Entropy-based Location Fingerprinting for WLAN Systems

Nayef Alsindi, Zdenek Chaloupka and James Aweya
Etisalat BT Innovation Center (EBTIC)
Khalifa University of Science, Technology and Research
Abu Dhabi, UAE

***Abstract*— Localization for indoor environments has gained considerable attention over the last decade due to the enormous potential in the technology and the significant challenges facing this area of research. One practical localization technique that relies on the available fixed wireless infrastructure is RF location fingerprinting. Received Signal Strength (RSS)-based location fingerprinting has been the dominant fingerprinting approach in the literature due to the simplicity and practicality of measuring the RSS in a variety of wireless technologies (such as IEEE 802.11 and UMTS). Recognizing the diminishing gains using the RSS-based techniques, researchers have recently shifted focus to proposing improvements at the physical layer by adopting the channel impulse response (CIR) as an alternate fingerprint. In this paper we propose a novel fingerprint structure that is based on the entropy estimation of the channel; which provides a more unique/robust fingerprint that is capable of distinguishing between locations more effectively. Through extensive frequency domain channel measurements and analysis in a typical indoor environment we further validate the proposed technique and compare it against RSS and CIR-based fingerprinting. We will show that the technique combines the advantage of RSS-based fingerprinting simplicity of structure (storage and pattern recognition requirements) and improves on the robustness of the CIR-based fingerprinting techniques. Finally we will illustrate that our entropy-based location fingerprinting can be practically integrated into the architecture of popular OFDM-based WLAN systems.**

Keywords-Location Fingerprinting; WLAN localization; indoor localization; entropy estimation; CIR location fingerprinting; RSS fingerprinting; AR modeling

I. INTRODUCTION

Localization for indoor environments has gained considerable attention over the last decade due to the enormous potential in the technology and the significant challenges facing this area of research [1]. The localization technology provides the fundamental basis for a myriad of location-enabled services in indoor environments such as locating personnel and objects in residential homes, guiding shoppers inside a mall for the latest discount offers, locating the elderly in nursing homes or children tracking in day-care centers.

The complexity of the indoor wireless channel, the lack of Line of Sight (LOS) paths, and severe multipath and shadow fading problems make it difficult to accurately locate objects. One practical and popular localization technique that relies on the available fixed wireless infrastructure is location fingerprinting. Location fingerprinting combines the use of both radio frequency measurements and pattern recognition algorithms in order to find the best correlation between the pre-measurements (offline phase) and the real-time measurements (online phase). RSS-based location fingerprinting has been the dominant fingerprinting approach in the literature due to the simplicity and practicality of measuring the RSS in a variety of wireless technologies (such as IEEE 802.11 and UMTS). The majority of the work in the literature focuses on the pattern recognition stage of location fingerprinting [2], [3], [4]. However, results reported in [5], [6] showed that complex pattern recognition algorithms, such as neural networks or kernel regression, perform as well as the simple nearest neighbor (NN) or k-nearest neighbor. Recognizing the diminishing gains using the RSS-based techniques, researchers have recently shifted focus to proposing improvements at the physical layer by adopting the channel impulse response (CIR) as an alternate fingerprint. The CIR-based location fingerprinting essentially turns the disadvantage of the multipath to a powerful advantage by realizing that at each location the measured CIR provides a unique signature. This can be used to significantly improve the pattern recognition stage. In [7], CIR-based location fingerprinting was first introduced for cellular technology. The statistics of the CIR as fingerprints were implemented in [8], where RMS delay spread, number of multipath components, and power were used in a neural network framework for localization in mines. CIR-based location fingerprinting for WLAN systems showed great potential by exploiting the log of the CIR through Ray Tracing (RT) simulations [9]. The challenge of using CIR-based fingerprinting for WLAN systems, however, is the limited bandwidth available which results in low time-resolution fingerprints. This results in difficulty to distinguish between fingerprints that are 1 meter apart. Recently, a RT based evaluation of CIR fingerprinting for different system parameters highlighted the impact of system bandwidth on the practical achievable accuracy [10]. Another issue with CIR fingerprinting is the requirement for matrix storage and computations at the pattern recognition stage; since for M Access Points (APs) the fingerprint at each location is an $M \times N$ matrix, where N is the number of samples of the CIR vector. As a result, there is a need for a more accurate, robust and computationally efficient fingerprint structure for WLAN systems.

