Article

Indoor Location Sensing with Invariant Wi-Fi Received Signal Strength Fingerprinting

Mohd Nizam Husen and Sukhan Lee *

Intelligent Systems Research Institute, Sungkyunkwan University, Suwon, Gyeonggi-do 440-746, Korea; mnizam@skku.edu

* Correspondence: lsh1@skku.edu; Tel.: +82-31-299-6471

Academic Editor: Leonhard M. Reindl

Received: 15 September 2016; Accepted: 5 November 2016; Published: 11 November 2016

Abstract: A method of location fingerprinting based on the Wi-Fi received signal strength (RSS) in an indoor environment is presented. The method aims to overcome the RSS instability due to varying channel disturbances in time by introducing the concept of invariant RSS statistics. The invariant RSS statistics represent here the RSS distributions collected at individual calibration locations under minimal random spatiotemporal disturbances in time. The invariant RSS statistics thus collected serve as the reference pattern classes for fingerprinting. Fingerprinting is carried out at an unknown location by identifying the reference pattern class that maximally supports the spontaneous RSS sensed from individual Wi-Fi sources. A design guideline is also presented as a rule of thumb for estimating the number of Wi-Fi signal sources required to be available for any given number of calibration locations under a certain level of random spatiotemporal disturbances. Experimental results show that the proposed method not only provides 17% higher success rate than conventional ones but also removes the need for recalibration. Furthermore, the resolution is shown finer by 40% with the execution time more than an order of magnitude faster than the conventional methods. These results are also backed up by theoretical analysis.

Keywords: indoor positioning; Wi-Fi fingerprinting; WLAN indoor localization; smartphone sensor; received signal strength (RSS)

1. Introduction

Indoor location-based services (LBS) have attracted much attention due to their commercial and social values. Applications of these services in the fields of robotics, logistics, medicine, tourism, and entertainment are expected to grow in the coming years. A variety of sensor technologies could be utilized as a medium to determine indoor location. Radio Frequency (RF) signals such as ultrasound [1], Bluetooth [2], radio frequency identification (RFID) [3], infrared [4], ultra-wide band (UWB) [5], ZigBee [6], and Wi-Fi [7] have been used in the past and present of LBS and indoor localization work. Among those mentioned, Wi-Fi signals using received signal strength (RSS) fingerprinting has been considered as one of the most popular indoor positioning solutions due to its low cost [8]. Wi-Fi RSS fingerprinting technique consists of two phases: offline site survey and online pattern matching. During the offline phase, the RSS values of several Wi-Fi access points are collected at different predefined calibration locations to create a fingerprint database named radio map. Then, during the online phase, RSS samples are collected in real-time and matched with RSS patterns pre-stored in radio map to estimate the user's location.

The utilization of Wi-Fi RSS for indoor location fingerprinting has drawn attention as an enabler of various location-based personal services with handheld and wearable communication devices. However, the instability of Wi-Fi RSS incurred by highly mutable channel characteristics, due to random spatiotemporal disturbances, hampers a wide-spread adoption of RSS-based location

fingerprinting to real world applications [8,9]. Conventional methods adopt the multi-class pattern classification framework for location fingerprinting, where the RSS sensed from available Wi-Fi sources is collected at each calibration location as statistical samples to form a pattern class. Fingerprinting is then performed by deriving the decision boundaries that provide minimum-error classification [10]. The problem is that the random spatiotemporal disturbances associated with RSS make class patterns heavily overlapped and time varying, causing significant classification errors while requiring frequent offline recalibration. Increasing the number of available Wi-Fi sources and/or decreasing the number of calibration locations may help reduce the problem somewhat, but at the expense of efficiency and/or fingerprinting resolution. Although Wi-Fi sources ineffective for classification can be identified and removed offline for improved efficiency [11], the problem described above remains fundamentally unchanged.

This paper presents a novel method to cope with the RSS instability, using the concept of “the minimally disturbed invariant RSS statistics”, or, simply, “the invariant RSS statistics”. The invariant RSS statistics introduced here are intended not only for enhancing the performance of classification but also for eliminating the need for offline recalibration without reducing fingerprinting resolution and efficiency. The main purpose is to determine a real-time location of a Wi-Fi sensor device based on the real-time RSS values, evaluating these against the database for the closest pre-defined calibration locations. As such, the proposed method is especially suited for smartphone-based indoor location fingerprinting where only the Wi-Fi sources existing in the environment are to be utilized. We conducted an experimental study and theoretical analysis to support the results. The contributions of this article are summarized below:

1. A new method is introduced by resorting to invariant RSS statistics as the reference in fingerprinting, together with the effective RSS readings chosen as input data, which make the proposed fingerprinting accurate and robust against random spatiotemporal disturbances.
2. The automatic removal of ineffective Wi-Fi signal sources in the process of soliciting effective RSS readings makes the proposed method efficient in fingerprinting with the in-situ reduction in the dimensions of decision space.
3. A proposed design guideline is presented as a rule of thumb for estimating the number of Wi-Fi signal sources required to be available for any given number of calibration locations under a certain level of random spatiotemporal disturbances. This will serve as a key guideline for the benefits of the society who wish to employ invariant RSS-based indoor localization using Wi-Fi fingerprinting.
4. Our method requires no recalibration once the invariant RSS statistics are set initially, unlike conventional methods which require recalibration after a certain period of time. This contributes better localization success rate with stable performance in time.

The rest of the paper is organized as follows: in Section 2, we review existing works in indoor localization and Wi-Fi fingerprinting. Section 3 elaborated thoroughly our proposed method and design guideline. Consequently, the experimental results are presented in Section 4. Then, in Section 5 we present the theoretical analysis to support our experimental results. Finally, Section 6 concludes the paper.

