# SENSOR SELECTION FOR MITIGATION OF RSS-BASED ATTACKS IN WIRELESS LOCAL AREA NETWORK POSITIONING

*Azadeh Kushki, Konstantinos N. Plataniotis, Anastasios N. Venetsanopoulos*

E-mail: {azadeh, kostas, anv}@comm.toronto.edu
The Edward S. Rogers Sr. Dept. of Electrical & Computer Engineering, University of Toronto

## ABSTRACT

Positioning in wireless networks has gained significant ground as an enabling technology for various applications such as event detection and context awareness. Since these positioning systems rely on radio features to locate a mobile, they are susceptible to non-cryptographic attacks resulting from malicious alteration of the propagation environment. This paper proposes a sensor selection scheme for increasing the resilience of fingerprinting-based positioning systems to RSS-based attacks in the context of Wireless Local Area Networks (WLAN). A distributed positioning scheme is proposed whereby an estimate is obtained from each WLAN access point (AP). Sensor selection is performed based on a nonparametric estimate of the Fisher Information. Experimental results indicate superior performance compared to existing methods and graceful performance degradation in presence of RSS attacks.

***Index Terms*—** Position measurement, radio position measurement, security, multisensor systems, distributed estimation.

## 1. INTRODUCTION

The mobility of users of wireless networks has motivated the development of a wide range of services offered as added value in existing communication infrastructures. One such service is *positioning*, or the automatic determination of a mobile's location, for use in higher level functions. These functions include event detection, network related services (*e.g.*, routing and security), as well as context awareness and location-based services (LBS).

Positioning using wireless sensors relies on the location dependency of radio features such as time of arrival (ToA), time difference of arrival (TDoA), angle of arrival (AoA), and received signal strength (RSS). Characterization of the relationship between physical positions and a given radio feature allows a positioning system to locate a mobile device through observation of radio signals received or transmitted by this device. The position-radio feature relationship is highly dependent on the propagation environment as shadowing, multipath, and non-line-of-sight propagation effects have a severe impact on radio features. For this reason, position systems relying on wireless radio features are susceptible not only to cryptographic attacks, such as impersonation, but also to non-cryptographic attacks launched by deliberate alterations of the propagation environment. Since the positioning estimates are often inputs to other systems, the sensitivity of positioning methods to malicious attacks is of great concern [1]. This paper considers the problem of sensor selection for improving the resilience of positioning systems to non-cryptographic attacks. In particular, we propose a distributed estimation and sensor selection scheme to address RSS-based attacks in the context of an indoor Wireless Local Area Network (WLAN) positioning system.

## 2. SYSTEM MODEL & RELATED WORK

WLANs are widely deployed in commercial and residential indoor environments and RSS readings can be readily obtained from Network Interface Cards (NIC) available on most mobile devices. These two factors make RSS-based WLAN positioning an attractive solution in many applications.

Fig.1 depicts a typical WLAN positioning setup containing $L$ WLAN access points (APs). Since these APs may belong to different networks, their exact coordinates are generally unknown to the positioning system, rendering ranging-based positioning techniques inappropriate.

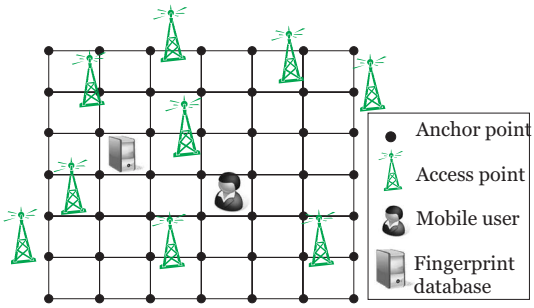



**Fig. 1**. The problem setup.

Positioning is carried out by utilizing the location dependency of RSS. In indoor environments, severe multipath and shadowing effects as well as non-line-of-sight propagation give rise to a time-varying and complex RSS-position dependence. For this reason, WLAN positioning systems characterize the RSS-position relationship implicity, through the use of a method known as *fingerprinting* or *scene matching*. In such an approach, training RSS measurements are collected at a set of $N$ anchor points with known coordinates. During the online operation of the system, the incoming readings from the mobile are matched against these fingerprints using various method including the Euclidean distance [2] and statistical and pattern recognition techniques [3]. Finally, a combination of the coordinates of the anchor points whose fingerprint records best match the observation are returned to the user.