In this paper we introduce a novel fingerprint that is based on the entropy estimation of the channel which provides an accurate fingerprinting structure that reduces computation

978-1-4673-1954-6/12/$31.00 ©2012 IEEE

complexity and storage requirements. Since a CIR in a given location is a realization of a random process (random in space and time) then from information theoretic point of view the entropy of CIR can represent unique information at each grid point. In addition to the robust localization capabilities, entropy-based location fingerprinting eliminates the need for manipulation and storage of matrices thus reducing complex computation in the pattern recognition stage and reducing storage requirements of the *offline* fingerprint database. We will first provide an overview of entropy estimation through Autoregressive (AR) modeling techniques. Then we will illustrate the performance gains of our proposed fingerprint using measurement data collected in an indoor environment. Note that in this paper we will focus on the performance gains of *fingerprinting structures* rather than *fingerprinting algorithms* (*pattern recognition*). As a result the simple *k*-NN algorithm will be used due to its simplicity and the fact that other research papers have shown that its performance is close to more complex pattern recognition algorithms.

The paper is organized as follows. In section II we describe a general formulation of location fingerprinting using the simple *k*-NN technique. In section III we introduce entropy estimation for CIR fingerprints that is based on AR modeling and highlight how it can be used in localization. In section IV we provide a brief discussion on the measurement methodology. In Section V we present the results and analysis of our proposed technique. Finally the paper is concluded in Section VI.

II. LOCATION FINGERPRINTING

In typical fingerprinting-based location systems an RF fingerprint database is created in an offline stage by constructing the fingerprints/signatures through measurement of RF channel parameters such as RSS in different locations across a grid. In the online phase a mobile terminal in an unknown location constructs a fingerprint by measuring channel parameters to all APs within its coverage. This measured fingerprint is then compared to the offline database and the position is estimated using pattern recognition techniques. The simplest pattern recognition technique is the nearest neighbor where the position is estimated by selecting the location of the fingerprint in the database that is the closest (smallest distance in vector space) to the online measured fingerprint.

The fingerprint database is typically created by gridding a room/area in a given indoor environment. The grid is composed of L locations that are spaced by Δ. The coordinates of a given location on the grid can be represented by $\mathbf{p}_j = [x_j, y_j]^T$ where x_j and y_j are the x- and y-coordinates of the j^{th} location and $j \in [1, L]$. The *offline* fingerprint/signature at each grid location is given by the vector $\mathbf{z}_j = [z_j^1, \cdots, z_j^M]^T$ and each element is a measured parameter of the channel (e.g. RSS) from one of the m^{th} APs where $m \in [1, M]$ and M is the total number of APs covering the indoor environment. Note that for RSS and entropy, $\mathbf{z}$ is an $(M \times 1)$ vector and each element is a scalar value. However for CIR fingerprints, $\mathbf{z}$ is a $(M \times N)$ matrix and each element is a CIR vector of length N samples. In the online stage a fingerprint/signature $\mathbf{v}$ is similarly captured by the mobile device where $\mathbf{v} = [v^1, \cdots, v^M]^T$.

In the pattern recognition stage, an estimate of the position $\hat{\mathbf{p}} = [\hat{x}, \hat{y}]^T$ can be determined by choosing the nearest neighbor. For vector fingerprints such as RSS and entropy, the nearest neighbor is the offline fingerprint with the minimum Euclidean distance to the online fingerprint which is given by

$$d_{\min} = \arg\min_j \left\| \mathbf{z}_j - \mathbf{v} \right\|. \quad (1)$$

For CIR fingerprints the nearest neighbor will have the highest correlation to the offline fingerprint or

$$\rho_{\max} = \arg\max_j corr(\mathbf{z}_j, \mathbf{v}) \quad (2)$$

where ρ is the correlation coefficient and *corr()* is the correlation function. If more than one nearest neighbors is available then for the *k*-NN algorithm the position is estimated by

$$\hat{\mathbf{p}} = \frac{\sum_{i=1}^{k} \alpha_i \mathbf{p}_i}{\sum_{i=1}^{k} \alpha_i} \quad (3)$$

where *k* is the number of nearest neighbors selected and α_i is a weight that can be set to distinguish between the selected positions. For vector fingerprints such as RSS and entropy, the weights can be given by $\alpha_i = 1/d_i$ where d_i is the Euclidean distance between the offline and online fingerprints. For CIR fingerprints the weights can be given by the correlation coefficient or $\alpha_i = \rho_i$.

III. ENTROPY FINGERPRINTS

In this section we will first introduce how entropy estimation of signals can be achieved and then outline how it will be used to create a novel fingerprint structure.