2. Related Work

In radio frequency-based indoor positioning, the RADAR system of Microsoft Research [10] was the first work that made use of the WiFi-based network to produce location fingerprints. In the training phase, an area is divided into a 1×1 m grid where the signal strength measurements of the access points are taken at each intersection. The mean of the signal strengths which have been obtained is recorded to create a radio map to be used in the online phase. In the online phase, when the user looks for its location, the mobile station will detect and record the signal strength from as many access points as possible. Then, the signal strength received will be compared to the radio maps to determine the location of the user. Although the authors assume that the radio map is always stable, there is an

issue in that this is not always the case [8]. Furthermore, by picking the access points' signal strength arbitrarily, it is hard to get the optimum location estimation.

In indoor localization, there are two methods for identifying the location: the deterministic and the probabilistic method [12]. Both methods have its own strengths. Bolliger [13] explored the issue of localization accuracy using multiple deterministic and probabilistic methods based on WLAN fingerprinting to improve the radio map precision in the database. Bolliger has developed an indoor positioning system named Redpin in the implementation of their approach. There is also a method using signal strength ratios between pairs of base stations [14]. They claim that this method is more stable among Wi-Fi clients than absolute signal strength. This approach, however, will increase the computational time. There are also other similar works [15,16] trying to use Wi-Fi RSS in the fingerprinting process.

Recently, research in WiFi-based indoor localization explored its feasibility using smartphones with various techniques [11,17–28]. In choosing the Wi-Fi signals to be used to uniquely classify a location, different authors used different approaches in their methods. However, none of them have yet successfully created a stable location fingerprint database because of the nature of Wi-Fi signals that fluctuates over time depending on environment disturbance.

Martin et al. [20] claim that they are the first to implement localization application using RSS of Wi-Fi signals in a smartphone. The authors apply the deterministic method in their work. They take all signals above -80 dBm and used their mean values in their location fingerprint database formation. A similar approach is also implemented by Shin et al. [21]. They developed an indoor Wi-Fi positioning system for Android-based smartphone and collecting Wi-Fi access points RSS indiscriminately. By this approach, the localization process is suffering from a good accuracy because of instability to the location fingerprints due to signal fluctuations.

Meanwhile, Chen [22] use Received Signal Strength (RSS) relations instead of RSS mean values. They claim that although RSS fluctuates, relations between values are more stable. Their approach is setting up rules at each reference location such as “less than”, “equal to”, or “greater than” between two related signals. Another method [24] proposed a way to improve location estimation by penalizes signals from unstable access points and signifies strong signals compared to weak signals. They criticized the use of Euclidean distance in nearest neighbor approach. However, all these approaches still will lead to the same problem where variation is too high when time varying is included during data collection, which will lead to unavoidable large overlaps in location classification.

Another approach in the selection of Wi-Fi signals is using a *predetermined number* of “first-come-first-get” signals [25]. By setting up a predetermined number (e.g., 8), they will compare the online measured first eight detected signals to the location fingerprint database to estimate the location. The authors apply the probabilistic method in this work. Although this approach will slightly increase the computational time, again it will suffer from instability of accuracy. The work in [19] focusing on orientation-based indoor localization using smartphones, where they managed to localize the smartphone by location and orientation. However, there is still rooms for improvements in the calibration phase to avoid inaccurate detection.

The key to formulating a stable location fingerprint is selecting the best appropriate signals. Samih Eisa et al. [11] has made an effort in this matter by removing useless signals. However, their approach still lacks in producing highly stable location fingerprints because the author's definition and identification methods of useless signals are not enough to select the best signals to form a good location fingerprints.

All the above mentioned past related researches have issues with random spatiotemporal disturbances and do not provide any design guideline to implement their approach. This paper proposes an accurate and robust Wi-Fi RSS fingerprinting against random spatiotemporal disturbances based on the concept of invariant RSS [28] and a design guideline that relates statistically among the number of Wi-Fi sources, the number of calibration locations, the level of disturbance, and the success rate.

3. Methodology

This section will elaborate the Wi-Fi received signal strength (RSS) model and the invariant RSS method, followed by the design guideline of implementing invariant RSS indoor localization.

3.1. Invariant Wi-Fi RSS Method

Figure 1 illustrates the overview of the processes in the proposed invariant RSS method for indoor localization using a smartphone. The proposed method consists of two phases. It starts with the training phase, where the essential *invariant* reference pattern classes are formed based on invariant RSS statistics. Then, the localization phase is carried out to estimate the user's location based on real-time Wi-Fi RSS sensing.

Figure 1. Proposed method of the smartphone-based indoor location fingerprinting based on invariant Wi-Fi received signal strength.

The m dimensional RSS vector, $\{s_i(t)\}_j$, at time t at a particular calibration location j due to m Wi-Fi sources, $i = 1, \dots, m$, is represented as:

$$\{s_i(t)\}_j = \{s_{1,j}(t), s_{2,j}(t), \dots, s_{i,j}(t), \dots, s_{m,j}(t)\}, j = 1, \dots, n \quad (1)$$

where the bold-face letter, s , is used to represent it as a random variable.

Here, we model $s_{i,j}(t)$, the RSS from the i th Wi-Fi source at the j th calibration location, as:

$$s_{i,j}(t) = \alpha_{i,j}(t) \times r_{i,j} + \delta_{i,j} \quad (2)$$

where $r_{i,j}$ represents the time-invariant RSS with no spatiotemporal disturbances present, $\alpha_{i,j}(t)$ the multiplicative signal alteration factor to account for the spatiotemporal disturbances of $r_{i,j}$, and $\delta_{i,j}$ the sensor noise. Note that we introduce $r_{i,j}$ as the ideal time-invariant signal attenuated by the distance from the signal source i to the location j through the invariant channel characteristics, taking only fixed building infrastructure and furniture layout into consideration.

The actual signal, $s_{i,j}(t)$, is then considered as the alteration of $r_{i,j}$ by $\alpha_{i,j}(t)$, $0 < \alpha_{i,j} \leq 1$, reflecting the stochastic channel characteristics due to randomly moving people and/or objects as well as randomly orienting smartphone users disturbing the channel during RSS measurement.

