers to the uplink. The values are given in dB, dBi or dBm as appropriate. We used FCC regulatory limits [8] in our analysis; ETSI limits can likewise be studied.

6.2 Loss Propagation Model

Selecting a suitable propagation model so as to compute a most realistic path loss P_L is one of the key aspects of this work. For a link without obstruction, the standard free space loss model is:

$$Loss = 32.45 + 20\log(D/\mathrm{km}) + 20\log(f/\mathrm{MHz})$$

But this oversimplified model is inadequate. We found the COST-Hata Model [7] most suitable. It is not site-specific to the extent of requiring every physical structure to be modelled. It had been fitted to observed power degradation curves through exhaustive field studies of mobile environments similar to what we are considering.

$$Loss = 46.3 + 33.9\log(f/\mathrm{MHz}) - 13.82\log(h_{base}/\mathrm{m}) - a(h_{mobile}/\mathrm{m}) + (44.9 - 6.55\log(h_{base}/\mathrm{m}))\log(D/\mathrm{km}) + C_m$$

The parameters are f for frequency, h_{base} for the base station height, h_{mobile} for the mobile node height, D for the distance, a is a function [7] and C_m is a constant.

6.3 Antenna Gain Pattern

In adaptive beamforming, the operations of phase shifting and amplitude scaling for each antenna element are carried out adaptively. A beampattern is basically dependent on the array geometry (i.e. the number of elements, the spacing and the aperture size), the transmitted carrier frequency, the gain pattern of each element, and the antenna weights. The array factor (AF) representing the gain for the antenna is described analytically [1] as follows, substituting $N = 4$ for a 4-element antenna.

$$AF = \sum_{n=1}^{N} e^{+j(n-1)(kd\cos\theta+\beta)}$$

where k is the wave number and β is the excitation phase. The expression for the radiation pattern is:

$$E = a_\phi j\eta \frac{kI_o e^{-jkr}}{4\pi r} \times AF$$

where the first term is the electric field description, which will be replaced with our normalized transmission power and an antenna gain of unity. The second term is the array factor describing of the radiation pattern in a vertical dipole. Choosing values so as to achieve appropriate tradeoffs between minimizing sidelobe power and minimizing mainlobe beamwidth, and assuming such antenna can be synthesized, we can derive the array factor in Fig. 2.

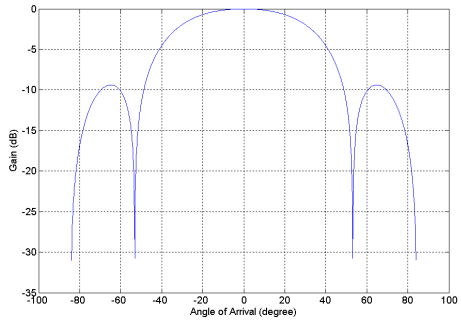


Fig 2. Array Factor of a 4-element Antenna Array

6.4 Attacker's Direction-Finding

Due to tolerances and known calibration issues in direction-finding systems, and environmental factors, localizing an emitter in the field is not as trivial as simulations suggest, but these upper-bound an attacker's effectiveness.

Position-localization via Received-Signal-Strength-Indicator is unreliable, because the transmitter power may be unknown and received power may fluctuate due to multipath propagation and shadowing. Time-Of-Arrival and Time-Difference-Of-Arrival methods require the signal to have a very short pulse-width to be effective, such as Ultra-WideBand signals. The reliable option left against WLAN-type systems is Angle-of-Arrival (AOA) estimation by two or more sensors, followed by triangulation.