A. Entropy Estimation

From information theory the entropy of a random variable X is given by the well known Shannon relation [11]

$$H(X) = -E_X\left[\log_2 p_X\right] = -\int_{-\infty}^{\infty} p_X(x) \log_2 p_X(x) dx \quad (4)$$

where $p_X(x)$ is the PDF of X. In practice, direct evaluation of (4) is difficult because it is not easy to compute or estimate the PDF from real data. As a result the key in accurate entropy estimation is the ability to estimate the PDF of a random variable in real-time. Typical methods rely on estimating the PDF through histograms [12], order statistics [13] or kernel

methods [14]. A more practical and efficient alternative technique to estimate the PDF of a random signal has been proposed by [15][16] where the PDF of a random variable X can be estimated simply by appealing to the theory of Power Spectral Density (PSD) estimation. An estimate of the PDF $\hat{p}_X(x)$ can be parameterized by a set of coefficients $\{a_k\}$ of an autoregressive (AR) model in the form of a PSD $S_W(x)$, where $1 \le k \le p$ is the number of parameters [15,16] or

$$\hat{p}_X(x) = S_W(x) = \frac{\sigma_w^2}{\left|1 - \sum_{k=1}^{p} a_k e^{-j2\pi kx}\right|^2} \tag{5}$$

where σ_W^2 is designed such that $\int_{-1/2}^{1/2} S_W(f)df = 1$ since PSDs are different from PDFs in that they do not usually integrate to 1. The bounded support [-1/2, 1/2] ensures that the random variable is constrained between these values, since a general PDF is not periodic with period one as is imposed by the AR model [15]. This can be easily achieved by normalizing the data by $k\sigma_W$ where σ_W is the standard deviation and k is a suitable parameter [15]. One way to clarify the modeling of the PDF by a PSD is to note that if $X(\omega, n)$ is a random process then we can find a process $W(\omega, n)$ that has a PSD that matches $p_X(x)$. One such process is $W(\omega, n) = e^{j(nX + \varphi(\omega))}$ where $\varphi(\omega)$ is uniformly distributed over $[0, 2\pi]$ and independent of X [16]. It can be easily shown that the autocorrelation of W, $R_W(k)$, is the first characteristic function of X. The relationship between the PDF and the PSD can be highlighted through the following Fourier Transform relationships

$$R_W(k) \xrightarrow{F} S_W(x) \tag{6a}$$

$$\phi_x(k) \xrightarrow{F} p_X(x) \tag{6b}$$

It is clear that if the autocorrelation of the process W, $R_W(k)$, is equal to the characteristic function of the random variable X, $\phi_X(x)$, then $p_X(x) = S_W(x)$. As a result to estimate the PDF $\hat{p}_X(x)$ we can find an AR model of the form in (5). In order to estimate the model parameters in (5) based on available data $\{x_1, x_2, \ldots, x_N\}$ it is well known that the Yule-Walker equations relate the AR model parameters to the autocorrelation function [17]. The autocorrelation function in this case is the samples of the characteristic function given by [15]

$$F^{-1}\{p_X(x)\} = \phi_x(k) = E\left[e^{j2\pi kx}\right] = R_W(k) \tag{7}$$

where $\phi_x(k)$ is the characteristic function and $R_W(k)$ is the autocorrelation function of the underlying process W related to the PSD $S_W(x)$. The autocorrelation function essentially becomes samples of the characteristic function and in the remainder of the paper $\phi_x(k)$ and $R_W(k)$ will be used interchangeably and they both refer to (7). To estimate the AR model parameters (ultimately estimate the PDF) we need to estimate the autocorrelation of the random process which can be given by the sample moment estimator

$$\hat{R}_W(k) = \frac{1}{N}\sum_{n=0}^{N-1} e^{j2\pi kx(n)} \tag{8}$$

where $k = 0, \ldots, p$, p is the model order and N is the number of samples in the data vector. The AR model parameters can then be found by solving the Yule-Walker equations [15,17]

$$\hat{\mathbf{R}}_W \hat{\mathbf{a}} = \hat{\mathbf{r}}_W \tag{9}$$

where

$$\hat{\mathbf{R}}_W = \begin{bmatrix} \hat{R}_W(0) & \hat{R}_W(-1) & \cdots & \hat{R}_W(-(p-1)) \\ \hat{R}_W(1) & \hat{R}_W(0) & \cdots & \hat{R}_W(-(p-2)) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{R}_W(p-1) & \hat{R}_W(p-2) & \cdots & \hat{R}_W(0) \end{bmatrix}$$

$\hat{\mathbf{a}} = \begin{bmatrix} \hat{a}(1) & \hat{a}(2) & \cdots & \hat{a}(p) \end{bmatrix}^T$ and

$\hat{\mathbf{r}}_W = \begin{bmatrix} \hat{R}_W(1) & \hat{R}_W(2) & \cdots & \hat{R}_W(p) \end{bmatrix}^T$.

