Analysis of a Device-free Passive Tracking System in Typical Wireless Environments

Ahmed E. Kosba, Ahmed Abdelkader
Department of Computer Engineering
Faculty of Engineering, Alexandria University, Egypt
Email: {a.kosba, a.abdelkader}@student.alx.edu.eg

Moustafa Youssef*
Wireless Intelligent Networks Center (WINC)
Nile University, Egypt
Email: mayoussef@nileu.edu.eg

***Abstract*—Device-free Passive (DfP) localization is a new concept in location determination where the tracked entity does not carry any device nor participate actively in the localization process. A *DfP* system operates by processing the received physical signal of a wireless transmitter at one or more monitoring points. The previously introduced *DfP* system was shown to enable the tracking of a single intruder with high accuracy in a *highly controlled* WLAN environment. In this paper, we propose and analyze different algorithms for *DfP* tracking in a *typical* indoor WLAN environment, *rich in multipath*. We also study the effect of the temporal and spatial changes in the environment on the accuracy of the system. In addition, we evaluate the effect of the different configurations of the wireless equipment placement on the *DfP* localization accuracy. Our results show that our proposed techniques can accurately track the user in typical environments, thus enabling a large number of *DfP* applications.**

***Index Terms*—Deterministic localization algorithms, device-free localization, probabilistic localization algorithms.**

I. Introduction

The rapid advancements in communication networks and the introduction of location based services (LBS) have motivated significant research effort in the area of location determination. This effort resulted in the development of many location determination systems, including the GPS system [1], ultrasonic-based systems [2], infrared-based (IR) systems [3], and radio frequency-based (RF) [4] systems. All these systems require a device to be attached to the tracked entity, which may be responsible to run part of the localization algorithm. Therefore, we refer to these systems as device-based active localization systems.

Recently, we introduced device-free passive *DfP* localization for standard wireless networks [5], [6]. *DfP* localization enables the detection and tracking of entities without requiring any attached devices or active participation on their part. In addition, it works with already installed wireless communication networks without requiring any specialized hardware, which increases the value of the underlying network by extending its services to support the localization function besides and without affecting the original communication function.

DfP systems are attractive for many practical applications including: intrusion detection and tracking, sensor-less sensing, low cost surveillance, and home automation.

* Also Affiliated with Alexandria University, Egypt

A *DfP* system consists of signal transmitters, such as access points (APs), signal receivers or *monitoring points* (MPs), such as standard laptops, and an application server (AS) which collects information about the received signals from each MP. The AS uses this information to perform the localization functions and initiate actions as needed.

The *DfP* concept utilizes the fact that RF characteristics are functions of the surrounding environment. Changes in the received physical signals can be related to changes in the environment. Among these changes, we are interested in the movement of entities inside an area of interest. Humans in particular, considerably affect RF waves in the 2.4 GHz band [7], e.g. IEEE 802.11b/g, as it contains the resonance frequency of water which comprises a significant fraction of the human body. Our first *DfP* system [5] revealed promising capabilities of very high probability of detection and accurate tracking of a single intruder in a highly controlled WLAN environment using simple algorithms.

In this paper, we propose and analyze the performance of four different *DfP* tracking algorithms in a typical wireless environment, rich in multipath. Our algorithms fall into two categories: deterministic and probabilistic. For the probabilistic case, both parametric and non-parametric distributions are used to estimate the complex relation between signal characteristics and the user's location. We evaluate the proposed algorithms in a typical environment and study the effect of the temporal and spatial changes in the environment on the accuracy of the system. We also compare different *configurations* of APs and MPs placement and suggest a simple approach for selecting a configuration that enhances the system performance.

The rest of this paper is organized as follows: Section II reviews related work in the area of location determination. Section III presents the proposed algorithms. Section IV analyzes the system's tracking performance. Section V studies the effect of changing the configuration on the tracking performance. Finally, Section VI summarizes the paper and sheds some light on future work.

II. Related Work

In WLAN device-based active localization, e.g. [4], [8] the system usually works in two phases: an offline training phase and an online operational phase. In the offline phase, the system learns about the relation between RF characteristics and the possible user's locations. Then, during the online phase, the system infers the user's location from the measurements of the received signals collected by the client device. The Received Signal Strength Indicator (RSSI) is the most commonly used signal characteristic as it is much easier to estimate by client devices than Time of Arrival (ToA) or Angle of Arrival (AoA), which require specialized hardware.

