Journal of Ubiquitous Systems & Pervasive Networks
Volume 17, No. 2 (2022) pp. 51-58

New and Reliable Points Shifting - Based Algorithm for Indoor Location Services

Tarek El Salti [a]*, Nelson Shaw [a], Joseph Chun-Chung Cheung [b], Edward R. Sykes [c]

[a]*School of Applied Computing, Sheridan College, Oakville, Ontario, Canada, L6H 2L1*
[b]*TELUS Communication Inc., 200 Consilium Place, Scarborough, Ontario, Canada, M1H 3E4*
[c]*Centre for Mobile Innovation, Sheridan College, Oakville, Ontario, Canada, L6H 2L1*

Abstract

Indoor localization is of great importance to several fields such as healthcare and asset tracking. However, many factors (e.g., multipath propagations) impact the quality of signals which are used to perform localizations. As a consequence, the precision and accuracy of the computed locations are heavily influenced. Therefore, the methodologies to compute indoor locations always need continuous refinements in terms of those metrics including the time complexity. For the last metric, It impacts the performance of mobile devices due to their limited resources. To address these challenges, a new set of fingerprinting algorithms was presented in this paper called *Fingerprinting Line-Based Nearest Neighbour*. This set shifts grid points potentially towards targets via a deterministic percentage. The running time of the set is upper bounded. Moreover, this paper presents the following: 1) an upper bound in terms of distance errors for the proposed algorithms, and 2) based on real experiments, the new algorithms (e.g., 90% shifting) improved the accuracy and precision, and had lower distance errors probabilities compared to those for the nearest neighbour-based algorithms (e.g., by 106% and 76%, respectively).

Keywords: *Indoor Location Services, Fingerprinting, Wi-Fi, Path Loss exponent, K-Nearest Neighbor*

1. Introduction

Recently, many researchers have been attracted by the massive deployments of the Internet of Things (IoT) [1–3]. Among the applications of IoT, location-based applications are significant as they improve the quality of life for fields such as healthcare and security [4]. Examples of these location-based services are health care applications (e.g., locating patients with Alzheimer in hospitals [4]).

Traditionally, technologies such as GPS are used to determine the location of objects (e.g., automobiles and individuals) in outdoor environments. However, due to the complexity of the indoor environments (e.g., walls and ceilings), these technologies are not sufficient to localize an object (e.g., a medical equipment in hospitals) in indoor environments [5]. In response to this limitation, Wi-Fi, Bluetooth, Ultra-Wide Band, RFID, and Zigbee technologies have been implemented [6, 7]. However, Wi-Fi indoor localization system is common due to its low infrastructure cost and high reliability for indoor environments [7, 8].

1.1. Motivation and Problem Background

Different techniques based on Wi-Fi signal's characteristics were used to estimate the locations of objects. These techniques include the Time of Arrival (ToA), Time Difference of Arrival (TDoA), Angle of Arrival (AoA), Channel State Information (CSI) and the Receive Signal Strength Indicators (RSSIs) [9–11]. However, TDoA and AoA require more devices [12]. In addition, non-RSSI-based techniques such as ToA, TDoA, and AoA have reduced performance in Non-Line of Sight (NLOS) conditions [12]. Finally, while CSI offers higher performance compared to RSSI, CSI can only be obtained through specific hardware and is not practical for smart phones and other mobile devices

*Corresponding author. Tel.: +1 905 845 9430 Ext. 2950
Fax: +1 905 815 4035; E-mail: tarek.elsalti@sheridancollege.ca

DOI: 10.5383/JUSPN.17.02.001

[11]. Compared to other techniques, RSSIs can be implemented without changes to existing wireless infrastructure [13]. However, RSSI-based techniques are also susceptible to multipath fading and noise [14]. In order to effectively deal with the noisy RSSI measurements, Kalman's filter has been implemented [15].

The indoor localization techniques that use RSSIs include ranging and fingerprinting [6, 16]. Ranging is a technique which involves trilateration. However, trilateration is not suitable for indoor localization due to the high complexity of indoor environments [17]. Another approach, known as position fingerprinting, is used to improve the indoor localization of a device and functions more effectively within indoor environments that are more spatially diverse [6]. Although fingerprinting requires additional setup (i.e., offline stage), compared to non-RSSI techniques (e.g., AoA), fingerprinting is not as susceptible to angle multipath [10, 18].