The Levinson-Durbin recursive algorithm can be used to solve for the coefficients. Similarly, an estimate of $\hat{\sigma}_W^2$ can be computed (once $\hat{a}_k$ are estimated) by

$$\hat{\sigma}_W^2 = \hat{R}_W(0) - \sum_{k=1}^{p} \hat{a}(k)\hat{R}_W(-k) \tag{10}$$

Once the AR parameters of the PSD that models the PDF are obtained then the estimate of entropy can be computed by

$$\hat{H} = -\int_{-1/2}^{1/2} \hat{p}_X(x)\log_2 \hat{p}_X(x)dx = -\int_{-1/2}^{1/2} S_W(x)\log_2 S_W(x)dx \tag{11}$$

A more practical expression can be obtained using Plancherel-Parseval formula to the right-hand side of (11) [16] which yields

$$\hat{H} = -\sum_{k=-\infty}^{\infty} R_W(k)C_W^*(k) \tag{12}$$

where $R_W(k)$ is the k^{th} correlation coefficient and $C_W^*(k) = F^{-1}\{\log_2 S_W(x)\}$ is the k^{th} component of its cepstrum [17]. Since $S_W(x)$ is real, both $R_W(k)$ and $C_W^*(k)$ have Hermitian symmetry and thus (12) can be given by

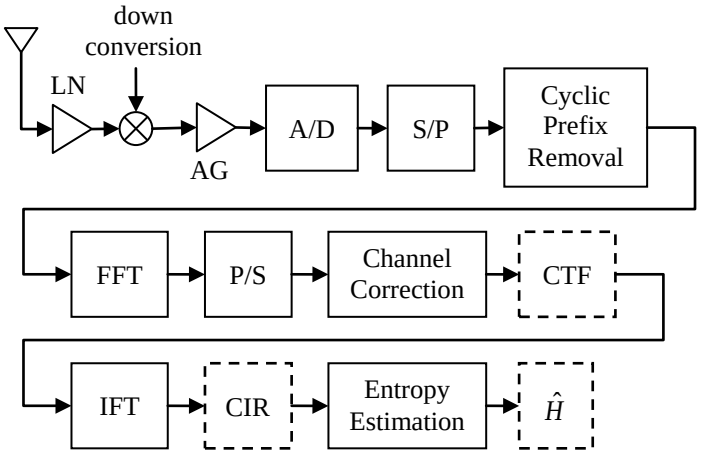


Figure 1: System diagram of entropy estimation in OFDM-based WLAN systems

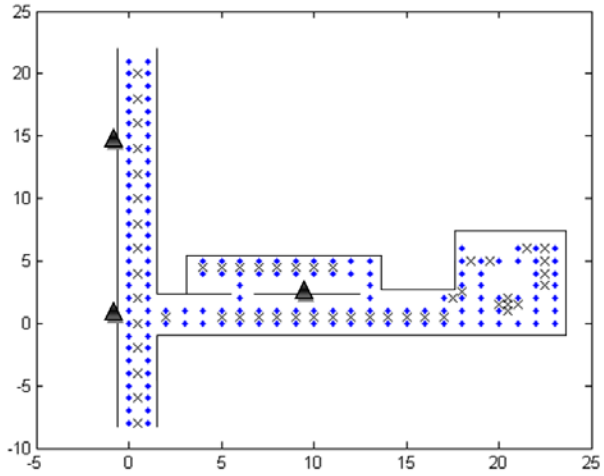


Figure 2: Floor plan and measurement locations. Triangle – Transmitter locations; dots – offline receiver locations; 'x' – online receiver locations.

$$\hat{H} = -2\,\mathrm{Re}\left\{\sum_{k=0}^{\infty} R_W(k) C_W^*(k)\right\}. \qquad (13)$$

From (5) it is clear that the entropy estimate will depend on the AR model order. A very well known model order selection technique is the Minimum Description Length (MDL) criterion where the idea is select a p that minimizes the following cost function [17]

$$C(p) = N \log \varepsilon_p + (\log N)\, p \qquad (14)$$

where N is the number of data samples, p is the model order and ε_p is the modeling error typically given by $\hat{\sigma}_W^2$ in (10).