In a typical indoor environment, fading effects due to multipath propagation and shadowing from obstacles present a hostile propagation medium for RF waves [7]. As a result, most WLAN based systems use a radio map to capture the complex relation between RSSI readings and the user's location. During the offline phase, the system constructs an active radio map that stores information about the RSSI from each AP at a selected set of locations. This radio map is searched in the online phase to find the location that best matches the physical signals received by the client device.

There are two basic classes of techniques to perform this matching operation: deterministic techniques and probabilistic techniques. Deterministic techniques represent the RSSI by a scalar value, most commonly the mean value, and use distance based methods to find the closest match. The *RADAR* [8] system, for example, uses a nearest neighbor search to estimate the user's location. On the other hand, probabilistic techniques use the training information to estimate RSSI distributions at each location and use probabilistic methods to find the match with the maximum likelihood. The *Horus* [4] system, for example, uses a Bayesian inversion based search to infer the user's location.

The fundamental difference between a *DfP* system and device-based active localization systems is that the user carries no devices and all measurements are collected at monitoring points and forwarded to a centralized application server for processing. Therefore, the passive radio map is constructed during the offline phase by recording the effect of a person on the RSSI measurements for each *stream*, i.e. AP and MP pair, when a user is located at the selected set of locations that covers the area of interest. Similarly, in the online phase, each MP periodically sends the RSSI received from each AP to the AS where the matching operation is performed.

Ultra-wide band radar [9], computer vision [10] and physical contact [11] based systems can be classified as device-free passive localization systems. However, they require specialized hardware with possible high installment costs, which makes them less appealing in practice.

III. Device-free Passive Localization

In this section, we present different algorithms for *DfP* localization. The algorithms work in two phases: an offline training phase and an online tracking phase.

A. Offline Phase: Passive Radio Map Construction

Given an area of interest that has N calibration locations and a set of n APs and m MPs ($m \times n$ streams), a person stands at each of the N locations for a period of time and RSSI values from each stream are processed and stored in a radio map. This is done only once during an offline phase. What the radio map stores depends on the localization algorithm being used as explained in the next subsection.

B. Online Phase: Tracking Algorithms

During the online phase, the system is required to determine the location of a passive user. Let k, the number of streams, equal $m \times n$. The AS periodically receives form the MPs a k-dimensional signal strength vector $\bar{s}$, where each entry $\bar{s}_{ij}$ denotes the signal strength received at MP i from AP j. The goal is to determine the radio map location that is closest to the passive user's location. We propose four algorithms for estimating the user's location that fall into two categories: deterministic and probabilistic.

1) Deterministic Algorithm: We propose a Nearest Neighbor in Signal Space (NNSS) algorithm. For this deterministic algorithm, at each location in the radio map, the system stores a vector, with each entry representing the mean of one of the k streams when the user is standing at this location. During the online phase, the algorithm selects the radio map location l that minimizes the Euclidean distance between the received signal strength vector $\bar{s}$ and the mean signal strength vector s'_l stored for that location.

$$\arg\ \min_l \sum_{i=1}^{m} \sum_{j=1}^{n} (\bar{s}_{ij} - s'_{l_{ij}})^2 \tag{1}$$

2) Probabilistic Algorithms: We propose three Bayesian inversion based probabilistic algorithms. The three algorithms select the radio map location l that maximizes the probability $P(l/\bar{s})$ expressed as:

$$\arg\ \max_l P(l/\bar{s}) \quad = \quad \arg\ \max_l P(\bar{s}/l).\frac{P(l)}{P(\bar{s})} \tag{2}$$

$P(\bar{s})$ can be factored out from Equation 2 as it does not depend on any location l. In addition, assuming a uniform user profile[1], i.e. all locations are equiprobable, Equation 2 reduces to:

$$\begin{aligned} \arg\ \max_l P(l/\bar{s}) &= \arg\ \max_l P(\bar{s}/l) \\ &= \arg\ \max_l \prod_{i=1}^{m} \prod_{j=1}^{n} P(\bar{s}_{ij}/l) \end{aligned} \tag{3}$$

We use three different algorithms for estimating the density function $P(\bar{s}_{ij}/l)$:

1) A Parametric Approach (Gaussian): Assumes that the RSSI values for each stream at each location follow a Gaussian distribution. The radio map stores the estimated mean and variance for each of these distributions.
2) A Non-parametric Histogram-Based Approach (Histogram): The radio map stores the RSSI histogram of each stream at each location.

