Wi-Fi Signals Database Construction using Chebyshev Wavelets for Indoor Positioning Systems

Anvar Narzullaev#, Zahriddin Muminov*, Gafurjan Ibragimov+

#Faculty of Science and Technology, Universiti Sains Islam Malaysia, 71800 Nilai, Negeri Sembilan, Malaysia
E-mail: anvar@usim.edu.my

*Faculty of Engineering and Technology, Nilai International University, 71800 Nilai, Negeri Sembilan, Malaysia
E-mail: zimuminov@gmail.com

+Faculty of Science, University Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia
E-mail: ibragimlov@gmail.com

Abstract— Nowadays fast and accurate positioning of assets and people is as a crucial part of many businesses, such as, warehousing, manufacturing and logistics. Applications that offer different services based on mobile user location gaining more and more attention. Some of the most common applications include location-based advertising, directory assistance, point-to-point navigation, asset tracking, emergency and fleet management. While outdoors mostly covered by the Global Positioning System, there is no one versatile solution for indoor positioning. For the past decade Wi-Fi fingerprinting based indoor positioning systems gained a lot of attention by enterprises as an affordable and flexible solution to track their assets and resources more effectively. The concept behind Wi-Fi fingerprinting is to create signal strength database of the area prior to the actual positioning. This process is known as a calibration carried out manually and the indoor positioning system accuracy highly depends on a calibration intensity. Unfortunately, this procedure requires huge amount of time, manpower and effort, which makes extensive deployment of indoor positioning system a challenging task. An approach of constructing signal strength database from a minimal number of measurements using Chebyshev wavelets approximation. The main objective of the research is to minimize the calibration workload while providing high positioning accuracy. The field tests as well as computer simulation results showed significant improvement in signal strength prediction accuracy compared to existing approximation algorithms. Furthermore, the proposed algorithm can recover missing signal values with much smaller number of on-site measurements compared to conventional calibration algorithm.

Keywords— indoor positioning; chebyshev wavelets; Wi-Fi fingerprinting; real-time locating systems.

I. INTRODUCTION

Tracking assets and equipment has been a long-standing issue for enterprises. In warehouses, factories, and hospitals, for example, staff and important assets should be located and tracked rapidly. A number of indoor positioning systems (IPS) grow steadily for the past decade thanks to various wireless technologies, including Wi-Fi, Bluetooth, ZigBee, Ultra-wideband, etc, which gave the enterprises the opportunity to leverage on some form of control over the manageability of its resources [1]. The number of systems that provide different services based on the user’s location, aka Location Based Services or LBS, is overgrowing and considered as a promising market, especially in advertising, healthcare, social networking, and entertainment. Currently, the most extensively used positioning system is the Global Positioning System (GPS). Global availability and high degrees of accuracy make the GPS a perfect candidate for LBS. However, GPS signals transmitted from satellites fail to penetrate indoors and crowded urban, urban areas [2]. Thus, many researchers focused on developing alternative positioning systems for such impenetrable environments. Different technologies, including Cellular, digital TV, Ultra-wideband, ultrasonic, Bluetooth, and Wi-Fi have been suggested to be a backbone for IPS. Among the technologies mentioned above, Wi-Fi networks based on IEEE802.11 standard are rapidly growing, widely deployed and most importantly supported by a wide range of mobile devices. Wi-Fi networks and hotspot are available almost everywhere. Wi-Fi chips can be found in all sorts of mobile devices: smartphones, laptops, tablets, TV and even home appliances. Initially intended to provide internet access to mobile users, IEEE802.11 WLAN emerged as a feasible solution for IPS [3].
Specific characteristics of an indoor environment and distinct requirements of time-based location estimation algorithms make received signal strength information (RSSI) based location estimation techniques more appealing for IPS [4]. The well-known RSSI based location estimation technique is the fingerprinting algorithm that compares certain signal properties such as RSSI against a database (DB) of properties previously collected at different locations. The closest match from DB is returned as the estimated position. Accuracies of a few meters are typically reported [5]. However, the main drawback of the algorithm is an immense calibration phase, that requires a huge amount of time, workforce and effort to build a signal database. Since the Wi-Fi signal characteristics highly depend on indoor environment parameters, newly collected signal samples may be invalid due to major transformations to the target area (e.g., renovation, construction). Thus, deploying large-scale indoor positioning systems is a challenging task.