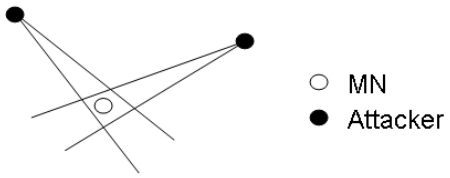


Fig 3. Geometrical error constraints using AOA

Signal Model

The signal model is as follows. There are $M \leq L$ narrowband sources (where L is the number of antenna elements) centred at frequency ω_0. We also assume the sources

are deterministic. Additive noise is modelled as a stationary zero-mean random process. To mathematically describe the effect of the translational invariance of the antenna array, we apply the conventional practice of treating the array as being comprised of two sub-arrays, Q_x and Q_y, which are identical in every respect, except that they are physically separated by a known displacement, Δ. The received signal model at the ith antenna element can be expressed as:

$$\begin{aligned} x_i(t) &= \sum_{k=1}^{M} s_k(t)a_i(\theta_k) + n_{xi}(t) \\ y_i(t) &= \sum_{k=1}^{M} s_k(t)e^{j\omega_0 \Delta sin\theta_k/c}a_i(\theta_k) + n_{yi}(t) \end{aligned}$$

where θ_k is the direction-of-arrival of the kth source relative to the direction of the translational displacement vector Δ.

Combining the outputs of each element in the two sub-arrays, the received data vector can be derived as:

$$\begin{aligned} \mathbf{x}_i(t) &= \mathbf{A}\mathbf{s}(t) + \mathbf{n}_x(t) \\ \mathbf{y}_i(t) &= \mathbf{A}\mathbf{\Phi}\mathbf{s}(t) + \mathbf{n}_y(t) \end{aligned}$$

where $\mathbf{s}(t)$ is the $M \times 1$ vector of impinging signals (wavefronts) as observed at the reference sub-array (in our case, we choose this to be Q_x), and the matrix $\mathbf{\Phi}$ is a diagonal $M \times M$ matrix of the phase delays, and is given by:

$$\mathbf{\Phi} = diag\{e^{j\beta_1}, \cdots, e^{j\beta_M}\}$$

where $\beta_k = \omega_0 \Delta sin\theta_k/c$

Finding the Angle-Of-Arrival (AOA)

The MUSIC (MUltiple SIgnal Classification) algorithm [21] was one of the first significant algorithms for doing AOA estimation, then ESPRIT (Estimation of Signal Parameters via Rotational Invariance Techniques) [19] emerged. MUSIC is powerful, however, ESPRIT is compelling because it is much less computationally intensive, and merely requires the array to be linear. We used the Total Least Squares-ESPRIT (TLS-ESPRIT) [19] version, which was introduced to make the estimation more practical when only a finite number of noisy measurements is available.

We apply the TLS-ESPRIT algorithm as follows:
(1) We define L as the number of elements in the direction-finding array. Since we are assuming a 4-element array, so $L = 4$. We next define a L-dimensional random vector $\bar{x}$ corresponding to L consecutive data samples. We estimate the correlation matrix $\hat{R}_{\bar{x}}$ from the data.

(2) We compute the generalized eigenvectors and eigenvalues of $\hat{R}_{\bar{x}}$, where k ranges from 1 to L :

$$\hat{R}_{\bar{x}}\bar{e}_k = \bar{\lambda}_k \Sigma_{\bar{\eta}} \bar{e}_k$$

(3) Usually there is a need to first estimate the number of sources, M. (Several algorithms exist to compute this.) The maximum number of signal sources whose direction can be estimated is limited to be 1 less than the number of antenna elements. Thus, a 4-element array is unable to estimate for more than 3 signal sources simultaneously.
(4) We next generate a basis spanning the signal subspace and partition it as $\bar{Q}$, where $\bar{Q} \in \mathbb{C}^{L \times M}$.

$$\bar{Q} = \Sigma_{\bar{\eta}} \begin{bmatrix} | & & | \\ \bar{e}_1 & \cdots & \bar{e}_M \\ | & & | \end{bmatrix} = \begin{bmatrix} Q \\ \times \cdots \times \end{bmatrix} = \begin{bmatrix} \times \cdots \times \\ Q' \end{bmatrix}$$

(5) We perform singular value decomposition on: $[\,Q \quad Q'\,]$ We then extract V as the right singular vector, and partition V into four $M \times M$ submatrices:

$$V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}$$

(6) Next we compute the eigenvalues λ_1, λ_2,... λ_M of the matrix Ψ_{TLS} which is equal to $-V_{12}V_{22}^{-1}$.
(7) We come to our last step, to compute the angles of arrival, where δ is the separation between adjacent antenna elements:

$$\theta_k = sin^{-1}\frac{Arg(\lambda_k)}{2\pi\delta}$$

We are not considering an adversary who might use physically steered directional antennas (such as Yagi antennas), because these devices would firstly, be cumbersome, and secondly, allow only one signal source to be tracked at a time per receiver.