As shown in Fig.1, fingerprints are stored on a central server and positioning is performed on the mobile to preserve privacy.

### 2.1. Attack Scenarios

Fingerprinting-based positioning systems are vulnerable to malicious attacks during both offline and online stages. During the offline fingerprint collection phase, an adversary can impersonate APs (Sybil

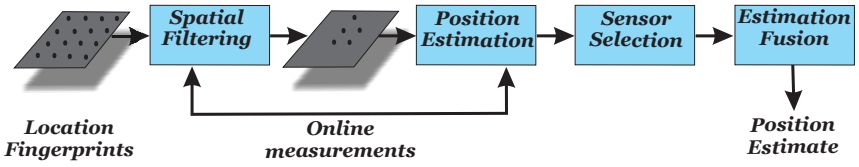



**Fig. 2**. System overview.

attack), or corrupt the training RSS data by jamming, partial attenuation or amplification, or by compromising APs. Impersonation can be remedied by authenticating of beacon nodes [4] and corruptions to training RSS values can be handled through collection of data over various days. Moreover, as the training is performed offline, validation and attack detection schemes [1] can be applied to the data before use in the system.

Assuming the secure transmission of the fingerprint data to the mobile client, this paper focuses on RSS-based attacks carried out on the observed radio signal during the online operation of the system. Such attacks again include impersonation, jamming, and modifying signal strengths. Moreover, we limit the scope of this paper to non-cryptographic RSS-based attacks where an adversary alters the RSS readings through attenuation or amplification [5].

Most prior works addressing RSS-based attacks have focused on range-based techniques where a position estimate is obtained from distances to at least three landmarks with known locations. The study of RSS-based attacks in the context of fingerprinting techniques, however, has been very limited. In [1] a method for detection of RSS-based attacks is proposed. The work of [5] proposes increasing redundancy in the system by increasing the number of APs. The traditional Euclidean distance is replaced with a median-based distance measure for comparing observation vectors to the fingerprints to reduce the effect of outlier APs. The work of [6] uses a set of static and mobile hidden base-stations for secure positioning.

We propose to improve the resilience of the fingerprinting-based methods through sensor selection. In particular, the proposed method selects a set of *reliable* APs from the set of available APs to perform positioning. Such a technique is especially effective in the WLAN context as the number of available APs is generally much larger than the minimum three needed for positioning.

Existing WLAN positioning methods choose three or more of the APs based on *a priori* knowledge. Selection methodologies include choosing a subset of APs with the strongest observation RSS to decrease the probability of outage [2], minimization of correlation between APs using divergence measures [3], and selection of the most discriminant APs using the the entropy-based Information Gain criterion [7]. Since such methods do not provide realtime quality assessment, they are not suitable for use in detection and mitigation of APs under attack. In contrast, this paper proposes the realtime selection of APs by using the *Fisher information* for assessing the quality of the information provided by each AP.

## 3. PROPOSED METHOD

Traditionally, RSS readings at a given location $i$ and time $t$ are treated as a vector $\mathbf{r}_i(t) = [r_i^1(t), \ldots, r_i^L(t)]'$ where $L$ corresponds to the number of available APs in the environment. In this paper, we propose a *distributed* scheme whereby RSS reading from each AP are treated individually. The individual estimates are then fused to provide the final position estimate. Such a distributed (estimation fusion) approach is advantageous over its centralized (feature fusion) counterpart as the quality of each AP estimate can be considered during the fusion stage.