Let us define the invariant RSS, $\hat{s}_{i,j}(t)$, as $s_{i,j}(t)$ with minimal random spatiotemporal disturbances, i.e., $\alpha_{i,j} \approx 1$. Then, from Equation (2):

$$\hat{s}_{i,j}(t) \approx r_{i,j} + \delta_{i,j} \quad (3)$$

Since the randomness of $\hat{s}_{i,j}(t)$ comes mostly from sensor noise, $\delta_{i,j}$, multiple measurements of $\hat{s}_{i,j}(t)$ form a cluster around $r_{i,j}$, the statistics of which is governed by that of $\delta_{i,j}$. Assuming that the statistics of $\delta_{i,j}$ are time-invariant, so is $\hat{s}_{i,j}(t)$. For convenience, the statistics of $\hat{s}_{i,j}(t)$ are represented as a random variable, $s_{i,j}$. Then, at the calibration location j with m available Wi-Fi sources, an m dimensional vector of invariant RSS statistics, $\{s_i\}_j$, $i = 1, \dots, m$, is defined. For the pattern classification point of view, $\{s_i\}_j$ represents the j th reference pattern class in the m dimensional space of Wi-Fi sources. The small blue ellipses in Figure 2 schematically represent three invariant reference pattern classes, $\{s_i\}_j$, $j = 1, 2, 3$, in two-dimensional Wi-Fi source space $i = 1, 2$. Notice that the red ellipses surrounding the blue ellipses represent spatiotemporally varying RSS statistics.

Figure 2. Schematic representation of three reference pattern classes in 2-dimensional Wi-Fi source space that illustrates the difference between the distribution of invariant reference pattern classes (small blue ellipses) and the spatiotemporally varying RSSs (large red ellipses).

Fingerprinting of the RSS vector, $\{s_i(t)\}$, captured at an unknown location is done by identifying the reference pattern class among $\{s_i\}_j$, $j = 1, \dots, n$, that maximally support $\{s_i(t)\}$. The support of $\{s_i(t)\}$ by the reference pattern class j is defined here as the number of $s_i(t)$ that fall in $s_{i,j}$ for $i = 1, \dots, m$, or the sum of the likelihood probabilities of $s_i(t)$ to belong to $s_{i,j}$, for $i = 1, \dots, m$. Note that, in counting the number or summing the likelihood probabilities, only those $s_i(t)$ having a sufficiently high level of statistical confidence as a member of $s_{i,j}$ are allowed to participate in the sum, while excluding others. That is, $s_i(t) \in s_{i,j}$, for $i = 1, \dots, m$, when the following equation is satisfied:

$$|s_i(t) - \mu(s_{i,j})| < \xi \sigma(s_{i,j}) \quad (4)$$

where $\mu(s_{i,j})$ and $\sigma(s_{i,j})$ are the mean and the standard deviation of $s_{i,j}$.

Fingerprinting is done based on the pattern classification rule as shown in Algorithm 1. Notice that, in the algorithm, when two or more calibration locations come up with the same support values, the decision is delayed and signals are recaptured until only a single calibration location produces the maximum support.

Algorithm 1. Pseudocode of the localization phase in estimating the user's location based on real-time spontaneous sensed Wi-Fi RSS vector.

Input: Spontaneous RSS Vector Measurement $\{s_i(t)\}$

Output: Estimated Location $L \in \{1, \dots, n\}$

```

1:  $m \leftarrow$  Number of Wi-Fi Sources
2:  $n \leftarrow$  Number of Calibration Locations
3:  $s_{i,j} \leftarrow$  A component of  $\{s_i\}_j$ , representing Invariant RSS Pattern for the  $i$ th Wi-Fi source at Calibration Location  $j$ 
4:  $\sigma(s_{i,j}) \leftarrow$  Standard Deviation of  $s_{i,j}$ 
5:  $\xi\sigma(s_{i,j}) \leftarrow$  Decision Margin
6: for  $j = 1$  to  $n$  do
7:   for  $i = 1$  to  $m$  do
8:     if  $s_i(t) \in s_{i,j}$ , i.e.,  $|s_i(t) - \text{Mean of } s_{i,j}| < \xi\sigma(s_{i,j})$  then
9:       Either  $\text{sum}(j) \leftarrow \text{sum}(j) + 1$ 
          or  $\text{sum}(j) \leftarrow \text{sum}(j) + \text{Pr}(s_i(t) | s_{i,j})$ 
10:    end if
11:  end for
12:  if  $\text{sum}(j) > \text{Maximum}$  (default:  $\text{Maximum} = 0.0$ ) then
13:     $\text{Maximum} = \text{sum}(j)$ 
14:     $L \leftarrow j$ 
15:  else if  $\text{sum}(j) = \text{Maximum}$  then
16:     $L \leftarrow 0$ 
17:  end if
18: end for
19: if  $L = 0$  then
20:   Reject and Recapture Signals
21: end if

```

The proposed method provides better performance over conventional ones because we use the concept of invariant RSS statistics in the formulation of invariant reference pattern classes which produce low variance in avoiding RSS instability. Furthermore, during the localization phase, only the invariant RSS, i.e., the $s_i(t)$ having a sufficiently high level of statistical confidence as a member of $s_{i,j}$, are selected to participate in classification in comparison to the conventional methods that takes all available RSS to participate in the classification. This will cause the proposed method deliver better performance with the low variance, thus, preventing RSS instability effects as encountered by the conventional methods. The other advantage of our proposed method is that by a proper choice of decision margin, a minimal false negative can be achieved with the expense of rejections. The slight rise in rejections is handled by repeating location queries which ultimately produces true positives. Experiments demonstrating the superior performance of our proposed method are detailed in Section 4.

3.2. Design Guideline of Invariant RSS Based Wi-Fi Fingerprinting

This section describes the formulation of a design guideline for the benefits of the society who wish to employ invariant RSS-based indoor localization using Wi-Fi fingerprinting in any indoor environment. The guideline represents how the relations between the numbers of Wi-Fi signals obtained with a certain number of calibration locations may contribute to the class/location separation degree that will classify and separate each calibrated location, at many different levels of random spatiotemporal disturbances. This useful guideline is produced based on statistical analysis. In formulating the design guideline, we develop a simulation algorithm. The simulation is developed

in order for us to control the level of random spatiotemporal disturbance and the number of Wi-Fi sources. In real life, the spatiotemporal disturbance is random and difficult to set and control, while the obtainable number of Wi-Fi sources is depending on the surrounding environment. Hence, simulation is the best solution in order to formulate the design guideline.