7 Radio Simulation

7.1 Path Loss

Using procedures outlined earlier, we calculate the Signal-to-Noise Ratios (SNRs) for a range of Effective Isotropic Radiated Power (EIRP) values for the uplink. The upper end is the FCC limit. The lower end can be considered as the mobile device exercising power control to limit its output power to just 20 dBm, and using an omnidirectional antenna. Many PC card wireless adapters today have omnidirectional antennas with a gain of 0 dBi (or equivalently, the mobile node could also be transmitting at less than 20 dBm, and using a high gain antenna). We assume a reasonable receiver noise figure (NF) of 10 dB for the base

station. The receiver antenna gain, G_{rx} is set at 6 dBi, again a typical figure, for many current BS antennas. Fig. 4 shows the degradation of SNR with increasing distance, for different uplink EIRP, using the previously stated assumptions.

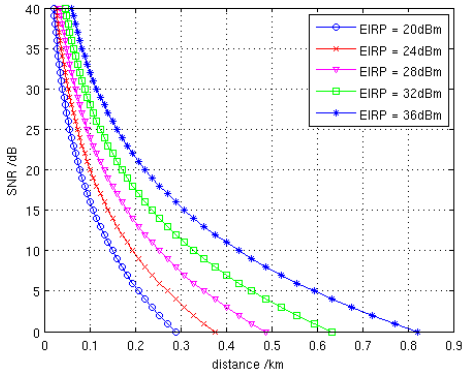


Fig 4. SNR versus distance for different EIRP

Note that an attacker who invests in a more expensive low-noise receiver will improve the SNR of the received signals for a given distance.

7.2 Angle-Of-Arrival Estimation Errors

We investigate the performance attainable using the TLS-ESPRIT algorithm to intercept wireless LAN type of signals. We make the assumption that the attacker uses a 4-element linear array for his receiving station. (If the attacker invests in more antenna elements for his system, he can obtain better performance, provided the channel characteristics do not change while he takes a longer time to collect the signals, having more elements.) Angle-of-Arrival estimation is carried out and the the mean-square-error (MSE) of the estimation is computed. The plot of the MSE of AOA estimates with different SNRs is shown in Fig. 5. The SNR refers to that experienced at the *attacker's receiver*. Again, we assume that the receiver NF is 10 dB.

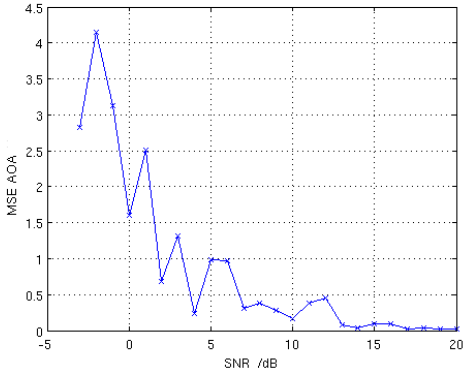


Fig 5. MSE of AOA estimation vs SNR

Fig. 5 shows that the AOA estimates converge to an acceptable accuracy when the SNR is 10 dB or better. At lower SNR levels, the fluctuation is quite significant, often giving more than 1 deg MSE. At SNR levels of 15 dB or more, the MSE flattens out. Thus, the adversary who attempts to do Angle-of-Arrival estimation using TLS-ESPRIT needs to be placed in a position where he can receive the MN's uplink signal at a SNR value of 10 dB or better, otherwise his estimate will be very error-prone. To give an idea of the error, if the MSE of the AOA estimate is 1.4 deg^2 and the attacker is 150 metres away from the MN, then the error in distance is around ± 5 metres.

The above simulations indicate the **upper bound** on the attacker's direction-finding ability with the said set-up. In real field conditions, there will be imperfections in equipment. For example, frequency stability of the clock used for demodulating and sampling the measured signals affects the AOA estimation accuracy [16]. This is shown in Fig. 6.