B. Entropy Fingerprint in WLAN Systems

In order to estimate the entropy of the channel, an estimate of the channel is required. In a generic IEEE 802.11a/g that is based on OFDM it is possible to estimate the CIR from the channel transfer function (CTF) estimation that is carried out for channel correction prior to symbol demodulation. **Figure 1** illustrates a typical OFDM receiver with entropy estimation. For a given CIR estimate it is possible to compute the entropy using the procedure described in the previous section. At a given location covered by *M* access points it is then possible to construct the entropy fingerprint as $\mathbf{z}_j = [\hat{H}_j^1, \hat{H}_j^2, \cdots, \hat{H}_j^M]^T$ where each element is an entropy estimate to one of the APs covering the environment. Note that the entropy fingerprint is a vector at a given location and thus shares RSS's simplicity, but is more accurate due to the information representation of entropy.

IV. Measurements Database

A. Background

In order to analyze the proposed approach, a database of measured CIRs was collected in the academic building of Khalifa University in Sharjah, UAE. Frequency-domain measurement techniques have been utilized to characterize the CIR for communication and geolocation applications in the past [18,19]. The measurement system provides estimates of the CTF in the frequency domain which can then be post-processed to obtain the CIR. Both entropy and RSS can then be derived from the measurements. The measurements were conducted in the WiFi 2.4 GHz band.

B. Measurement System

The measurement system is based on Rhode and Schwartz's Vector Network Analyzer (VNA) that has the capability of measuring S-21 parameter (CTF) up to 11 GHz. In order to increase the dynamic range of the system a 30 dB power amplifier at the transmitter and 17 dB low-noise amplifier at the receiver were used. The overall system dynamic range is 120 dB. The antennas used were omni-directional 2.4 GHz with 5 dB gain. The height of the transmitter and receiver antennas was fixed to 1.5 m. The VNA, transmitter and receiver were connected by low-loss RF cables.

C. Measurement Procedure

Part of the ground floor of the academic building was gridded with points spaced by 1 meter. The area covered two main corridors, the RF lab and the student area. **Figure 2** illustrates the floor plan and the measurement locations. Three AP locations were measured (these are actual AP locations of the existing WiFi network). Note that in many buildings in UAE, extensive number of APs is generally used to provide coverage to the users. This is mainly due to the fact that the buildings have concrete exterior and interior walls. In addition a heat insulation layer is usually inserted in the middle of the concrete blocks to further insulate the buildings from the external heat. As a result the measurement database and results are unique from that point.

Since the VNA has only a transmitter and a receiver port, multiple AP transmission had to be emulated by repeated movement of the transmitter from AP locations 1 to 3. That is in the first round the AP transmitter is located in TX1 location and the receiver covers all the offline and online points. In the

second time the transmitter is placed in the second TX2 location and the measurements are repeated and so on. Since the fingerprints were not measured at the same time from the three APs, care was taken to ensure uniform and controlled channel environments. That is measurements were conducted in times of low activity (early morning/evening) to ensure that the channel exhibits stationary properties.

At each location, 40 sequential measurement snapshots were recorded which is important to capture the variation in time for different motion scenarios. During the measurements, two motion scenarios were measured: the first is stationary where there are no movements around the vicinity of the transmitter or receiver. For the second, 3 people walked around the TX/RX path to emulate random motion that could be encountered in indoor environments. This is important since in practical scenarios the motion of individuals/objects can significantly change the fingerprints and degrade the localization. In general it is possible to assume that in the offline stage no motion exists (since the survey is usually conducted late at night/weekends) and motion will affect the online fingerprints only.

Finally, the measurements were conducted to cover the entire 2.4 GHz band. Specifically the measured bands that were analyzed are for the 1st non-overlapping option specified by the IEEE 802.11g standard that is channels 1, 6 and 11. In this paper we analyze the performance of the fingerprints using the measurements from channel 1 since it has the minimum interference from nearby operational 802.11 APs. In fact by looking at the data we could verify that band 6 experienced the most interference and analysis of the impact of interference on the fingerprints is our future work.