3.2.1. Design Guideline Development Procedure

Figure 3 depicts the steps that have been carried out in formulating the design guideline. In the procedure, we simulate the invariant reference RSS propagations, spontaneous RSS propagations, identify the effective invariant RSS after applying spatiotemporal disturbances, and compute the class separation degree of the calibrated reference locations.

Figure 3. Procedure to formulate design guideline.

• Step 1: Invariant Reference RSS Generation

In collecting the invariant reference RSSs at each calibration location j , invariant reference RSS is defined by $\{s_i(t)\}_j$ with $\alpha_{i,j} \approx 1$. We generate the signal propagation values based on the $\hat{s}_{i,j}(t)$ as defined in Equation (3). The value of $\hat{s}_{i,j}(t)$ is based on $r_{i,j}$, where $r_{i,j}$ is the ideal time-invariant signal attenuated by the distance from the Wi-Fi signal source i to the location j through the invariant channel characteristics, taking only fixed building infrastructure and furniture layout into consideration. In the invariant RSS values generation, we set the following number of Wi-Fi signal sources m and the number of calibration locations n as defined in Table 1 to the same setup area.

Table 1. Setup Parameter of the Invariant Reference RSS Generation.

No. of Wi-Fi Signal Sources	No. of Calibration Locations
20	7
50	
80	
20	15
50	
80	
20	20
50	
80	

As an example, when we simulate the area with seven calibration locations and 20 Wi-Fi sources as illustrated in Figure 4, we generate all 140 invariant reference RSSs. To get a statistical datapoint, we

generate the reference RSS values repeatedly for 300 times. Then, the *mean* and *standard deviation* of each RSS at each calibrated location are calculated.

Figure 4. The visual representation of setup area to generate invariant reference RSS with 20 Wi-Fi signal sources $i = 1, 2, \dots, m$; and seven calibration locations $j = 1, 2, \dots, n$; ($m = 20, n = 7$).

- **Step 2: Invariant Reference Pattern Classes Formulation**

After the invariant reference RSS are collected, the invariant RSS statistics, $\{s_i\}_j$, as mentioned in Figure 1, are set and the invariant reference pattern classes are defined. At each calibration location j , with all available Wi-Fi signal sources, $i = 1, 2, \dots, m$, will produce an m dimensional invariant RSS statistics vector, $\{s_i\}_j$. From the invariant RSS statistics, the reference pattern classes are formulated and stored in a database.

- **Step 3: Spontaneous RSS Generation**

In order to simulate the spontaneous RSS gained at an unknown location during location query in a real-life situation, we developed another simulation module to act out the situation. The spontaneous RSS is the signals added by temporally varying signal disturbance, $\alpha_{i,j}$, due to the moving people and objects that may disrupt the channel and/or the orientation of the smartphone user disturbing or blocking the channel at the time of Wi-Fi sensor measure the RSS. The effect of $\alpha_{i,j}$ on spontaneous RSS is represented by a multiplication coefficient to invariant reference RSS by executing Equation (2). The explanation of each parameter in the equation are already mentioned in Section 3.1. The level of $\alpha_{i,j}$ values is selected in 0.1 intervals. At each level of $\alpha_{i,j}$, the associated α to each Wi-Fi source is distributed in the uniform random distribution of sigmoid function as:

$$f_1(\alpha) = \frac{1}{1 + e^{-\alpha}}, \quad 0 < \alpha < 1 \quad (5)$$

The reason to apply sigmoid function on α rather than just binary values or exponential function is to reflect the real life behavior of the disturbance towards Wi-Fi signals. Based on the experimental data analysis, it is observed that the effect of random spatiotemporal disturbance over RSS is modeled closest to sigmoid function in comparison to the others.

A sigmoid function values based on the number of Wi-Fi signal sources m , the number of calibration locations n , and the temporally varying signal disturbance level α as:

$$f_2(\alpha) = f_1(\alpha) mn = \frac{1}{1 + e^{-\alpha}} mn, \quad 0 < \alpha < 1 \quad (6)$$

Then, we generate uniform random numbers and assign the numbers from left to the right of the sigmoid curve values. This represents the varying disturbances for each Wi-Fi RSS. Finally, the invariant reference RSS values of each Wi-Fi sources which already gained from the previous

process is multiplied to the respective temporally varying signal disturbance by executing Equation (2). The outcome of the computation is the spontaneous RSS gathered at each calibrated location.

• Step 4: Effective Invariant RSS Selection

The selection of the effective signals to be used in fingerprinting classification and the removal of the ineffectual signals to discriminate and uniquely identify each calibrated location is the key factor for the success of a Wi-Fi fingerprinting method. The process is based on the condition in line 8 of the pseudocode as depicted in Algorithm 1, where the effective invariant RSS is the individual spontaneous RSS within the decision margin. In the simulation, we defined the effective invariant RSS as the α values being uniform randomly selected in the range between 0.9 and 1.0. The threshold value of 0.9 is determined based on the analysis of experimental data. From the data analysis, it was observed that in classifying an RSS as invariant or not, 0.9 produces the best result in comparison to 0.8 and 0.7. Hence, in defining the invariant and non-invariant RSS from the number of Wi-Fi signal sources m and the number of calibration locations n is:

$$g(m, n) = \begin{cases} 1, & f_2(\alpha) \geq 0.9 \\ 0, & f_2(\alpha) < 0.9 \end{cases} \quad (7)$$

• Step 5: Class Separation Degree Computation

Class separation degree refers to the level of discrimination power between pattern classes. It is calculated as the following process:

$$\frac{N}{\sum_{k=1}^{n-1} k} \quad (8)$$

where N is the number of calibration locations that could be discriminated or separated by the intensity of the effective invariant RSS, n is the number of calibration locations, and the denominator is the total number of combination pairs.