Many researchers have been conducting studies to reduce the calibration burden and make IPS more feasible and practical. A crowdsourcing system, where users are encouraged to fix the positioning errors by pinpointing correct location has been proposed [6], [7]. Several researchers proposed signal propagation models that effectively estimate RSSI values indoors [8], [9]. However, most of these models require comprehensive layout information, such as the number of floors, obstacles (walls, doors, partitions) and most fundamentally the physical location of wireless stations. Propagation models without such prerequisites reported having a low prediction accuracy, which is not suitable for IPS.

In this manuscript, we propose a novel, hybrid approach to construct RSSI signal DB by combining conventional signal sampling procedure with Chebyshev wavelets based prediction algorithm. The conventional sampling procedure is very straightforward: A technician equipped with a mobile device moves around the building and collect RSSI samples at pre-allocated reference points (RPs). Collected RSSI samples, after some processing, are stored in the fingerprints DB. Indoor positioning system accuracy significantly affected by the number of RPs. Higher the number, better the accuracy. Thus, the number of RPs may go up to several hundred per floor for an average building. The main objective of this research is to develop a time efficient calibration algorithm that will create an accurate and complete fingerprints DB from a few on-site measurements using Chebyshev wavelet transform.

Recently, various researches considering the Wavelet transforms as an effective tool for signal processing. Wavelet Transforms are being used to represent a discrete signal in a more redundant form, often as preconditioning for data compression. Practical applications of the Wavelet transform can be found in data analysis, digital communications, signal and image processing, and many others. The ability to perform the local analysis is a critical feature of wavelet functions. Many methods fail to detect various signal aspects, such as trends, breakpoints, and discontinuities, whereas they can be shown by wavelet analysis [10]. Our computations show that using wavelet transforms in constructing an accurate signal strength DB for IPS is significant. Because continuous and piecewise continuous functions can be approximated well by the Chebyshev polynomials [10] to any desired precision, over a prescribed interval, the present study is based on the Chebyshev wavelets [11], though there are various types of wavelets. We obtain a system of linear algebraic equations to find the Chebyshev wavelet transform coefficients, although, in many works devoted to discrete Wavelet transforms, the coefficients are produced using an inner product [12].

The proposed algorithm has been carefully evaluated through field tests and computer simulations. Real-world experiments were conducted in one of the biggest shopping malls in Klang Valley of Malaysia. Obtained results showed a significant reduction in calibration time while providing high positioning accuracy compared to conventional and contemporary calibration algorithms.

The rest of the manuscript is organized as follows. Section II describes the basic principles of RSSI fingerprinting algorithm and RSSI approximation methods. In Section III we describe the experimental setup and present field test results and analyses. Concluding remarks are given in Section IV.

II. MATERIAL AND METHOD

A. RSSI-fingerprinting Location Estimation

Fingerprinting localization technique is based on comparing the unique RSSI samples collected from surrounding sources at a particular location with a database of pre-collected data. RSSI-fingerprinting has two phases: offline (calibration) and online (positioning). The calibration is the process of constructing signals database by sampling RSSI values from surrounding APs at different locations. During the positioning phase, the mobile user’s position is estimated by comparing the samples collected by the user’s mobile device with the values stored in the database. The outcome is the likeliest location of the mobile user [13].

An advantage of WLAN based fingerprinting is the rather small number of APs that required to estimate the mobile users’ position. Also, due to the rapid growth of wireless networks in indoor environments, existing wireless infrastructure can be utilized for positioning as well, seriously reducing the IPS deployment costs [15]. On the other hand, creating a fingerprint DB is often time-consuming and, furthermore, the newly collected DB may not be reliable if there are notable changes in the indoor environment. Consequently, large-scale deployments of fingerprinting localization for indoors become non-trivial [16].