An overview of the proposed method is depicted in Fig.2. As previously mentioned, in the proposed distributed approach, the position estimates are formed based on single APs. Positioning in two-dimensional space, however, requires at least three APs. In order to resolve ambiguities arising from the use of a single AP, a *spatial filtering* step [3] is first performed to filter out anchor points far from the current position of the mobile and localize the spatial search area as described in Section 3.1. Next, the minimum mean squared error (MMSE) estimate of the mobile position is formed for each AP using a nonparametric estimate of the conditional probability density function as detailed in Section 3.2. Finally, a set of reliable APs is selected based on the Fisher information associated with each estimate and the results are fused to provide the final position estimate. Sensor selection and fusion methods are discussed in Section 3.3.

### 3.1. Spatial Filtering

Spatial filtering aims to resolve ambiguities that arise due to the use of a single AP and is performed based on the observation that points that are close in the physical space, receive coverage from similar sets of APs [3]. Denote a coverage vector as $\mathbf{I_p} = [I_\mathbf{p}^1 \ldots I_\mathbf{p}^L]$ where $I_\mathbf{p}^i = 1$ if AP $i$ provides coverage at $\mathbf{p}$. The set of anchor points retained by the filtering operation is denoted as $\{\mathbf{p}_{(1)}, \ldots, \mathbf{p}_{(N')}\}$ where $\mathbf{p}_{(1)}, \ldots, \mathbf{p}_{(N)}$ denotes an ordering of the anchor points based on the Hamming distance between each anchor point's coverage vector and that of the observation. The parameter $N'$ indicates the number of anchor points retained after spatial filtering. Section 4 reports on the effect of this parameter on positioning results.

### 3.2. Position Estimation

The fingerprint matrix for an anchor point $\mathbf{p}_i = (x_i, y_i)$ is defined as $\mathbf{F}(\mathbf{p}_i) = [\mathbf{r}_i(1); \ldots; \mathbf{r}_i(n)]$ where $\mathbf{r}_i(t) = [r_i^1(t), \ldots, r_i^L(t)]'$ is the vector of RSS readings from $L$ APs at time $t$ when the mobile resides at $\mathbf{p}_i$ [3], and $n$ is the number of time samples collected at $\mathbf{p}_i$. These samples are averaged to generate a single representative training RSS value per anchor point. Thus, the set of location fingerprints after spatial filtering is $\{(\mathbf{p}_1, \overline{r}_1^a), \ldots, (\mathbf{p}'_N, \overline{r}_{N'}^a)\}$ where $\overline{r}_i^a = \frac{1}{n}\sum_{t=1}^{n} r_i^a(t)$ is the mean RSS from the $a^{th}$ AP at $\mathbf{p}_i$.

Given an observation $r^a$ from AP $a$, the MMSE estimate of position $\mathbf{p}$ is the posterior mean $E(\mathbf{p}|r_a)$. This estimate is obtained from the nonparametric estimate of the posterior density $f(\mathbf{p}|\mathbf{r}_a)$ [3]:

$$f(\mathbf{p}|\mathbf{r}_a) = \sum_{i=1}^{N'} \frac{K_{\mathbf{H}_r}(r^a - \overline{r}_i^a)}{\sum_{i=1}^{N'} K_{\mathbf{H}_r}(r^a - \overline{r}_i^a)} K_{\mathbf{H}_\mathbf{p}}(\mathbf{p} - \mathbf{p}_i) \quad (1)$$

where $K_\mathbf{H}(\cdot)$ is the $d-$dimensional Gaussian kernel defined as

$$K_\mathbf{H}(x) = \frac{1}{(2\pi)^{d/2}|\mathbf{H}|^{d/2}} \exp\left(-\frac{1}{2} x \mathbf{H}^{-1} x'\right). \quad (2)$$