In calibration locations $n = 7$, there are 21 possible combination pairs. So, to compute the class separation degree, we compute how many combination pairs out of the 21 that are successfully separate the calibration locations. The same method is used in 15 and 20 calibration locations setup where there are 105 and 190 pairs, respectively.

Figure 5. An example of the visualization in computing the class separation degree at ($n = 7, m = 20$). This example showing the computation of the class separation degree at different levels of disturbance: $\alpha = 0.1, 0.2, 0.3, 0.4, 0.5$, and 0.7 .

Figure 5 depicts the conception example of the class separation degree calculation process at $n = 7$, $m = 20$. There are 21 possible combinations pairs in seven calibration locations which represented by 21 (white) cells as shown in the figure. At a different level of random spatiotemporal disturbance α , the class separation degree is computed. The numbers presented in each cell represent the number of Wi-Fi signals that can discriminate the two intersects locations. Blank cell denotes no signals may discriminate both locations. Notice that the class separation degree starts to be 100% when $\alpha_{i,j} = 0.4$, and remains 100% when $\alpha_{i,j}$ is incremented. Although it remains 100% as it is incremented, the numbers of Wi-Fi signals that can discriminate/separate each calibrated location are also getting higher. This shows that higher values of $\alpha_{i,j}$ will give better location separation options.

3.2.2. Proposed Design Guideline

After implementing the processes as mentioned in the previous subsection, we present the result as depicted in Figure 6 which significantly beneficial as a guideline to employ indoor localization using invariant RSS-based location fingerprinting. Based on this result, we conclude that the higher the number of Wi-Fi sources m , the higher the class separation degree. It means more obtainable signals at any particular location are better to uniquely identify that location. When the number of calibrated locations n is increased, it will lower the class separation degree as proved in the graph where the pattern shifted to the right. This is particularly happening at a lower number of Wi-Fi signal sources $m = 20$. The result also shows that at $m = 80$, the similar class separation degree is gained when $n = 7$, $n = 15$, and $n = 20$. That means when the accessible Wi-Fi sources m is 80, it is enough to discriminate all the maximum 20 calibration locations without deteriorating the class separation degree.

Figure 6. Design guideline for the number of Wi-Fi signal sources (m) and level of disturbance (α) for a given number of calibration locations (n) to achieve certain class separation degree.

As a recommendation to employ indoor localization using invariant RSS-based location fingerprints, an optimum number of obtainable Wi-Fi signal sources is above 50 when the number of calibration location n is 20 to get a good class separation degree (above 90%). A lower number of m is tolerable as the number of n is gradually decreased.

4. Evaluation and Results

Prior to the theoretical investigation on the implication of the proposed method, an experimental evaluation of the performance of proposed method is conducted in comparison with such well-known conventional methods as RADAR [10] and *Removal of Useless Wi-Fi Sources* [11].

4.1. Experimental Setup

The experiment is set up in a 400 m² area on a floor of a building. There are 25 Wi-Fi signals available for fingerprinting with 20 calibration locations assigned based on 4 m × 4 m as the initial resolution. The resolution will be increased and decreased during the comparative investigation of different calibration location resolutions (resolution between 1 m × 1 m and 6 m × 6 m). The calibration locations are the point that has been used to collect the readings of the Wi-Fi RSS data. Figure 7a depicts the floor map of the experimental area with all 20 calibration locations (Loc. 1 to Loc. 20). Each calibration location is within equal distance from each other based on the specified resolution and it is being marked on the floor for easy identification. In the training phase, 100 invariant RSS data are collected at each calibration location to form invariant reference class patterns. To collect invariant RSS data with minimal random spatiotemporal disturbances, the data are collected after midnight when no random disturbance such as human is present. Furthermore, a smartphone is placed on a pinnacle of the head of a remotely controlled robot to avoid any blocking effect of the Wi-Fi sensor. The smartphone collects Wi-Fi RSS at a particular calibration location and the RSS data is sent to a server, and the readings are tagged to the location associated with it. Figure 7b illustrates the robot that has been used during data collection. However, during the testing stage in the localization phase, the smartphone could either be held by a human or attached to a robot to obtain a real spontaneous RSS measurement. Meanwhile, for performance comparison, the conventional non-invariant reference class patterns are formed with 100 spontaneous RSS data collected at each calibration location in normal office hours when people may be present in rooms and corridors. A Samsung Galaxy S4 smartphone with Android version 4.4.2 is used in the experiment.

Figure 7. (a) The initial experiment floor map area with marked calibration locations. Later, the number of calibration locations are increased and decreased according to its resolution; (b) The robot attached with a smartphone Wi-Fi sensor used in sensing invariant Wi-Fi RSS data collection. The collected data are sent through the network and stored in a server.

4.2. Comparison of Success Rate and Its Temporal Variation

The first experiment is to compare the proposed method with the well-known conventional methods: *RADAR* [10] and *Removal of Useless Wi-Fi Sources* [11], in terms of the success rate at each calibration location and the possible variation of the success rate in time. The latter aims at investigating whether there is any degradation of performance with respect to the lapse of time, dictating recalibration. Success rate here is defined as the percentage of the total number of counts with correct recognition decision during testing divided by the number of total tests. The correct recognition decision is the precisely detected location as predicted by the localization system. To this end, the experiment has been carried out over the period of 18 weeks, where, each week, 100 localization experiments are conducted for four different orientations at each calibration location. Note that, for the sake of fair comparison, the same reference calibration locations of 4 m × 4 m resolution, as well as the same normal office hour environments, are adopted in the experiments.