The metrics used to evaluate the indoor localization capabilities are accuracy, precision, and time complexity. For the first metric, accuracy is the distance error between the location estimated and the actual location. Accuracy is computed using the Euclidean Distance formula in 2-D. Another metric to consider is precision, which refers to the distribution of the estimated locations and the relative distances between them. Precision is calculated by taking the standard deviation of the accuracy. Finally, time complexity refers to the run time complexity of an algorithm. The Big O Notation, which is a common metric that computes time complexity, must be considered for fingerprinting algorithms because of the limited resources of mobile devices such as the lifetime battery. There are certain indoor positioning systems that achieve high accuracy but have high computational complexity and are not suited for mobile devices indoors [23].

One of the first approaches to take advantage of fingerprinting is the RADAR system [19]. However, the average accuracy of the system is considered low for indoor localization [19] [20]. Another related work done by Zhang et al. [21] implements new Nearest Neighbour (NN)-based algorithms and considers the use of the Path Loss exponents (PLe) as features for the grid points. In other words, the authors replaced the RSSI fingerprints by the PLe features. This is due to the fact that the RSSI measurements are considered unstable due to noise [21]. However, the time complexities of these algorithms were not considered [17].

A more recent work proposed by Hoang *et al.* [22] designs and implements an extension of the KNN to address both the time complexity and accuracy. While the work demonstrates that the algorithm achieved a high accuracy, the precision of the algorithm was not considered. Therefore, based on the above challenges, there is a need to improve and consider accuracy, precision, and time complexity for indoor localization. This articles focuses on the NN-based approach.

1.2. Contributions Overview

In this paper, an extended set of the K-Nearest Neighbour (K-NN) and the Nearest Neighbour version 3 (NNv3) [17, 19] referred to as the Fingerprinting Line-Based Nearest Neighbor (FLBNN) algorithms was proposed and developed [24]. The new set is based on the shifting capability of the grid points to improve the accuracy and precision. For the time complexity, the run time of the new set is upper bounded. In addition, the new algorithms were proven to have a theoretical upper bound for the distance errors.

Our real-world experiments demonstrated that the new set of algorithms (e.g., 90% shifting percentage) achieved significant accuracies compared to those for the Nearest Neighbour (NN) algorithms such as the Soft-Range-Limited KNN (SRL-KNN) and the Nearest Neighbor version 2 (NNv2) [17, 22]. The new set also achieved higher precision compared to those for the same NN-based algorithms such as the SRL-KNN and the Nearest Neighbor version 4 (NNv4). Lastly, the new set of algorithms achieved higher probabilities of being more accurate compared to those for the same existing NN-based algorithms. This was measured based on the Cumulative Distribution Function (CDF) for a set of distances. All these experiments were based on the Cisco's best practices regarding the deployment of Access Points (APs).

The remainder of this article is mentioned as follows. Section 2 provides a detailed background which spans some preliminaries and literature review. Section 3 proposes the new fingerprinting-based methodology. Section 4 evaluates and analyzes the proposed and existing algorithms. Section 5 discusses the findings from the previous section. Lastly, Section 6 provides some concluding remarks and recommends several future directions.

2. Background

2.1. Assumptions

The following assumptions were used in this research for an interior space: The grid points were deployed at a rectangular area based on the Cisco's best practices for indoor localization [25]; a centroid was calculated based on its surrounding four grid points and it was also considered a grid point; the Cartesian coordinates were in 2-D [26]; a mobile phone was referred to as a target node; the IEEE 8.2.11ac was the communication protocol used between the target node and its surrounding Access Points (APs) (i.e., anchor nodes); Cisco routers were used to collect the RSSI measurements as they provide more granular data regarding RSSI measurements than other types of routers [27]; the traditional Kalman filter was used because the distribution of the noise follows Gaussian distribution [28]; linear functions were used as collected RSSI measurements did not change over time for each grid point during the offline stage [28]; the time complexity of the new algorithm was only analyzed for the online stage; the distances computed between the target node and the APs and between the grid points and APs were based on the Path Loss Exponent (PLe) [21]. These exponents were used as features instead of the RSSI measurements as these measurements are considered unstable [21].