The $d \times d$ diagonal matrix $\mathbf{H}$ defines the width of the kernel function in each direction. The diagonal elements $h_1, \ldots, h_d$ are estimated to minimize the Asymptotic Mean Integrated Square Error between the true and estimated densities. These parameters are obtained as $h_i = 1.06\hat{\sigma}_i N^{1/5}$ with $\hat{\sigma}_i$ is the sample variance of the training data in the $i^{th}$ dimension. Using the Gaussian kernel in (1), the resulting density can be interpreted as a Gaussian mixture:

$$f(\mathbf{p}|\mathbf{r}_a) = \sum_{i=1}^{N'} w_i \exp\left(-\frac{1}{2}(\mathbf{p}-\mathbf{p}_i)\mathbf{H}_\mathbf{p}^{-1}(\mathbf{p}-\mathbf{p}_i)\right), \quad (3)$$

where

$$w_i = \frac{K_{\mathbf{H}_r}\left(r_a - \overline{r}_a^i\right)}{\sum_{i=1}^{N'} K_{\mathbf{H}_r}\left(r_a - \overline{r}_a^i\right)}. \quad (4)$$

The MMSE estimate of the position (*i.e.*, the conditional expectation) and its covariance can be obtained from (3) by noting that the Gaussian mixture can be approximated with a single Gaussian with the mean and covariance shown below [8]

$$\hat{\mathbf{p}}_a = E\{\mathbf{p}|r_a\} \approx \sum_{i=1}^{N'} w_i \mathbf{p}_i \quad (5)$$

$$C_a = cov\{\mathbf{p}_a|r\} \approx \sum_{i=1}^{N'} w_i \left(\mathbf{H}_\mathbf{p} + (\mathbf{p}_i - \hat{\mathbf{p}}_a)(\mathbf{p}_i - \hat{\mathbf{p}}_a)'\right) \quad (6)$$

### 3.3. Sensor Selection & Fusion

Assuming that each estimate obtained from (5) is a corrupted version of the true mobile position, the following observation model is used for each AP

$$\hat{\mathbf{p}}_a = \mathbf{p} + \epsilon_a, \qquad a = 1, \ldots, A. \quad (7)$$

where $\epsilon_a = \mathbf{p} - \hat{\mathbf{p}}_a$ is the estimation error associated with the $a^{th}$ AP. These corrupted measurements must now be fused to obtain the final position estimate $\hat{\mathbf{p}}$. To this end, the mean and covariance of the estimation error from each AP are considered:

$$E\{\epsilon_a\} = E\{\mathbf{p} - \hat{\mathbf{p}}_a\} \quad (8)$$

$$cov\{\epsilon_a\} = C_a + \underbrace{E\{\mathbf{p} - \hat{\mathbf{p}}_a\}E\{\mathbf{p} - \hat{\mathbf{p}}_a\}'}_{\text{bias term}} \quad (9)$$

The estimation covariance $C_a$ provides a measure of confidence in a given estimate $\hat{\mathbf{p}}_a$: the smaller the covariance, the higher the probability that the estimate is close to the estimation mean. The estimator of (5), however, is biased and $E\{\epsilon_a\} \neq 0$ [3]. Unfortunately, the estimation bias is unknown in (8) and $cov\{\epsilon_a\}$ cannot be calculated directly. For this reason, the minimum variance unbiased estimate (MVUE) [9], cannot be applied readily. Instead, we apply a two step procedure: we first select a set of APs with small estimation covariance and then fuse this subset.

Sensor selection is performed based on the Fisher Information obtained from each AP using the estimation covariance only. Define the score for the $a^{th}$ AP as $J_a = tr(C_a^{-1})$ and let $\hat{\mathbf{p}}_{(1)}, \ldots, \hat{\mathbf{p}}_{(L)}$ denote the ordering of the APs based $J_a$. Given this ranking, a predetermined number of APs with the highest confidence scores are selected. In this paper, the selection is performed adaptively by choosing the APs whose score is at least 50% of the highest confidence score. The set of selected APs is then

$$\mathcal{A} = \{a | J_a \leq \frac{1}{2}\max_i J_i\} \quad (10)$$

The next step is to fuse the estimates from the selected APs to obtain the position estimate. The use of the sample mean for fusion of the individual estimates is not appropriate as the individual estimates may be biased. Instead, we use the *median* as the fusion operator to provide robustness to outlier APs. Mathematically, $\hat{\mathbf{p}} = median_{a\in\mathcal{A}}\hat{\mathbf{p}}_a$.