V. Results and Analysis

A. Overview

In this section the localization performance of the proposed entropy fingerprinting technique is presented. Further we compare the performance with two popular fingerprinting techniques, namely RSS and CIR. RSS fingerprints are obtained from the CIR which is typically modeled as

$$g(\tau) = \sum_{k=1}^{L_p} \beta_k e^{j\phi_k} \delta(\tau - \tau_k) \qquad (15)$$

where β_k, ϕ_k and τ_k are the amplitude phase and delay of the k^{th} path, respectively. Note that for lower bandwidth systems such as WLAN, *clusters* of path arrivals are detected instead of individual path; that is due to the low time-resolution and inability to resolve the multipath components. In the measurement data, CIR is obtained from the frequency domain CTF by inverse Fourier Transform. RSS can then be computed by detecting K multipath components that are above the noise threshold (typically -100~-110 dBm) and it is given by

$$RSS = 10\log_{10}\left(\sum_{i=1}^{K} |\beta_i|^2\right). \qquad (16)$$

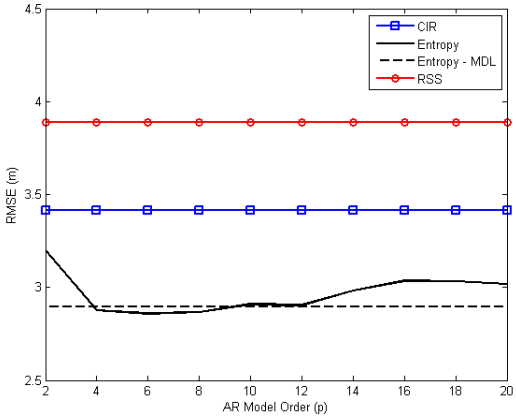
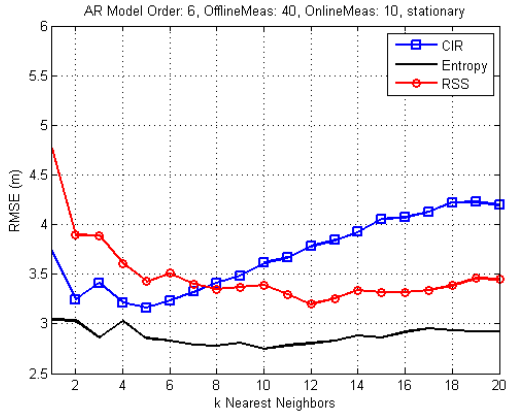


Figure 3: Impact of AR model order selection on the performance of entropy-based location fingerprinting.

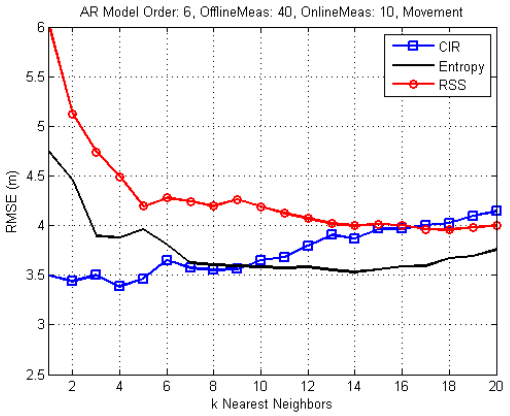
B. Performance of Entropy-based Location Fingerprinting

The main parameter that is important to entropy estimation using the AR modeling approach is the model order. In order to find the appropriate model order, the AR modeling has to be repeated for a selected number of model orders and then the one with the lowest modeling error according to MDL (14) is chosen. In order to test the importance of the model order on the performance we analyzed the RMSE against different model orders from 2-20 and compared it to the MDL. **Figure 3** summarizes the results. The RMSE for RSS and CIR are included for comparison as well and they are constant since they do not depend on the AR model order. The results obtained are for 40 time measurements (offline), 10 time measurements (online), $k = 3$ nearest neighbors and the stationary scenario. Entropy-based fingerprinting in both MDL and specific model order outperforms RSS and CIR. It is clear that the model order in the range of 4-6 provides the best performance. In fact results presented in [20] highlighted that the wireless multipath channel can be modeled as an AR process and showed that a model order of 2 is sufficient. Since model order of 6 provides the best performance then in the remainder of this paper it will be used instead of the MDL approach. This further reduces the computational requirements of repeated model order testing. In addition a model order of 6 reduces the computational requirements of solving the Yule-Walker equations (9) for many different model orders. Since this model order is sufficient then it will be used for the remainder of all the entropy estimation, simulations and results.

In Figure 4, the location fingerprinting performance (RMSE) using the *k*-NN approach is presented for stationary and moving scenarios with weights that are related to the Euclidean distance and correlation for RSS/entropy and CIR, respectively. The results are obtained by averaging the measured CIR at each location: an average of 40 time measurements for the offline points and 10 time measurements for the online/test points. For estimating entropy, the AR model order was set to 6.