2.2. Fingerprinting Stages

The fingerprinting technique consists of offline and online stages. The offline phase involves the collection of signal information at a known grid point in the indoor environment. RSSI measurements are collected at each grid point to construct a radio map. Each grid point has a set of RSSI values $RSSI_j$ = ($RSSI_1$, $RSSI_2$, … , $RSSI_n$) where j is the AP index and n is the number of RSSI measurements collected at each grid point during the offline phase. The online stage involves the calculation of RSSI measurements in real-time and comparing them to the offline stage RSSI measurements at each grid point [29]. Although the focus of this work was on RSSI-based algorithms, several other techniques for indoor localization including non-RSSI-based techniques were explored in order to cover the breadth of the localization area.

Table 1. Fingerprinting-Based Algorithms

	Accuracy	Precision	Time Complexity
Nearest Neighbour [17, 19, 32]	2.94 m - 4 m	≈ 2 m	O(g * m)
K-Nearest Neighbour [17, 19, 20, 32]	2.13 m - 3.25 m	≈ 1.3 m	O(g * m)
Weighted K-Nearest Neighbour [20, 30]	1.66 m - 3.06 m	X	O(r * a * log(k) + d)
Weighted Fusion [17] [30]	1.54 m	X	X
Path-Loss-Based Fingerprint Localization [21]	3.5 m - 4 m	< 5 m	X
Dual-Scanned Fingerprint Localization [21]	≈ 3 m	< 5 m	X
Nearest Neighbour version 2 [17]	≈ 3 m	≈ 1.3 m	O(t * m)
Nearest Neighbour version 3 [17]	≈ 3.75 m	≈ 1.9 m	O(g * m)
Nearest Neighbour version 4 [17]	≈ 2.75 m	≈ 0.9 m	O(t * m)
Soft-Range-Limited K-Nearest Neighbour [22]	≈ 0.60 m	X	O(g * m)

X:Unknown

2.3. Nearest Neighbour - Based Algorithms

Kalman's filter and Particle filters are used to filter noisy RSSI measurements. Particle filters are extended Kalman's filters which approximates the nonlinear functions to linear functions using Taylor Series [13]. Fingerprinting can take advantage of the Kalman's filter or Particle filter for effectively dealing with multipath fading [14].

One of the first approaches to take advantage of fingerprinting is the RADAR system, which uses the Nearest Neighbor algorithm (NN), the K-Nearest Neighbor algorithm (K-NN), and Wi-Fi signals in order to estimate the indoor location of a user [19]. However, the average accuracy of the system was approximately 3 m, which is not ideal for accurately estimating indoor locations [19, 20, 22]. The NN's accuracy depends on the number of grid points that construct a radio map. The algorithm also does not consider the areas between the grid points and instead chooses a grid point that is the localized position. This leads to imprecise and inaccurate localized positions. For K-NN, the accuracy of the algorithm is influenced by the fact that noisy RSSIs in offline and/or online stages may not allow the algorithm to choose consecutive K nearest grid points. As a consequence, high errors in distances between computed centroids and target nodes (i.e., inaccuracies) may exist.

Another version of the K-NN algorithm referred to as Weighted KNN (WKNN) is based on Statistical Learning Theory [20, 22]. The algorithm calculates the weighted average of RSSI-based distances between a set of selected grid points and the current location [20, 22]. This approach improved the accuracy from 3.12 m to 3.06 m, however, the accuracy is still relatively low [20]. In addition, the accuracy of the approach was not statistically verified [20]. Lastly, this algorithm inherits K-NN's drawback (i.e., nonconsecutive K nearest grid points). To improve the accuracy metric, a new weighted fusion fingerprinting algorithm was proposed [30]. With the proposed algorithm, the accuracy was improved to approximately 1.5 m [30]. However, the design of the algorithm does not consider the precision factor.

Another work by Zhang et al. [21] implements the Path-Loss-Based Fingerprint Localization algorithm (PFL) and Dual-Scanned Fingerprint Localization algorithm (DFL) in order to improve the accuracy and precision of indoor localization [21]. However, the time complexities of these algorithms were not considered [17, 21]. Furthermore, PFL is impacted by outliers (i.e., RSSIs) and DFLs' thresholds may direct the algorithm to inaccurate areas. To examine time complexity and improve accuracy, the work done by El Salti et al. [17] proposes the NNv2, NNv3, and NNv4 algorithms [17]. However, the accuracy of indoor localization can be further improved, as well as the complexity of NNv4.