B. RSSI Approximation

The time-consuming calibration phase is the main drawback of RSSI-fingerprinting location estimation algorithm. Re-calibration is another well-known issue for such systems. As signal propagation is sensitive to environmental parameters, even slightest changes to the indoor characteristics may result in signal alterations. Numerous researches have attempted to solve the re-calibration issue by proposing various signal propagation prediction models and approximately estimating the RSSI values [8]. While some of these models are computationally simple, such as one-step model [16], they fail to reflect the
indoor propagation features carefully. On the other hand, more accurate propagation models like multi-wall model [17] and new empirical model [18] are complex in computation and require more detailed layout information, such as, the number of obstacles (walls, floors, doors etc.) the signal passes through on their way from an AP to a given location, obstacles' density and even an angle of penetration. Also, the main drawback of prediction models is that they do not take into account some important signal propagation effects, such as diffraction, reflection, and scattering [19].

Lately, several researchers attempted to reduce the calibration efforts via signal interpolation. These methods generally approximate the missing RSSI data by interpolation of the measurements at nearby reference points. In [20] authors presented an efficient database construction method, from relatively sparse measurements, based on low-rank approximation. Additionally, to eliminate the signal interference effect, the basic low-rank complement model is combined with the signal spatial consistency model. However, this method starts with the assumption that the positioning area is represented as a rectangle (to create RSSI measurements matrix) which is not practical for real-life scenarios. Another interesting approach is presented in [21]. In this research, authors combined signal sampling procedure with the Nelder-Mead simplex algorithm. In order to take into account signal fading caused by walls and other obstacles, researches refine the approximation model for each cell of the target area by a higher-order Voronoi diagram. Although the proposed method reduces the calibration time, the location accuracy of the model is a bit on the downside. In our previous study [22] we presented an interpolation based calibration algorithm that utilizes a modified one-slope model (MOSM). Instead of using fixed reference path-loss values, the proposed method requires a small number of reference samples collected in a target area. Collected samples are used to estimate the location of Wi-Fi access points, first. Next, the RSSI database is constructed using reference samples and AP location information. While the algorithm provides good RSSI prediction and location estimation accuracy, the computational complexity of the method is rather sophisticated. Although there are numerous approximation methods as mentioned above, there is still a room for improvement in terms of increased accuracy, reduced efforts, and complexity.

C. Methodology

The proposed RSSI database construction algorithm involves two procedures: collecting reference Wi-Fi signals and RSSI estimation via Chebyshev wavelets. The objective of this research is to create an accurate and comprehensive signal database from a minimal number of on-field measurements.

We start by identifying the paths within the target area, that is used by people to navigate the building. Next, we allocate multiple reference points along each path and start measuring RSSI values at each RP concerning individual APs. Next step is to create path-specific approximation functions using Chebyshev wavelets.

The Fourier analysis can be used to represent the frequency domain characteristics of the stationary signals with given periods. However, if the signal has non-stationary characteristics such as abrupt asynchronous changes, drifts, and frequency trends, then Fourier analysis, in general, is not suitable. Wavelet analysis, at the same time, uses adaptive windows to deal with these problems. Long windows retrieve low-frequency information, whereas short windows are used for high-frequency information [23]. One of the dominant characteristics of the wavelet transform is the ability to carry out flexible localized analysis. Discrete Wavelet is a family of functions generated by single function called the mother wavelet. This family is double indexed and is defined by writing

$$\Psi_{k,n}(x) = 2^\frac{k}{2} \psi(2^k x - n), \quad k, n \in \mathbb{Z},$$

(1)

where \( \{\psi_{k,n}\} \) form a wavelet basis for \( L^2(\mathbb{R}) \).

Chebyshev wavelets \( \psi(k^n, n, m, x) \), have four arguments, where \( n = 0, 1, \ldots, N-1 \), \( m = 0, \ldots, M \), is a degree of Chebyshev polynomials of the first kind and \( x \) belongs to the real set:

$$\Psi_{n,m}(x) = \begin{cases} \frac{2}{a} x - 2n + 1, & \frac{n-1}{a} \leq x < \frac{n}{a} \\ 0, & \text{otherwise}, \end{cases}$$

where \( M \) and \( N \) are any positive integers and

$$T_m(x) = \begin{cases} \frac{1}{\sqrt{\pi}}, & m = 0 \\ \frac{2}{\pi} T_m(x), & m > 0 \end{cases}$$

Here, \( T_m(x) \) are Chebyshev polynomials of the first kind of degree \( m \) which are orthogonal concerning the weight function \( w(x) = (1-x^2)^{-\frac{1}{2}} \), and satisfy the following recursive formula:

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$

(2)

where \( m = 1, 2, \ldots \).