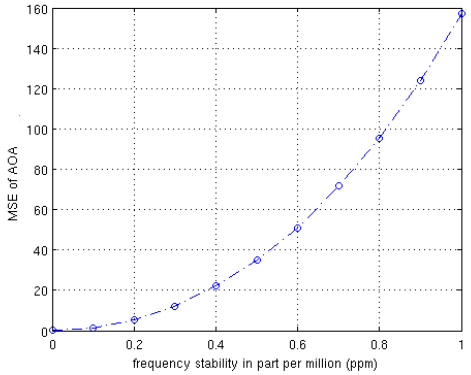


Fig 6. MSE of AOA estimation vs frequency stability

Briefly, the frequency offset (i.e. phase rotation observed at the receiver) is caused by differences in the oscillator reference frequencies at the transmitter and observer. The frequency shift depends on the frequency stability of the reference clocks, specified in parts per million (ppm). If both the transmitter and receiver have different clock accuracies, then the maximum frequency accuracy in terms of ppm is

$$f_{ppm(max)} = f_{ppm(max)rx} + f_{ppm(max)tx}$$

The maximum phase rotation over a burst of N samples with a sample rate r_s is:

$$\theta_{max} = (2\pi \times f_c \times f_{ppm(max)rx}) \times \frac{N}{r_s}$$

where f_c is the carrier frequency. Consequently, the incremental phase rotation for each received sample is:

$$\theta_i = \frac{\theta_{max}}{N}$$

For example, if we have maximum frequency offset f_Δ of 300 Hz, the frequency stability will be:

$$f_{ppm} = \frac{f_\Delta}{f_c} = \frac{300Hz}{2.4GHz} = 0.125 \times 10^{-6} (\text{i.e. } 0.125 \text{ ppm})$$

Fig. 6 shows the very significant deterioration of AOA estimation accuracy as the frequency stability worsens from 0.1 ppm to just 1 ppm. From 0.3 ppm upwards, errors make the AOA estimate rather unusable. Obtaining 0.1 ppm frequency stability is costly; the attacker is forced to make significant investment in equipment.

7.3 Security Margin

For a fixed uplink EIRP of 20 dBm, Fig. 7 shows the distance for successful communication between the MN and the BS, and the distance for successful AOA estimate of the MN by an attacker.

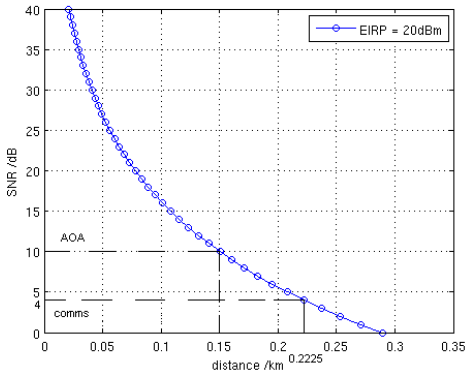
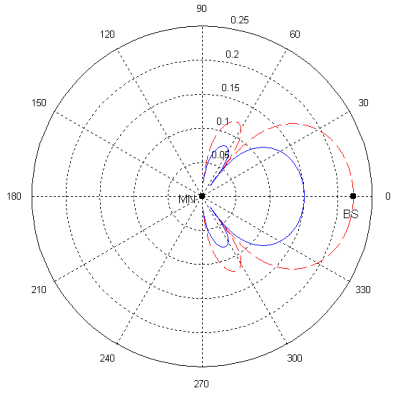


Fig 7. SNR required for AOA estimation and for successful comms

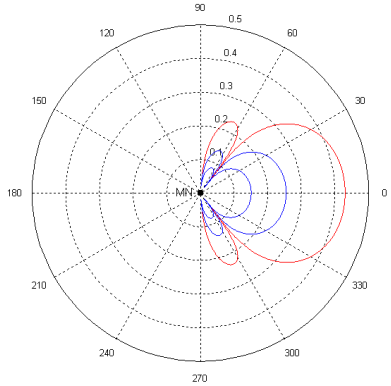
For transmission with a data rate of 1 Mbps with a bit-error-rate better than 10^{-5}, a SNR level of around 4 dB is required at the receiver [14]. As shown in Fig. 7, the BS is allowed to be 222.5 metres away from the MN. On the other hand, the attacker needs to be as close as 150 metres away from the MN to be able to perform good AOA. This difference in linear distance can be thought of as a sort of ***security margin***.