The result based on 18 weeks of comparative testing is illustrated in Figure 8. The figure shows that the proposed invariant RSS-based method performs superior to conventional methods from the start of testing with the 93% success rate in comparison to the respective 76% and 71% by RADAR [10] and *Removal of Useless Wi-Fi Sources* [11]. The superior performance of the proposed method becomes even more evident as testing continues. For instance, after 18 weeks, the performance of conventional methods is degraded by 14% and 16% respectively, while that of the proposed method remains virtually constant. Note that, in between the week of 12 and 13, the reference pattern classes for the conventional methods are recalibrated to see how much it improves their success rate. It is found that the recalibration enables their success rate back to 77% and 71% respectively, similar to that of their first week performance, as expected. After the recalibration, it is also observed that the conventional methods return back to a pattern of performance deterioration.

Figure 8. The success rate of the proposed method (green) in comparison with that of the conventional methods (red and blue), where the temporal variation over 18 weeks span is clearly shown. It is verified that the recalibration applied to the conventional methods puts their success rate back to the initial one.

4.3. Comparison of Success Rate at Different Resolution of Calibration Locations

To have a comparative evaluation on the possible resolution of fingerprinting, the success rate of the proposed method at six different resolutions of calibration locations, ranging from 1 m to 6 m with 1 m interval, is compared with that of the conventional methods. The number of the calibration locations are increased and decreased when the resolution is higher and lower, respectively. The result is shown in Figure 9, where the proposed method is shown to achieve 90% success rate at 3.7-m resolution while the conventional methods fail to achieve 90% success rate even at 6-m resolution.

4.4. Comparison of Performance from Samples Collected over Different Length of Time

In the preceding experiments, the number of samples used for the formation of invariant reference pattern classes of our proposed method is $n = 100$ for each calibration location, which was taken over a period of two hours for all calibration locations of $4\text{ m} \times 4\text{ m}$ resolution. The same number of samples is used for existing methods (RADAR and *Remove Useless APs*) for the formation of their location fingerprints. Our next investigation is to apply an increased number of samples from a longer period of time with the same resolution, which will be applied to all three approaches for performance comparison. The goal of this investigation is to check whether time length of collecting samples in the formation of reference pattern classes affecting the success rate.

Figure 9. A comparative evaluation of the possible resolution of fingerprinting, where the proposed method (green) is able to achieve 90% success rate at 3.7-m resolution while the conventional methods (red and blue) fail even at 6-m resolution.

We continue our investigation by increasing the number of samples n of each calibration location to $n = 200$ which was dispersedly collected in a one week time period. Then, we re-apply our approach by defining the invariant RSS statistics and forming the invariant reference pattern classes. To evaluate the performance comparison, we also apply the same number of samples from the same longer period of time in the formation of location fingerprints of existing approaches. We then repeat the same procedure for $n = 300$ which was collected in two weeks' time period.

Figure 10. Performance comparison with a different number of samples collected over different time length. The result shows no degradation in our proposed method as the time length is increased in comparison to the conventional methods.

Figure 10 illustrates the comparison results. We observed that our invariant RSS approach produced the highest success rate from all tested approaches and it is stable regardless of the number of samples over a different length of time used. On the contrary, the conventional approaches suffer from success rate degradation as the number of samples increased over time. From our analysis, the reason behind this result is our approach only takes invariant RSS which maintain low variance although as time increased, as opposed to existing conventional approaches that take all available RSS where the variance is increased as the time is extended.

4.5. Computational Complexity Analysis

The computational complexity of Wi-Fi based fingerprinting is, in general, a function of the number of Wi-Fi sources m and of calibration locations n . Here, we represent the computational complexity as $O(\alpha m \times n)$, $\alpha \leq 1$ with α representing the percentage of the Wi-Fi signals received that are selected to participate in classification. For instance, [11] excludes weak Wi-Fi signals from classification such that α becomes around 0.8. On the other hand, the proposed method excludes non-invariant Wi-Fi signals such that α becomes around 0.3, as shown in Figure 11. Note that, recent work by Sánchez-Rodríguez et al. [27] shows that a computational complexity of $O(1)$ is feasible in terms of floating point computation based on multiple weighted decision trees although the decision trees may grow with $O(\alpha m \times n)$.

Figure 11. Comparison of average execution time for different methods: RADAR with $\alpha = 1$, Remove Useless APs with $\alpha = 0.8$, and Invariant RSS proposed with $\alpha = 0.3$.

4.6. Application Model Implementation

An actual implementation of the proposed method has been employed to localized human in a human-robot interaction case. Figure 12 illustrates the information flow of the smartphone application based on the proposed method's algorithm for initial human-robot interaction. The smartphone user makes a location query through the application, the application will then capture spontaneous RSSs through the Wi-Fi sensor, check for effective invariant RSS, and then estimate the location. The user estimated location, the user request to the robot (if any), and the user identification will then sent to a robot for further action.

Figure 12. Smartphone's indoor localization application information flow for initial human-robot interaction.

5. Theoretical Analysis

To understand the theoretical implication of the proposed method, a simulation study is conducted. To this end, the spontaneous RSS distributions at individual calibration locations are simulated for each Wi-Fi source same as a Gaussian distribution with their variances same as σ . Furthermore, the means of a pair of spontaneous RSS distributions are assumed separated by $\lambda\sigma$, $\lambda > 0$, as illustrated in Figure 13 with the class separability of $\mu_1 - \mu_2 = 3.0\sigma$.

Figure 13. Schematic representation of different class distributions enlightening class separability, decision margins, correct decision probability, erroneous decision probability, and class boundary.

Meanwhile, the invariant RSS distributions at individual calibration locations are modeled as having their means same as those of corresponding spontaneous RSS distributions but their variances represented indirectly by the decision margin, $\gamma\sigma$ (Note: $\gamma\sigma = \xi\sigma(s_{i,j})$ in Algorithm 1), $\gamma < \lambda$, around their means, as shown in Figure 13. Note that the above simulation model used for RSS distributions is not realistic but intended to show the performance of proposed method under harsh conditions. However, in our experiments, the sensor noise $\delta_{i,j}$ is found to be bounded by 3 dBm independent of calibration locations. Therefore, the decision margin $\xi\sigma(s_{i,j})$ in Algorithm 1 or $\gamma\sigma$ in Figure 13 is set as 3 dBm in real experiments.