The function may approximate signal function \( u(t) \) in the interval \([0, a]\)

$$u(t) = \sum_{n=0}^{N-1} \sum_{m=0}^{M} u_{n,m} \Psi_{n,m}(x)$$

(3)

where \( u_{n,m} \) are unknown coefficients to be found.

Unknown coefficients \( u_{n,m} \) will be calculated using the measured signal strength values \( u(t) \) at the points

$$t_{i,j} = \frac{i-1}{N} a + \frac{j-1}{MN} a, \quad t_{i,j} \in \left[ \frac{i-1}{N} a, \frac{i}{N} a \right],$$

where \( i = 1, \ldots, N \) and \( j = 0, \ldots, M \).
Let \( a_{ij} \) be measured RSSI values \( u(t) \) at the points \( t_{ij} \), i.e. 
\[ a_{ij} = u(t_{ij}). \]

Since the support \( \Psi_{n,m}(x) \) is the interval 
\[ \text{supp}(\Psi_{n,m}(x)) = \left[ \frac{n-1}{N} a_i, \frac{n}{N} a_j \right], \]
for any \( n_1 \) and \( n_2 \), \( n_1 \neq n_2 \), the supports \( \text{supp}(\Psi_{n,m}(x)) \) and \( \text{supp}(\Psi_{n_2,m}(x)) \) are disjunctive.

Therefore, we obtain from (2) that 
\[ u(t_{ij}) = \sum_{m=0}^{M} a_{i,m} \Psi_{i,m}(t_{ij}), \quad i = 1, \ldots, N, \quad j = 0, \ldots, M. \]

Consequently, using last equalities and 
\[ a_{i,j} = u(t_{ij}) \]
to find unknown coefficients \( u_{i,j} \), we obtain the following \( N \) systems of \( M+1 \) equations:
\[ \sum_{m=0}^{M} u_{1,m} \Psi_{1,m}(t_{1,j}) = a_{1,j}, \quad j = 0, \ldots, M, \]
\[ \sum_{m=0}^{M} u_{2,m} \Psi_{2,m}(t_{2,j}) = a_{2,j}, \quad j = 0, \ldots, M, \]
\[ \vdots \]
\[ \sum_{m=0}^{M} u_{N,m} \Psi_{N,m}(t_{N,j}) = a_{N,j}, \quad j = 0, \ldots, M, \]
\[ (4) \]

Any subsystem of (4) 
\[ \sum_{m=0}^{M} u_{n,m} \Psi_{n,m}(t_{n,j}) = a_{n,j}, \quad j = 0, \ldots, M, \]
\[ (5) \]
can be written as a matrix equation:
\[ \Psi_n U_n = A_n, \]
\[ (6) \]
where \( \Psi_n \) is the following \((M+1) \times (M+1)\) matrix
\[ \Psi_n = \begin{pmatrix}
\Psi_{n,0}(t_{n,0}) & \Psi_{n,2}(t_{n,0}) & \cdots & \Psi_{n,M}(t_{n,0}) \\
\Psi_{n,0}(t_{n,1}) & \Psi_{n,2}(t_{n,1}) & \cdots & \Psi_{n,M}(t_{n,1}) \\
\vdots & \vdots & \ddots & \vdots \\
\Psi_{n,0}(t_{n,M}) & \Psi_{n,2}(t_{n,M}) & \cdots & \Psi_{n,M}(t_{n,M}) \\
\end{pmatrix}, \]
\[ U_n = (u_{n,0}, \ldots, u_{n,M})^T \]
is unknown vector and
\[ A_n = (a_{n,0}, \ldots, a_{n,M})^T \]
is a known vector.

Due to the fundamental theorem of algebra, the determinant \( \det(\Psi_n) \) of \( \Psi_n \) is non-zero which implies that inverse matrix \( \Psi_n^{-1} \) exists, and consequently, solution of (5) is
\[ U_n = \Psi_n^{-1} A_n, \quad n = 1, \ldots, N. \]
\[ (7) \]

Substituting the coefficients found in (6) into (3), we obtain an approximation for the signal function \( u(t) \). Note that, while the parameters \( M \) and \( N \) are arbitrary, in our numerical simulations we get good results in the case \( 1 \leq M - N \leq 2 \).