7.4 Trade-off of Data Rate with Security

We calculate the SNR levels around a MN transmitting with the said shaped beam (from Fig. 2). We assume it uses a fixed EIRP of 20 dBm. Fig. 8(a) shows the plan view of the BS position relative to the MN. The BS receives with a SNR of 4 dB at this range (- outer lobe). The concentric grid lines show the distances (in km) away from the MN. The



(a) Security Margin



(b) Data Rate/Security Trade-off

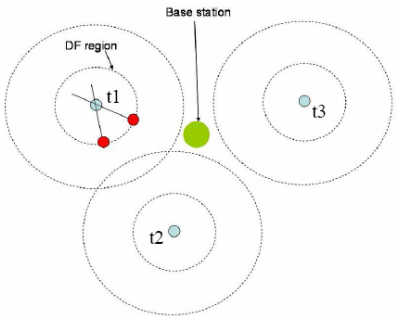
Fig. 8 Plan View of Radiation Coverage

adversary is required to be located within the *inner* lobes, to be able to direction-fix the MN (it follows from Fig. 7).

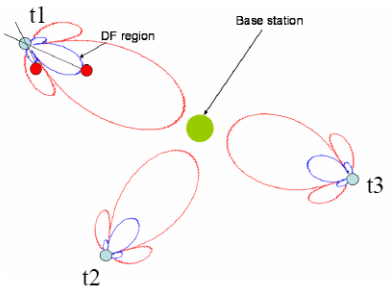
In the earlier graphs, the MN is transmitting with a power level just sufficient to communicate at 1 Mbps with a BS 222.5 metres away. We now consider the security impact of increasing the data rate while retaining the same BS distance in Fig. 8(b). A SNR of 4 dB at the BS receiver is required for a 1 Mbps data rate for 802.11b, but a SNR of 8 dB is required for a 5.5 Mbps data rate [14]. A higher data rate necessitates a higher transmit power, invariably increasing the mobile node's radio signature even with a shaped beam, and enlarging the vulnerable area. The lobes in Fig. 8(b) correspond to the contours at which SNR = 10 dB for EIRP values of 36, 28 and 20 dBm. Transmitting with an EIRP of 20 dBm, the MN can be direction-fixed from 150 metres away, whereas if it transmits with an EIRP of 28 dBm (still well within FCC regulations), giving it a data rate better than 5.5 Mbps at the BS at the same position, it can be direction-fixed from 250 metres away, both distances referring to the main lobe direction. Clearly, data rate trades off with location privacy, and needs to be carefully managed.

8 Integrated Radio and Mobility Simulation

We used the IBM City Simulator [15] to generate node mobility output to drive the location anonymity analysis under adaptive beamforming (ABF) and omni-directional (OMNI) antenna radiation patterns. It simulates realistic motion of people moving in a city, carrying mobile wireless devices. We placed 100 mobile nodes on a grid, all communicating with one base station. The adversary places an increasing number of receivers at random points, whose locations are fixed for the length of the simulation. The job of the adversary receiver equipment is to collect as much signal direction information as they can. The adversary aims to learn as much location information as possible. We wish to examine how the change of radiation pattern affects such information collection, and in turn location privacy.



(a) Omni Beams



(b) ABF Beams

Fig. 9 Integrated Radio and Mobility Simulation

Each mobile node exercises power control and optimizes transmission power so as to achieve an SNR of about 4 dB at the base station's receiver. The radio footprint that results is derived from procedures outlined earlier. An adversary receiver needs to be within the radiation zone of a SNR of 10 dB or better to be able to carry out direction-finding. These are simplistically represented in Fig. 9. t_1, t_2 and t_3 represent the different time snapshots of one moving node.

9 Location Privacy Performance

Location privacy attacks depend on being able to track mobile nodes at a frequency high enough to reveal movement information. We consider the attacks in phases:

(1) Attacks that require the location information of as many nodes as possible at given points in time, without linking.