[1] If the user profile is known, it can be used as in Equation 2.

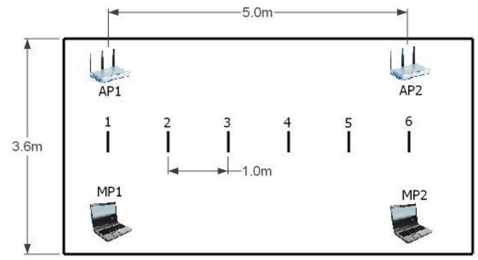


Fig. 1. Experiment 1 Layout.

3) A Non-parametric Kernel-Based Approach (Gaussian Kernel): The radio map stores an estimate for the density function of the RSSI values. This estimate is based on the Gaussian Kernel function described in [12].

In the next section, we study the performance of the system using the above algorithms. We refer to the deterministic NNSS algorithm by *deterministic* while we refer to each probabilistic algorithm by its density estimation method.

IV. Tracking Performance

In this section, we describe the environment in which we conducted our experiments. We also present the metrics we used to evaluate the tracking accuracy of the system. Finally, we analyze the results of our experiments.

A. Experimental Testbed

Our experiments were conducted in a typical IEEE 802.11b home environment with two access points and two monitoring points. We used two Cisco Aironet 1130AG series access points and two HP laptops running Windows XP Professional each equipped with a D-Link AirPlus G+ DWL-650+ Wireless NIC. The experiment considered a single room covered with typical furniture. The room was discretized into six locations spaced one meter apart, representing the radio map locations. We chose to have a small scale experiment, in the same scale as the one used in the controlled experiment [5], for a fair comparison. Figure 1 shows the layout of the first experiment.

To construct the passive radio map, a person stood at each location and each MP recorded 300 samples, taking one sample each second. The person kept standing in the same orientation for all locations. We study the effect of orientation change in Section IV-C2.

B. Evaluation Metrics

We used two metrics to analyze the tracking performance: the average distance error and the probability of error.

1) Average Distance Error: The average distance error between the estimated and true user's locations over all test cases.

2) Probability of Error: The probability that the location estimated by the *DfP* tracking system does not match the true location exactly. This is directly related to the number of times the system identified the user's location correctly.

TABLE I
CROSS-VALIDATION TRACKING RESULTS FOR EXPERIMENT 1 WITH 90% CONFIDENCE.

Algorithm	Average Distance Error	Error Probability
Gaussian Distribution	0.201($\pm$0.0092)m	0.0835($\pm$0.0036)
Histogram	0.176($\pm$0.0091)m	0.0727($\pm$0.0037)
Gaussian Kernel	0.175($\pm$0.0091)m	0.0723($\pm$0.0037)
Deterministic	0.214($\pm$0.0088)m	0.0874($\pm$0.0033)

C. Performance Evaluation

We tested the tracking performance of our system using two techniques: cross-validation and independent test sets. The cross-validation experiment represents the ideal environment, while the independent test sets present the effects of the temporal and spatial changes in the environment on the accuracy of the system.

1) Cross-Validation: Ideal Environment: We performed our analysis using 10-times 10-fold cross-validation as recommended in [13]. The training data set is divided randomly into ten subsets. Then, each individual subset is used as testing data, and the other nine subsets are used for training the system. This process is repeated ten times. Finally, the results are averaged over all tests.