## 4. EXPERIMENTS & RESULTS

In this section, we evaluate the resilience of the proposed method to RSS amplification and attenuation attacks and report experimental results based on RSS values collected in a real environment.

### 4.1. Attack Model

In prior literature [10], a linear attack model is considered whereby attacked RSS values are simulated by perturbing the original measurements by a deterministic constant across all APs. In this paper, a similar attack model is used with the exception that RSS values are corrupted by additive Gaussian noise instead of a deterministic constant. This approach is motivated by the path loss model relating the received signal power to the transmitted power [11]

$$P_r(dBm) = P_t(dBm) - 10\log_{10}K - 10\gamma\log_{10}\left(\frac{d}{d_0}\right) - \psi(dBm) \quad (11)$$

In (11), $P_r$ and $P_t$ are the received and transmitted powers respectively, $K$ is a constant relating to antenna and average channel attenuation, $d_0$ is a reference distance for antenna far-field, and $\gamma$ is the path loss exponent. Finally, $\psi \sim \mathcal{N}(0, \sigma_\psi^2)$ reflects variations caused by random attenuation (log-normal shadowing) in the environment. While the path loss model is not always accurate for estimating distances in indoors, it provides insights into possible attack models. RSS attacks may be launched through alteration of the transmitted power or the propagation environment [10]. To simulate an RSS attack caused by the alteration of the environment, we perturb the received power by adding Gaussian noise with variance $\sigma_{noise}$, essentially modifying the random component of the model due to environmental attenuation.

### 4.2. Experimental Setup

Experimental results are reported for the dataset used in [3]. In this dataset, a total of 33 APs are detectable throughout the floor, with an average of 9.6 APs covering each anchor point. The measurements are collected using a Toshiba Satellite laptop with a Pentium M processor, an Intel PRO/Wireless 2915ABG Network Adapter, and Windows XP operating system. RSS measurements are obtained by a publicly available network sniffer software, NetStumbler[1].

Fig.3 shows the experimentation area layout and depicts the anchor points as black circles. For each anchor point 100 samples are collected at a rate of 1 sample/second. The laptop orientation is indicated by arrows in Fig.3. Test measurements are collected over two days, different than the training days, for 44 test points situated on and off the training points with four different orientations of the laptop. A total of 60 samples were collected per test point per orientation at a rate of 1 sample/second.

[1] http://www.netstumbler.org

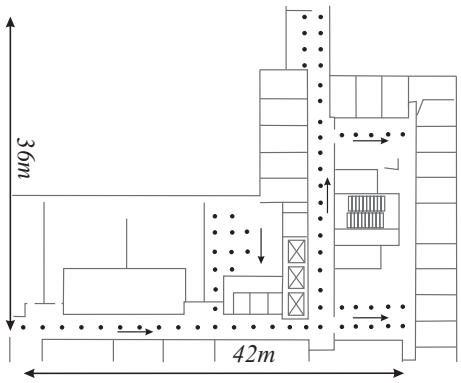



**Fig. 3**. Map of the experimentation environment.

The positioning error is reported numerically as the Root Mean Square Error (RMSE) averaged over the test points.

### 4.3. Results

Fig.4 shows the average RMSE as a function of the number of anchor points retained during the spatial filtering step. The results indicate that spatial filtering is effective in counteracting the negative effects of positioning with a single AP by excluding non-relevant anchor points during positioning. Due to the time variance of the environment, and consequently RSS values, for $N' < 0.2N$ all survey points are excluded and no estimate is available.