(2) *Location linking attacks* - that attempt to bind location information to unique nodes, hence linking as many positions and movements of a mobile node as it can.

9.1 Direction-Finding and Triangulation

In the first phase, the attacker attempts to gather as many direction estimates of the mobile nodes as possible, out of around 20×10^3 samples. We assume a detection if an attacker receiver falls within the beam coverage thresholded at SNR = 10 dB of a mobile node. If the attacker has two or more receivers successfully direction-fixing a victim node at a time instant, he can derive a triangulation of the node at that point in time. The results for both omnidirectional and adaptive beamforming beams are shown in Fig. 10. ABF is shown to be 6 to 7 times more covert than OMNI radiation pattern when the adversary uses a low number of receivers. Even when the adversary invests in a large number of receivers (eg. 10), the former still performs 3 times better. We also see that going from direction-finding (DF) to position-localization (PL), the attacker's success rate degrades more sharply against the ABF beams than for the OMNI beams.

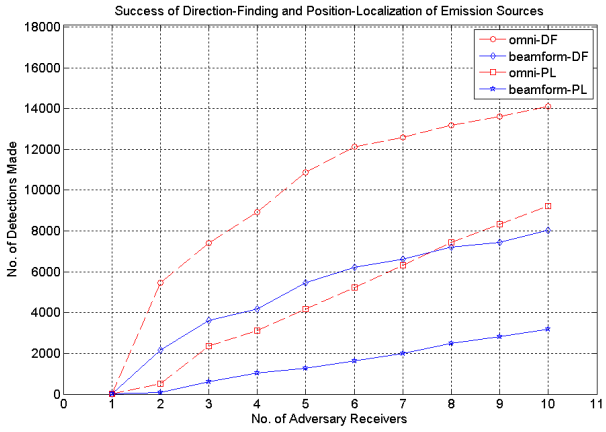


Fig 10. Direction-Finding and Position Localization Success

9.2 Location Linkability Attacks and Blackout Window

Location linking is a powerful attack by which the attacker uses location information obtained at various points

in time to de-anonymize an otherwise pseudonymous node and determine the path that it took. To carry this out, the attacker needs to know not only the location information of as many nodes as possible at various times, but to link them, he can only 'lose' them for a small window of time. We refer to this time as the ***blackout window B***, and use this as a metric of the robustness of different linking algorithms. The larger the value of 'B' that an algorithm can tolerate, the less its performance is impeded by factors such as a reduction in the radiation area. In the absence of precise simulations of linking algorithms (such as [12]), we approximate a linking attack's robustness to the size of its tolerable blackout window. (In view of the different multi-target tracking algorithms already existing, this is not a fine-grained approximation, but it arguably serves our current purpose to just compare performance between different beam patterns.)

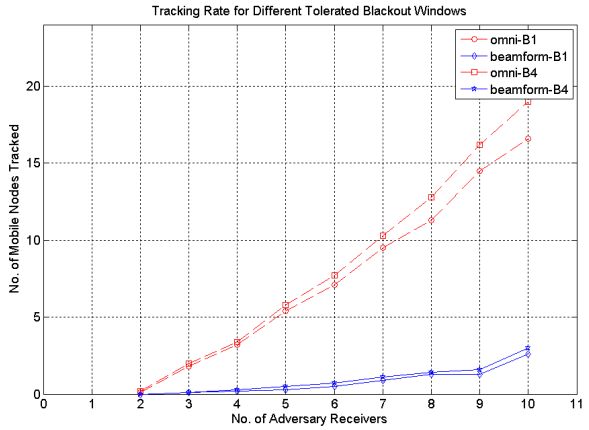


Fig 11. Tracking Rate

Conceptually, if S_t is the set of nodes whose movement the attacker has been able to track until time t, and let l_t be the corresponding location information, then at any time $t < t_k \leq t + B$, the attacker will be able to link position l_{t_k} with S_t with a probability close to 1. A linking attack algorithm loses tracks when it encounters a blackout larger than its B. Fig. 11 shows the tracking success for the two beam patterns for two sizes of B (specifically 1 and 4). ABF is shown to be 6∼7 times better in resisting tracking. For the adversary, using more robust linking algorithms yields benefits equivalent to deploying more receivers.