Accuracy: Figure 2 shows the distance error CDF of the four algorithms using cross-validation. Table I summarizes the results. The results show that probabilistic techniques perform better than the deterministic one. As shown in the table, the Gaussian kernel and histogram methods are almost equivalent and are better than the Gaussian approximation and the deterministic algorithms. In addition, the confidence intervals of the Gaussian distribution and the deterministic algorithm are overlapping. A paired-t significance test shows that the Gaussian distribution method is statistically significantly better than the deterministic algorithm. Therefore, we can conclude that probabilistic algorithms outperform the deterministic algorithm. In addition, the non-parametric techniques perform better than the parametric one. This can be explained by noting that the non-parametric estimation attempts to find an estimate for the density function of the signal distribution, whereas the parametric technique assumes that the signal strength follows a specific distribution. This can lead to higher errors, especially if the original signal distribution deviates from the assumed one.

Number of streams: To study the effect of the number of streams on the accuracy of the system, we tested all possible combinations of APs and MPs. Figure 3 shows the average distance error for these combinations. We can see from the figure that as the number of streams is increased (either by the addition of either MPs or APs), the system performs better. In addition, the placement of APs and MPs directly affects the system performance, as we study in more details in Section V.

2) Independent Test Sets: Temporal and Spatial Changes: To better evaluate the accuracy of the system in typical environments, we tested the system using independent test sets that show the effect of the time change and user's orientation. Table

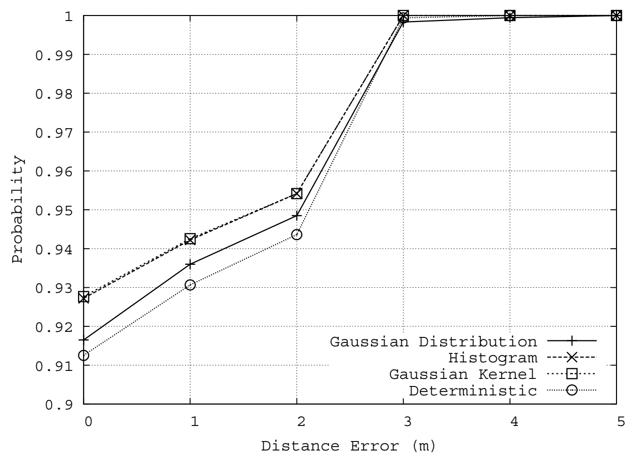


Fig. 2. Distance error CDF for Experiment 1 using cross-validation.

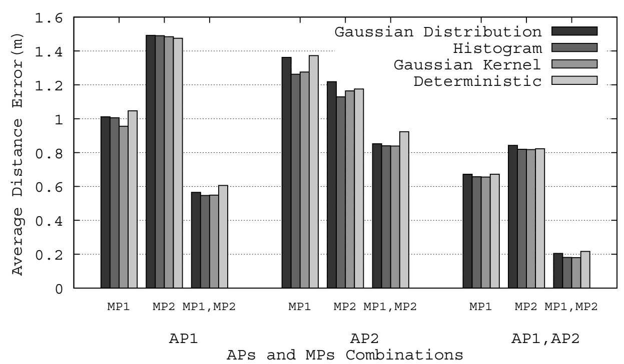


Fig. 3. The average distance error results of Experiment 1 for all APs and MPs combinations.

II shows the performance of the system under two independent test sets using the Gaussian Kernel method compared to the cross-validation results. We also compare the results to that of a random location selection method that returns one of the radio map locations randomly as a baseline for comparison. Each independent data set contains 60 samples per location taken with a rate of one sample per second. The goal of this study is to identify the factors that affect the accuracy of the system in practice.

The results show that, as expected, the cross-validation results overestimate the actual accuracy of the system, since they are extracted from the same training data. Changing the user orientation or the time the system is used reduces the accuracy of the system. This is expected as the environment may have changed significantly. In both cases, the system accuracy is significantly better than that of a random selection method (more than 3.3 times better in the worst case).

Compared to our previous study in a highly controlled environment [5], the performance of the *DfP* system degrades due to the noisy channel. However, the proposed techniques' performance, with a worst case average distance error of 0.586m is still suitable for accurate tracking of intruders.

It should be noted that the above study evaluated the system

TABLE II
RESULTS OF THE GAUSSIAN KERNEL METHOD UNDER INDEPENDENT TEST SETS COLLECTED FOR EXPERIMENT 1

Test set	Time	Orientation	Avg. distance error
Cross-Validation	Same	Same	0.175m
Set 1	Different	Same	0.358m
Set 2	Different	Different	0.586m
Random Selector	N/A	N/A	1.94m

without any special optimization. Extending the system to work operationally requires the addition of some enhancing features to suite practical scenarios. For example, continuous space estimation used in [4] is needed to overcome the limitation of discretized locations. In case of large areas, the system should give higher probabilities for the locations that are closer to the current user location. These optimizations will be necessary when considering large-scale deployments of the system.