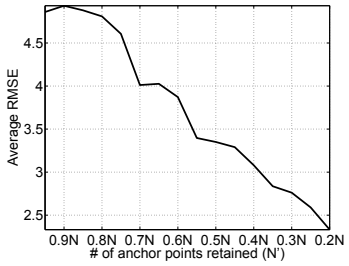
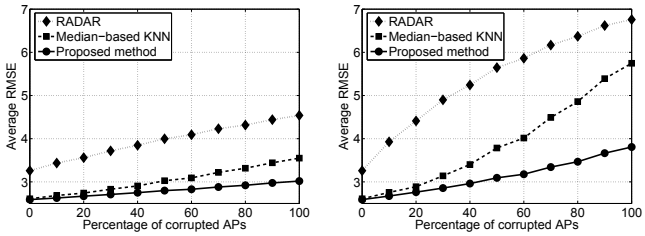



**Fig. 4**. Effect of spatial filtering on positioning accuracy. $N$ is the total number of anchor points.

Next, the resilience of the proposed method to RSS-based attack is compared to that of the K-nearest neighbour method of RADAR [2], and the median distance-based method of [5]. Fig.5 shows the average RMSE as a function of the percentage of APs corrupted for spatial filtering with $N' = 0.25N$. The Euclidean distance based KNN method performs the poorest as no provision for attack mitigation is taken. The median-based technique of [5] requires that at least half of the APs report true RSS values. As seen from Fig.5 this method performs poorly once the percentage of corrupted APs exceeds 50%, especially as the noise variance is increased. For the proposed method, the average RMSE remains relatively unchanged for the case of $\sigma_{noise} = 5$ even when all APs are corrupted. For higher noise variance, the method exhibits a graceful degradation in performance as compared with the other techniques. These results indicate that the nonparametric estimate of the Fisher information provides an effective means for assessing the quality of each AP.




(a) $\sigma_{noise}$=5dBm. (b) $\sigma_{noise}$=10dBm.

**Fig. 5**. Average RMSE for different percentage of corrupted APs.

## 5. CONCLUSION

In this paper, we proposed a distributed scheme for RSS-based WLAN positioning. Resilience to RSS-based attacks was achieved through the selection of APs used in positioning based on a nonparametric estimate of the Fisher Information matrix. Experimental result indicate graceful performance degradation even when RSS measurements from all APs are corrupted. An interesting direction for future work is the incorporation of position predications in sensor selection. An anticipatory design can further enhance the selection of both the anchor points and APs involved in positioning.

## 6. REFERENCES

[1] Y. Chen, W. Trappe, and R. P. Martin, "ADLS: Attack detection for wireless localization using least squares," in *Proc. of PerComW*, 2007, pp. 610–613.

[2] P. Bahl and V. Padmanabhan, "RADAR: an in-building RF-based user location and tracking system," in *Proc. of IEEE InfoCom*, vol. 2, 2000, pp. 775–784.

[3] A. Kushki, K. Plataniotis, and A. Venetsanopoulos, "Kernel-based positioning in wireless local area networks," *IEEE Trans. on Mobile Computing*, vol. 6, no. 6, pp. 689–705, 2007.

[4] D. Liu, P. Ning, and W. Du, "Attack-resistant location estimation in sensor networks," in *Proc. of IPSN*, 2005, pp. 99–106.

[5] Z. Li, W. Trappe, Y. Zhang, and B. Nath, "Robust statistical methods for securing wireless localization in sensor networks," in *Proc. of IPSN*, 2005, pp. 91–98.

[6] S. Capkun, M. Cagalj, and M. Srivastava, "Secure localization with hidden and mobile base stations," in *Proc. of InfoCom*, 2006, pp. 1–10.

[7] Y. Chen, Q. Yang, J. Yin, and X. Chai, "Power-efficient access-point selection for indoor location estimation," *IEEE Trans. on Knowledge and Data Engineering*, vol. 18, no. 7, pp. 877–888, 2006.

[8] Y. Bar-Shalom, X.-R. Li, and T. Kirubarajan, *Estimation with applications to tracking and navigation*. New York : Wiley, 2001.

[9] Z. Quan and A. H. Sayed, "Innovations-based sampling over spatially-correlated sensors," in *Proc. of ICASSP*, vol. 3, 2007, pp. 509–512.

[10] Y. Chen, W. Trappe, and R. Martin, "Attack detection in wireless localization," in *Proc. InfoCom*, 2007, pp. 1964–1972.

[11] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.