9.3 Information-Theoretic Location Privacy

There are several ways one can express the quality of anonymity a system provides. Qualitative tags such as *absolute privacy, beyond suspicion, probable innocence, etc*, have been used in Crowds [18] and in other schemes. While this helps clarify, comparison across systems is difficult. In our analysis, we use entropy to measure the amount of information the attacker is missing for him to link node identity with location and movement. Our method is predicated on: "Anonymity of a system may be defined as the amount of information the attacker is missing to uniquely identify an actor's link to an action" [22, 2]. In information theoretic terms, the anonymity of the system $\mathcal{A}$, is the entropy $\mathcal{E}$, of the probability distribution over all the actors α_i, that they committed a specific action. Hence,

$$\mathcal{A} = \mathcal{E}[\alpha_i] = -\sum_i Pr[\alpha_i] log_2 Pr[\alpha_i]$$

This expresses in bits the uncertainty experienced by the attacker. For a number of nodes d_{t_k} that the attacker can track from t_0 to t_k, there are $N - d_{t_k}$ nodes whose whereabouts the attacker is uncertain about, and assumed non-trackable. The entropy of the privacy-enhancing system is:

$$\mathcal{A} = -\sum \frac{1}{N - d_{t_k}} log_2 (\frac{1}{N - d_{t_k}})$$

Thus, if a linking algorithm is very sensitive to blackouts, then as $B \to 0$, $d_{t_k} \to 0$. This is the case of maximal anonymity where the entropy of the privacy-enhancing system is $\mathcal{A} = -log_2(\frac{1}{N})$, assuming uniform probability.

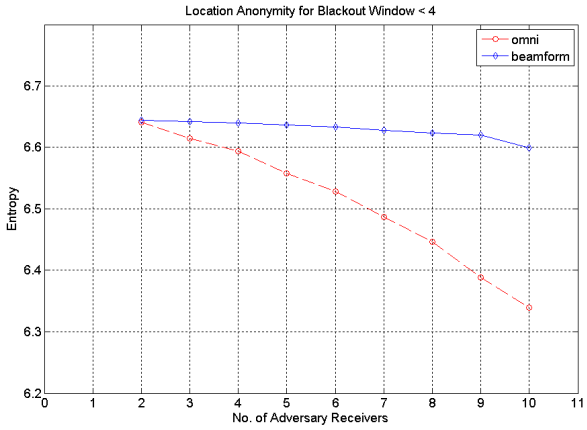


Fig 12. Information-Theoretic Location Privacy of System

An attack using a robust linking algorithm ($B = 4$) is shown in Fig. 12. According to the metric, ABF outperforms OMNI, providing better information-theoretic location privacy. The different slope of the curves indicates that attacker investment in more receivers yields less steeply increasing benefits for the former than the latter.

10 Future Work

Other candidate wireless systems and beam patterns can be evaluated under our flexible framework. Spread spectrum schemes, such as frequency hopping, which extend

covertness into the time and frequency domains (requiring secret keying between the parties), may also be evaluated as another privacy-enhancing layer, on top of beamforming, as a plug-in to our framework. When we are able to consider more specific tracking algorithms, our procedure for measuring location privacy and the metric could be refined.

11 Conclusions

We have shown how an low-level passive adversary can locate victim nodes using AOA despite the presence of higher level countermeasures. We have analyzed the requirements of a location-privacy-enhancing system, and proposed an appropriate architecture, in which mobile nodes use multiple element antennas to adaptively beamform. We composed an end-to-end evaluation framework combining radio and mobility modelling, and used it to show that our solution substantially improves location privacy. As is well known, there are no definitive solutions in the escalating interplay between attack and defence; but our method substantially increases the cost to the attacker, who is forced to deploy more resources. We made extensions to mobile location privacy and a metric to measure it, providing the concept of 'blackout window' to describe robustness of linking algorithms. Simulating direction-finding on a WLAN-type system, we showed that there is a security margin for mere communications versus direction-finding. Other wireless technologies would be equally amenable to analysis under our flexible and extensible framework.