V. CONFIGURATION EFFECT

In this section, we study the effect of changing the configuration of APs and MPs placement on the *DfP* system performance. First, we present a hypothesis about the relation between different configurations and the *DfP* tracking performance. Then, we validate the hypothesis through an experimental study.

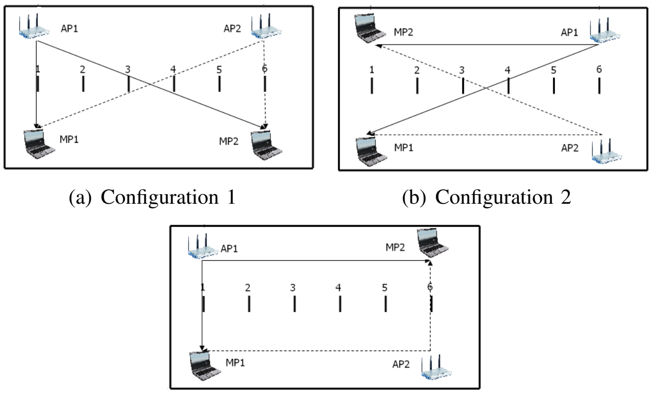
A. Hypothesis

DfP systems depend mainly on the effect of human presence on signal strength. Therefore, in order to improve the accuracy of a *DfP* system, the human effect on the signal strength received at the monitoring points should be maximized. This can be achieved by selecting configurations in which the lines of sight between the APs and MPs are affected directly. In addition, radio map locations should be selected to vary the human effect on the monitoring streams to better discriminate between locations.

B. Experimental Study

In this section, we validate the above hypothesis through experiments. We present a comparison between three different configurations, then we study the location set of the best of these configurations.

1) Different Configurations: In addition to the configuration tested in the previous section, we performed two additional experiments, with two new configurations, to study the configuration change effect. Each experiment had a different placement of two access points and two monitoring points. Figure 4 shows the three configurations we considered. Figure 5 shows the average distance error of the *DfP* system for the three configurations. It can be noted that Configuration 1 has the least error values. As shown in Figure 4(a), four locations in Configuration 1 intersect the lines of sight of the streams, whereas in the other two configurations fewer locations intersect the lines of sight. These results are in line with the hypothesis mentioned above about configuration selection.



(a) Configuration 1 (b) Configuration 2

(c) Configuration 3

Fig. 4. Configurations of the three experiments that were performed to study the effect of changing the APs and MPs configuration. Lines between APs and MPs represent the direct lines of sight for the different streams.

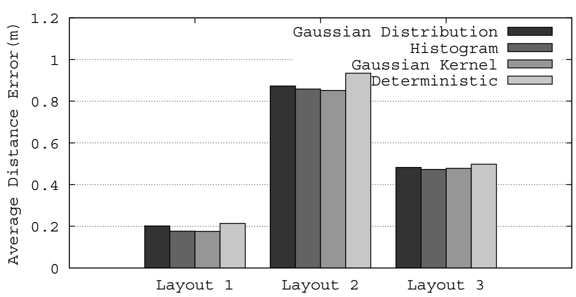


Fig. 5. System performance under the three configurations in Figure 4.

2) Location Set Selection: Figure 6 compares the average distance errors between different location sets for Configuration 1. It is noted that the system performs better when the locations are selected to affect the lines of sight directly (locations $\{1, 3, 4, 6\}$).

From the above results, it can be observed that the choice of the APs and MPs configuration has more effect on accuracy than the localization algorithm itself. This has to be confirmed however for larger areas of interest.