III. RESULTS AND DISCUSSION

A. Experiment setup

Real-life experiments, as well as, Wi-Fi signals sampling were conducted on level 4 of the Mines Shopping complex located in Selangor, Malaysia. Figure 1 shows the overall map of the test area with five navigation paths (dash lines). The total length of the paths is over 600m. There are 682 reference points, with 1.2m interval Android devices, with custom made signal sampling application were used for data collection. Only 20 APs (out of total 250) were chosen based on signal quality and coverage. In order to obtain a reliable signal database, at each RP we collected 20 signal samples for each detected AP. The average of the 20 measurements is recorded as the final RSSI value for this RP. All collected data along with approximation results are available at navigation.upm.edu.my.

We compared the proposed method (CHEB) with conventional measurement based calibration algorithm (CONV), as well as, with two recent interpolation based methods. The first method is VORO described in [21], and the other one is MOSM presented in [22]. VORO requires two parameters, \( p \) denoting power to the weights and \( Q \) denoting the number of the closest RPs for the refinement. We set \( p = 4 \) and \( Q = 2 \), as recommended by authors in [21]. As for MOSM we fix the number of anchor points required to locate each AP as \( N_{AP} = 5 \). In the case of the proposed model, we set the values of \( N \) and \( M \), such as the \( N \times M \) is equal to the total number of RPs. We evaluated the performance of all methods by gradually reducing the number of reference points.

![Fig. 1 Field-tests area layout](image-url)
B. RSSI approximation accuracy

First, we will analyze the performance of the proposed model in terms of RSSI approximation accuracy. In this experiment, we intend to recover initially measured 682 reference values with less number of measurements.

Fig. 2 compares the power spectrum of measured and approximated signals with the different number of RPs. Results show that with half of the measurements ($RP=341$) our method can recover the original signal almost identically. Relatively accurate results are achieved when the number of RPs are four ($RP=170$) and eight ($RP=85$) times less than the original data.

Fig. 3 shows the cumulative distribution function (CDF) of RSSI approximation error. Here we reduced the number of RPs up to 15 times ($RP = 45$), and the results show that 70% of the time proposed method can approximate RSSI values with less than 3dB error. Next, we compared the proposed method (CHEB) with existing interpolation methods (VORO, MOSM).

Fig. 4 shows the mean approximation error of three methods with different number of RPs. It can be observed that even with the smallest number of RPs, the proposed method significantly outperforms both, MOSM and VORO.

C. Positioning accuracy

Next, we evaluate the positioning accuracy of the proposed method. We compared the proposed method with conventional (CONV) measurement based method, as well as, with previously mentioned interpolation based algorithms.
Here we adopt the K-nearest neighbor (KNN) algorithm [24] for location estimation, and we fix the parameter $K=4$.

Fig. 5 shows the positioning results. Again the proposed method significantly outperforms the existing algorithms. When using 341 RPs, the mean error was 1.05m for CHEB, 2.45m for MOSM, 2.98m for VORO and 2.84m for CONV. For the worst case scenario with $RP=45$, results were 4.62m for CHEB, 6.74m for MOSM, 10.32m for VORO and 11.20m for CONV.

IV. CONCLUSION

In this research, we proposed a new Wi-Fi fingerprint DB construction method based on signal approximation using Chebyshev wavelets. Compared to conventional, measurement-based algorithm, the proposed method requires only a few reference RSSI samples to be collected, and thus significantly reduces the calibration effort. Also, field test results showed that the proposed method achieves better approximation accuracy than existing interpolation methods, such as VORO and MOSM. The constructed fingerprint database was used for positioning, and the results showed that the positioning accuracy is significantly improved compared to conventional, as well as, state-of-the-art interpolation algorithms. In conclusion, we can state that the proposed approximation method using Chebyshev wavelets can create an accurate RSSI fingerprint database for indoor positioning systems.

ACKNOWLEDGMENT

This research has been supported by National Fundamental Research Grant Scheme FRGS of Malaysia, No.01-01-17-1921FR and Universiti Sains Islam Malaysia Internal Grant, No. 0510000-PPP-FST-15218-00.

REFERENCES