(a)



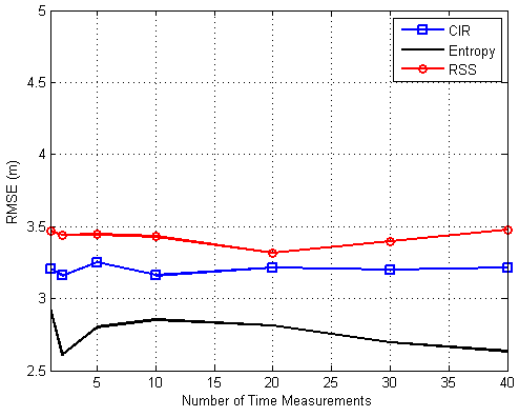
(b)

Figure 4: Performance against number of nearest neighbors selected. (a) Stationary – no motion. (b) Motion

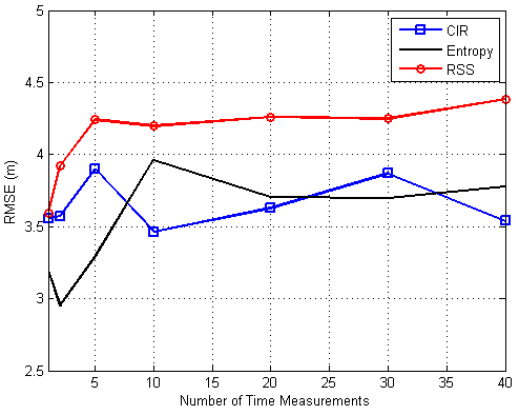
In general for RSS and entropy, the results show that increasing the number of nearest neighbors used to estimate the position improves the performance. However this is not the case for CIR and this is mainly attributed to the fact that for different points along the corridors, the CIRs are strongly correlated with each other. This correlation can be explained by the structure of the corridors that imposes a certain signature (multipath cluster arrivals). Thus a nearest neighbor with high correlation can be in fact in the opposite side of the corridor.

In the stationary scenario, entropy outperforms CIR and RSS due to its ability to capture the information in the CIR more accurately thereby enabling more accurate fingerprints. For the motion scenario entropy outperforms RSS but CIR performs better for $k < 9$ due its robustness to motion which was validated experimentally in [21]. However for higher number of neighbors entropy outperforms CIR for the same reason as discussed for the stationary case.

It is clear that the motion around the transmitters/receiver can have significant impact on the performance since it can alter the fingerprints. Thus it is necessary to analyze the performance against different time averages. In practice more time measurements can be taken for the offline points (since it's conducted in the survey stage) compared to the online points which require real-time measurements and localization. Figure 5 illustrates the impact of the number of time measurements on the RMSE of the location estimate for $k = 5$ nearest neighbors. The number of offline time measurements was fixed at 40 while the online measurements were varied from 1 to 40. Again in the stationary scenario (Fig 4 (a)), entropy outperforms both RSS and CIR. In addition, as the number of measurements increase entropy RMSE further decreases. This is attributed to the fact that averaging the CIR provides a "cleaner" signal from which entropy can be estimated – which results in more unique fingerprints. For the motion scenario, RSS degrades with increasing number of time measurements because RSS is an averaging notion and motion affects a significant portion of the 40 time measurements. Entropy's performance is close to the CIR but does not exhibit the deterioration with time measurements that affects RSS fingerprinting.



(a)



(b)

Figure 5: Performance against number of time measurements used to estimate the average of the CIR. For the offline fingerprints 40 time measurements and 1-40 measurements for the online fingerprints. (a) Stationary – no motion. (b) Motion.

VI. CONCLUSION

In the past several years, RSS fingerprinting has gained attention as a practical solution to the indoor localization problem. Research work focused on improvements in the pattern recognition stage. However, comparative studies showed diminishing improvements in accuracy mainly due to the weakness of RSS as a fingerprint. In this paper, we have introduced entropy as a novel fingerprinting structure that characterizes the multipath information in the CTF/CIR between a transmitter and a receiver.

Using AR modeling we have illustrated that it is possible to estimate entropy accurately and incorporate it in a location fingerprinting framework. We further verified the effectiveness of entropy as a fingerprint by conducting and analyzing measurements in a typical office environment. The results show that in stationary conditions between the transmitter and receiver (no movements) entropy outperforms both RSS and CIR. Increasing the number of nearest neighbors improves entropy further. In addition increasing the number of time measurements (averaging) can enhance entropy estimation and thus uniqueness of the fingerprint. In motion scenarios RSS exhibits the worst performance, while entropy can match and outperform CIR in most scenarios.