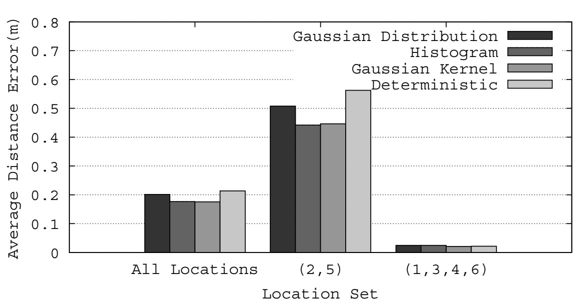


Fig. 6. Average distance errors for different location sets in Configuration 1. Locations $\{1, 3, 4, 6\}$ intersect the lines of sight, while locations $\{2, 5\}$ do not.

VI. Conclusions and Future Work

In this paper, we presented the design and analysis of a *DfP* tracking system in a typical environment, rich in multipath. We presented deterministic and probabilistic *DfP* tracking algorithms and evaluated their performance. The results showed that the probabilistic algorithms outperform the deterministic algorithm. They also showed that the non-parametric probabilistic technique outperformed the parametric one. We studied the effect of the time factor and user's state on the accuracy of the system. We argued that both APs and MPs configuration and radio map location set selection affect the accuracy of the system significantly and showed the effect of the different configurations on the accuracy of a *DfP* system.

For future work, we will work on devising a robust method for selecting configurations that enhances the system performance for a given environment. We will also consider large-scale deployments of *DfP* systems and study how the system can be extended to multi-person tracking and how it can adapt to changes in the monitored environment.

Acknowledgment

This work is supported in part by a grant from the Egyptian Science and Technology Development Fund (STDF).

References

[1] P. Enge and P. Misra, "Special Issue on Global Positioning System," in *Proceedings of the IEEE*, January 1999, pp. 3–172.

[2] N. B. Priyantha, A. Chakraborty, and H. Balakrishnan, "The Cricket Location-Support System," in *MobiCom '00: Proceedings of the 6th annual international conference on Mobile computing and networking*. New York, NY, USA: ACM Press, 2000, pp. 32–43.

[3] R. Want, A. Hopper, V. Falcao, and J. Gibbons, "The Active Badge Location System." *ACM Trans. Inf. Syst.*, vol. 10, no. 1, pp. 91–102, 1992.

[4] M. A. Youssef and A. Agrawala, "The Horus WLAN Location Determination System," in *Communication Networks and Distributed Systems Modeling and Simulation Conference*, 2005, pp. 205–218.

[5] M. Youssef, M. Mah, and A. Agrawala, "Challenges: Device-free Passive Localization for Wireless Environments," in *MobiCom '07: Proceedings of the 13th annual ACM international conference on Mobile computing and networking*. ACM, 2007, pp. 222–229.

[6] M. Moussa and M. Youssef, "Smart Devices for Smart Environments: Device-free Passive Detection in Real Environments," in *The IEEE International Workshop on Intell. Pervasive Devices (PerDev09)*, 2009.

[7] M. Youssef and A. Agrawala, "Small-scale Compensation for WLAN Location Determination Systems," in *WCNC 2003*, vol. 3, 2003, pp. 1974–1978.

[8] P. Bahl and V. N. Padmanabhan, "RADAR: An In-Building RF-based User Location and Tracking System," in *IEEE INFOCOM 2000*, vol. 2. IEEE, March 2000, pp. 775–784.

[9] S. S. Ram, Y. Li, A. Lin, and H. Ling, "Human Tracking Using Doppler Processing and Spatial Beamforming." IEEE 2007 Radar Conference, 2007.

[10] J. Krumm, S. Harris, B. Meyers, B. L. Brumitt, M. Hale, and S. A. Shafer, "Multi-Camera Multi-Person Tracking for Easyliving," in *Proceedings of the Third IEEE International Workshop on Visual Surveillance*, 2000, pp. 3–10.

[11] R. J. Orr and G. D. Abowd, "The Smart Floor: A Mechanism for Natural User Identification and Tracking," in *Proceedings of ACM CHI 2000 Conference on Human Factors in Computing Systems*, ser. Short talks: multimodal interaction, vol. 2, 2000, pp. 275–276.

[12] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. Chapman & Hall/CRC, April 1986.

[13] R. R. Bouckaert and E. Frank, "Evaluating the Replicability of Significance Tests for Comparing Learning Algorithms," in *PAKDD*. Springer, 2004, pp. 3–12.