Acknowledgments

We are grateful to our colleagues in the Digital Technology Group and in the Security Group for providing a stimulating and supportive environment in which to conduct this research. We also thank Jon Crowcroft and the anonymous reviewers for their valuable comments.

References

[1] C. Balanis. *Antenna Theory Analysis and Design*. John Wiley and Sons Inc., 1997.

[2] A. R. Beresford and F. Stajano. Location privacy in pervasive computing. *IEEE Pervasive Comp.*, 3(1):46–55, 2003.

[3] A. R. Beresford and F. Stajano. Mix zones: User privacy in location-aware services. *IEEE PerSec*, 2004.

[4] A. Brown and N. Gerein. Test results of a digital beamforming gps receiver in a jamming environment. In *Proceedings of ION GPS 2001*, Salt Lake City, Utah, September 2001.

[5] J. J. Caffery and G. L. Stuber. Overview of radiolocation cdma cellular systems. *IEEE Comms. Mag.*, April 1998.

[6] D. Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM*, 24(2):84–88, Feb 1981.

[7] COST. *COST Action 231: Digital mobile radio towards future generation systems: Final Report*, 1999.

[8] FCC. Fcc part 15.247 operation within the bands 902-928 mhz, 2400-2483.5 mhz, and 5725-5850 mhz, 2002.

[9] L. C. Godara. Application of antenna array to mobile communications, part ii: Beam-forming and directional-of-arrival consideration. *Proc. of the IEEE*, 85(8), Aug 1997.

[10] M. Gruteser and D. Grunwald. Anonymous usage of location-based services through spatial and temporal cloaking. *MobiSys*, 2003.

[11] M. Gruteser and D. Grunwald. Enhancing location privacy in wireless lan through disposable interface identifiers: A quantitative analysis. *1st WMASH*, 2003.

[12] M. Gruteser and B. Hoh. On the anonymity of periodic location samples. In *Proc. of SPC 2005*, Germany, 2005.

[13] L. Huang and B. Hoh. Enhancing wireless location privacy using silent period. In *5th Workshop on Privacy Enhancing Technologies (PET)*, Dubrovnik, Croatia, 2005.

[14] IEEE. Supplements to ieee standard for information technology- telecommunications and information exchange between systems-local and metropolitan area networks-specific requirements-part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications: Higher-speed physical layer extension in the 2.4ghz band. *IEEE Std 802.11b-1999*, 1999.

[15] J. Kaufman, J. Myllymaki, and J. Jackson. Ibm city simulator spatial data generator 2.0, 2001.

[16] M. Lin and I. Wassell. Impact of channel sounder frequency offset on the estimation of channel parameters. In *IEEE Vehicular Technology Conference 2006 Fall*, September 2006.

[17] A. Pfitzmann and M. Kohntopp. Anonymity, unobservability and pseudonymity – a proposal for terminology. In H. Federath, editor, *Designing Privacy Enhancing Technologies, Proc. Int'l Workshop Design Issues in Anonymity and Observability, LNCS 2009*, pages 1–9. Springer-Verlag, 2001.

[18] M. K. Reiter and A. D. Rubin. Crowds: anonymity for web transactions. *ACM Transactions Information Systems Security*, 1(1):66–92, 1998.

[19] R. Roy and T. Kailath. Esprit - estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37(7):984–995, July 1989.

[20] R. Roy, A. Paulraj, and T. Kailath. Esprit - a subspace rotation approach to estimation of parameters of cissoids in noise. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 34:1340–1342, 1986 1986.

[21] R. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation*, 34:276–290, March 1986.

[22] A. Serjantov and G. Danezis. Towards an information theoretic metric for anonymity. In *Proceedings of Workshop on Privacy Enhancing Technologies*, 2002.

[23] C. E. Shannon. A mathematical theory of communication. *Bell System Tech. J.*, 27:379–423, July 1948.

[24] F.-L. Wong and F. Stajano. Location privacy in bluetooth. In R. Molva, G. Tsudik, and D. Westhoff, editors, *Proceedings of the 2nd ESAS, LNCS 3813*, pages 176–188, Visegrad, Hungary, July 2005. Springer-Verlag.