Finally the value of entropy as a fingerprint is based on the fact that it can capture the information inherent in CIR fairly accurately and reduces the burden of matrix correlations/computations that CIR suffers from. Essentially it provides the uniqueness of CIR (in most cases better than CIR) with the simplicity of RSS fingerprint structure (vector fingerprint processing).

ACKNOWLEDGMENT

We would like to thank Prof. Saleh Al-Araji and Dr. Nazar Ali and Nuha Al-Khanbashi from Khalifa University for their help with the measurements.

REFERENCES

[1] K. Pahlavan, X. Li and J. Makela, "Indoor geolocation science and technology," *IEEE Commun. Mag.* , vol. 40, no. 2, pp. 112-118, Feb. 2002.

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

[3] T. Roos, P. Myllymaki, H. Tirri, P. Misikangas and J. Sievanen, "A probabilistic approach to WLAN user location estimation," *International Journal of Wireless Information Networks (IJWIN)*, vol. 9, pp. 155-164, July 2002.

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

[5] V. Honkavirta, T. Perala, S. A.-Loytty and R. Piche, "A comparative survey of WLAN location fingerprinting methods," in *Proc. of the 6th Workshop on Positioning, Navigation and Communication (WPNC'09)*, pp. 243-251, 2009.

[6] T. Lin and P. Lin, "Performance comparison of indoor positioning techniques based on location fingerprinting in wireless networks", in *Proc. International Conference Wireless Network, Communications and Mobile Computing*, vol. 2, pp. 1569-1574, June, 2005.

[7] S. Ahonen and P. Eskelinen, "Performance Estimations of Mobile terminal location with database correlation in UMTS networks," in *Proc. Of Int'l Conf. on 3G Mobile Communication Technologies*, pp. 25-27, June, 2003.

[8] C. Nerguizian, C. Despins and S. Affes, "Geolocation in mines with an impulse response fingerprinting technique and neural networks," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 603-611, Mar. 2006.

[9] Y. Jin, W.-S. Soh, W.-C. Wong, "Indoor localization with channel impulse response based fingerprint and nonparametric regression", *IEEE Transactions on Wireless Communications*, vol. 9, no. 3, pp. 1120-1127, Mar. 2010.

[10] N. Al-Khanbashi, N. Alsindi, S. Al-Araji, N. Ali and J. Aweya, "Performance evaluation of CIR based location fingerprinting," *in Proc. IEEE Int'l Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 9-12 Sept., 2012, Sydney, Australia.

[11] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379-423; 623-656, July/Oct. 1948 [Online] Available http://cm.bell-labs.com/cm/ms/what/shannonday/paper.html.

[12] R. Moddemeijer, "On estimation of entropy and mutual information of continuous distributions," *Signal Processing*, vol. 16, no. 3, pp. 233-246, 1989.

[13] O. Vasicek, "A test of normality based on sample entropy," *J. R. Stat. Soc. Ser. B*, vol. 38, pp. 54-59, 1976.

[14] P. Viola, N. N. Schraudolph and T. J. Sejnowski, "Empirical entropy manipulation for real-world problems," in *Advances in Neural Information Processing Systems 8*. Cambridge, MA: MIT Press, 1996.

[15] S. Kay, "Model-based probability density function estimation," *IEEE Signal Processing Letters*, vol. 5, no. 12, Dec. 1998.

[16] J.-F. Bercher, C. Vigant, "Estimating the entropy of a signal with applications", *IEEE Trans. on Signal Processing*, vol. 48, no. 6, pp. 1687- 1694, June 2000.

[17] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*, John Wiley & Sons, Inc., 1996.

[18] S. J. Howard and K. Pahlavan, "Measurement and analysis of the indoor radio channel in the frequency domain," *IEEE Trans. Instrum. Meas.*,vol. 39, no. 5, pp. 751–755, Oct. 1990.

[19] N. Alsindi, B. Alavi, K. Pahlavan, "Measurement and modeling of ultrawideband TOA-based ranging in indoor multipath environments," *IEEE Trans. on Vehicular Tech.*, vol. 58, no. 3, March 2009.

[20] S. Howard and K. Pahlavan, "Autoregressive modeling of wide-band indoor radio propagation," *IEEE Trans. on Commun.*, vol. 40, no.9, pp. 1540-1552, Sept. 1992.

[21] J. Zhang, M. H. Firooz, N. Patwari and S. K. Kasera, "Advancing wireless link signatures for location distinction," *In Proc. of the 14th ACM Int'l Conf. on Mobile Computing and Networking (MobiCom) '08*, New York, USA, 2008.