Recall that, in our proposed method, the location query returns the calibration location q , if q satisfies:

$$q = \text{Max}_j \{ \text{Cardinal} [s_{i,j}(t) \in \xi\sigma(s_{i,j}), i = 1, \dots, m], j = 1, \dots, n \} \quad (9)$$

The location query, Equation (9), is simulated by applying Bernoulli trial to the above RSS model as follows: First, let the probabilities of a Wi-Fi source fall within the decision margins of the respective right and wrong calibration locations be P_1 and P_2 (refer to Figure 13). Then, the probability that l and m Wi-Fi sources fall within the right and wrong calibration locations, respectively, can be computed by:

$$\sum_m \binom{n}{l+m} (P_1 + P_2)^{l+m} (1 - P_1 - P_2)^{n-l-m} \binom{l+m}{l} Q^l (1-Q)^m \quad (10)$$

where $Q = P_1 / (P_1 + P_2)$. Note that $l < m$, $l > m$, and $l = m$ implies respectively a false, a success, and a rejection in recognition. Figure 14 illustrates the false and rejection probabilities in terms of various decision margins and class separabilities. False probability is defined as the probability of the wrong detection is made, while rejection probability is defined as the probability of no decision is made because there is not enough information to make a decision. The figure shows that, given a class separability, the minimal false rate required can be achieved with a proper choice of decision margin, although the rejection rate goes up. The increased rejection rate can be handled by a suitable number of query repetitions. In practice, it suffices at most several query repetitions to reach a minimal false rate of decision.

Figure 14. Rejection and False probability comparison for class separability between 1.5σ and 3.0σ .

Next, Figure 15 depicts the precision-recall curve to show the effect of decision margin, representing the invariant region in our proposed method, on various class separabilities for the number of Wi-Fi sources given as 25. It represents how decreasing decision margin from the class boundary, the end of each precision-recall line, heightens the precision while lowering the recall. With the invariant region that decreases decision margin, the proposed classifier heightens the precision so that false positives seldom occur. Note that, in our classifier, resulting a decrease in recall increases rejections instead of increasing false negatives, such that repeating location queries ultimately produces true positives.

Figure 15. Precision-Recall curve of the proposed approach that shows the classifier performance with 25 Wi-Fi sources.

Figure 16 shows how the precision-recall curve changes due to the increase in the number of Wi-Fi sources from 25 to 50. It clearly shows the dramatic improvement of the precision-recall curve with the increase in the number of Wi-Fi sources. Figures 15 and 16 also indicates that, by properly choosing decision margin or invariant region and the number of Wi-Fi sources, we can set the precision and the recall or the rejection rate optimal for the given class separabilities.

Figure 16. Precision-Recall curve changes due to increasing in the number of Wi-Fi sources from 25 to 50. The arrows demonstrate the respective changes from 25 to 50 Wi-Fi sources at different class separability.

6. Conclusions

As demonstrated by the experiment, the proposed method yields more than 17% higher success rate and more than 40% improvement in the resolution of fingerprinting (3-m vs. 5-m at 85% success rate), compared to conventional methods [10,11]. Furthermore, the proposed method provides solid stability of performance in time, contrary to conventional methods experiencing performance degradation in time. Besides the better performance in success rate, accuracy, and stability, the proposed method also has lower execution time, where the average execution time is more than 10 times faster than the conventional methods. The proposed method is expected to find many smartphone-based indoor location fingerprinting applications.

Acknowledgments: This research was supported, in part, by the National Research Foundation (NRF) of Korea (NRF-2013M1A3A3A02042335), funded by the Korean Ministry of Science, ICT and Planning (MSIP), in part, by the Korea Evaluation Institute of Industrial Technology (KEIT) (2015-10060160), funded by the Korean Ministry of Trade, Industry and Energy (MOTIE), and, in part, by the Technology Innovation Program (or Industrial Strategic Technology Development Program, report number: 10048320), funded by MOTIE.

Author Contributions: Mohd Nizam Husen designed and performed the experiments, implemented and tested the proposed prototype; Sukhan Lee proposed the original idea and performed the theoretical analysis; Sukhan Lee and Mohd Nizam Husen analyzed the data, wrote and revised the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Medina, C.; Segura, J.C.; De la Torre, A. Ultrasound Indoor Positioning System Based on a Low-Power Wireless Sensor Network Providing Sub-Centimeter Accuracy. *Sensors* **2013**, *13*, 3501–3526. [[CrossRef](#)] [[PubMed](#)]
2. Zhuang, Y.; Yang, J.; Li, Y.; Qi, L.; El-Sheimy, N. Smartphone-Based Indoor Localization with Bluetooth Low Energy Beacons. *Sensors* **2016**, *16*. [[CrossRef](#)] [[PubMed](#)]
3. Ni, L.M.; Liu, Y.; Lau, Y.C.; Patil, A.P. LANDMARC: Indoor Location Sensing Using Active RFID. *Wirel. Netw.* **2004**, *10*, 701–710. [[CrossRef](#)]
4. Lee, S.; Ha, K.N.; Lee, K.C. A Pyroelectric Infrared Sensor-based Indoor Location-Aware System for the Smart Home. *IEEE T. Consum. Electr.* **2006**, *52*, 1311–1317. [[CrossRef](#)]
5. Gezici, S.; Tian, Z.; Giannakis, G.B.; Kobayashi, H.; Molisch, A.F.; Poor, H.V.; Sahinoglu, Z. Localization via Ultra-Wideband Radios: A Look at Positioning Aspects for Future Sensor Networks. *IEEE Signal Proc. Mag.* **2005**, *22*, 70–84. [[CrossRef](#)]
6. Zhao, Y.; Dong, L.; Wang, J.; Hu, B.; Fu, Y. Implementing Indoor Positioning System via ZigBee Devices. In Proceedings of the 42nd Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 26–29 October 2008; pp. 1867–1871.
7. Deng, Z.; Wang, G.; Qin, D.; Na, Z.; Cui, Y.; Chen, J. Continuous Indoor Positioning Fusing WiFi, Smartphone Sensors and Landmarks. *Sensors* **2016**, *16*, 1427. [[CrossRef](#)] [[PubMed](#)]
8. Curran, K.; Furey, E.; Lunney, T.; Santos, J.; Woods, D.; McCaughey, A. An Evaluation of Indoor Location Determination Technologies. *J. Locat. Based Serv.* **2011**, *5*, 61–78. [[CrossRef](#)]
9. Kaemarungsi, K.; Krishnamurthy, P. Analysis of WLAN's Received Signal Strength Indication for Indoor Location Fingerprinting. *Pervasive Mob. Comput.* **2012**, *8*, 292–316. [[CrossRef](#)]
10. Bahl, P.; Padmanabhan, V.N. RADAR: An In-building RF-based User Location and Tracking System. In Proceedings of the IEEE 19th International Conference on Computer Communications, Tel Aviv, Israel, 26–30 March 2000; pp. 775–784.
11. Eisa, S.; Peixoto, J.; Meneses, F.; Moreira, A. Removing Useless APs and Fingerprints from WiFi Indoor Positioning Radio Maps. In Proceedings of the IEEE 4th International Conference on Indoor Positioning and Indoor Navigation, Belfort-Montbéliard, France, 28–31 October 2013; pp. 1–7.
12. Dawes, B.; Chin, K.-W. A Comparison of Deterministic and Probabilistic Methods for Indoor Localization. *J. Syst. Softw.* **2011**, *84*, 442–451. [[CrossRef](#)]
13. Bolliger, P. Robust Indoor Positioning through Adaptive Collaborative Labeling of Location Fingerprints. Ph.D. Thesis, ETH Zurich, Zurich, Switzerland, February 2011.

14. Kjærgaard, M.B. Indoor Location Fingerprinting with Heterogeneous Clients. *Pervasive Mob. Comput.* **2011**, *7*, 31–43. [[CrossRef](#)]
15. Narzullaev, A.; Park, Y.; Yoo, K.; Yu, J. A Fast and Accurate Calibration Algorithm for Real-Time Locating Systems Based on the Received Signal Strength Indication. *AEU Int. J. Electron. Commun.* **2011**, *65*, 305–311. [[CrossRef](#)]
16. Milioris, D.; Tzagkarakis, G.; Papakonstantinou, A.; Papadopouli, M.; Tsakalides, P. Low-Dimensional Signal-Strength Fingerprint-Based Positioning in Wireless LANs. *Ad Hoc Netw.* **2014**, *12*, 100–114. [[CrossRef](#)]
17. Narzullaev, A.; Park, Y. Novel Calibration Algorithm for Received Signal Strength Based Indoor Real-Time Locating Systems. *AEU Int. J. Electron. Commun.* **2013**, *67*, 637–644. [[CrossRef](#)]
18. Gutierrez, N.; Belmonte, C.; Hanvey, J.; Espejo, R.; Dong, Z. Indoor Localization for Mobile Devices. In Proceedings of the 11th IEEE International Conference on Networking, Sensing and Control, Miami, FL, USA, 7–9 April 2014; pp. 173–178.
19. Husen, M.N.; Lee, S. Indoor Human Localization with Orientation Using Wi-Fi Fingerprinting. In Proceedings of the ACM 8th International Conference on Ubiquitous Information Management and Communication, Siem Reap, Cambodia, 9–11 January 2014.
20. Martin, E.; Vinyals, O.; Friedland, G.; Bujcsy, R. Precise Indoor Localization using Smart Phones. In Proceedings of the ACM Multimedia, Firenze, Italy, 25–29 October 2010; pp. 787–790.
21. Shin, B.-J.; Lee, K.-W.; Choi, S.-H.; Kim, J.-Y.; Lee, W.J.; Kim, H.S. Indoor Wifi Positioning System for Android-Based Smartphone. In Proceedings of the International Conference on Information and Communication Technology Convergence, Jeju Island, Korea, 17–19 November 2010.
22. Chen, Q. A Rule-Based Approach to Indoor Localization Based on Wifi Signal Strengths. Ph.D. Thesis, Hong Kong University of Science and Technology, Hong Kong, China, August 2012.
23. Luo, Y.; Hoeber, O.; Chen, Y. Enhancing Wi-Fi Fingerprinting for Indoor Positioning Using Human-Centric Collaborative Feedback. *Hum. Centric Comput. Inf. Sci.* **2013**, *3*. [[CrossRef](#)]
24. So, J.; Lee, J.-Y.; Yoon, C.-H.; Park, H. An Improved Location Estimation Method for Wifi Fingerprint-Based Indoor Localization. *Int. J. Softw. Eng. Appl.* **2013**, *7*, 77–86.
25. Chan, E.C.L.; Baci, G.; Mak, S.C. Orientation-based Wi-Fi Positioning on the Google Nexus One. In Proceedings of the 6th IEEE International Conference on Wireless and Mobile Computing, Networking and Communications, Niagara Falls, ON, Canada, 11–13 October 2010; pp. 392–397.
26. Husen, M.N.; Lee, S. Indoor Location Wi-Fi Fingerprinting Using Invariant Received Signal Strength. In Proceedings of the International Conference on Engineering Technology and Entrepreneurship, Kuala Lumpur, Malaysia, 16–18 November 2015.
27. Sánchez-Rodríguez, D.; Hernández-Morera, P.; Quinteiro, J.M.; Alonso-González, I. A Low Complexity System Based on Multiple Weighted Decision Trees for Indoor Localization. *Sensors* **2015**, *15*, 14809–14829. [[CrossRef](#)] [[PubMed](#)]
28. Husen, M.N.; Lee, S. High Performance Indoor Location Wi-Fi Fingerprinting Using Invariant Received Signal Strength. In Proceedings of the ACM 10th International Conference on Ubiquitous Information Management and Communication, Danang, Vietnam, 4–6 January 2016.