A more recent work discusses the use of a Soft-Range-Limited KNN algorithm (SRL-KNN) in order to address both the time complexity and accuracy [22]. The conducted experiment demonstrates that the SRL-KNN algorithm achieved a higher accuracy compared to that for K-NN with the same time complexity [22]. While the work demonstrates that the algorithm achieved a high accuracy, the precision of the algorithm was not considered as part of the design. For the time complexity of most of the mentioned algorithms, it was not also considered. Lastly, and to the best of our knowledge, the mentioned algorithms did not consider the best practices regarding the deployment of the anchor nodes. The goal of these best practices is to construct an efficient indoor location system.

Table 1 compares the mentioned algorithms in terms of the following factors: 1) accuracy, 2) precision, and 3) time complexity. As shown in the table, the time complexity of NN, KNN, and SRL-KNN is O(g * m), where the term g refers to the number of grid points and the term m refers to the number of RSSI measurements collected during the offline stage (i.e., this number of RSSI measurements depends on the number of APs) [17] [19] [22] [32]. The WKNN runs in O(r * a * log(k) + d). The term r and a refer to the number of observations, and the number of anchor nodes, respectively. For the term d and k, they refer to the number of physical dimensions, and the number of nearest neighbours, respectively. For the accuracy and precision measurements in the table, they may change depending on the designed environments. Based on these challenges, we proposed an algorithm that improves and considers the accuracy, precision, and time complexity for indoor localization to effectively estimate the location of a target indoors. For the testing real environment, the Cisco's best practices were followed for the deployment of APs for indoor localization [25].

3. Methodology

The Fingerprinting Line-Based Nearest Neighbor (FLBNN) algorithm, proposed in our work, is an extension of the K-Nearest Neighbor (K-NN) and the Nearest Neighbor version 3 (NNv3) algorithms [17, 19]. In order to implement the FLBNN algorithm, an offline stage needs to be constructed. The offline stage involves the collection of the Receive Signal Strength Indicator (RSSI) measurements at each grid point in the indoor environment. The number of these RSSIs depends on the number of deployed Access Points (APs). Afterwards, each of these RSSIs were replace by the Path Loss exponents (PLe). This establishes a radio map.

At the online phase, the FLBNN algorithm was executed. Algorithm 1 presents the FLBNN algorithm in the online stage. From Lines 1 to 10, the FLBNN algorithm searches for the k-nearest neighbor (e.g., first, second, third and fourth nearest neighbors).

Algorithm 1: Fingerprinting Line-Based Nearest Neighbor

Input : G - A graph of y Grid Points with d PLe Values (PLe1, PLe2, PLe3..., PLed), T - Target with d PLe values (PLe1, PLe2, PLe3..., PLed), and dp - The displacement percentage value between 0 and 100.
Output: Q - the estimated location coordinates.

```
1  foreach i of y do
2      foreach j of d do
3          dist[j] = calculateDifference(T[j], G[i][j]);
4      end foreach
5      totalDistG[i] = calculateTotalEuclideanDistance(G[i], dist);
6  end foreach
7  sortByShortestDistance(totalDistG);
8  for k = 1 to 4 do
9      chosenGridPt[k] = totalDistG[k];
10 end for
11 S = calculateCentroid(chosenGridPt);
12 M = calculateMidpoint(chosenGridPt[0], chosenGridPt[1]);
13 N = shiftPoint(chosenGridPt[0], S, M, dp);
14 Q = calculateCentroid(N);
15 Return Q;
```

Line 9 searches for the nearest neighbor Z to the target node T. The nearest neighbor Z is added to the list of chosen neighbors. Fig. 1(a) presents the Target point T located within some grid points. For simplicity, the nearest grid points are G1, G2, G3, and G4. However, in some cases, the points may not be necessarily consecutive such as G2, G7, G10, and G19.

In Line 11, the algorithm computes the Centroid S from the k-nearest neighbor. Afterward, in Line 12, the algorithm then chooses the two closest grid points (i.e., first and second k-nearest neighbor) to the Target point T (see Fig. 1(a), where $G1$ and $G2$ are the closest and second closest grid points to T respectively) and calculates the Midpoint M between these grid points. From Line 13, the algorithm displaces one or more points (i.e., $G1$, M, or S) based on a determined displacement percentage (dp) (see Fig. 1(b)). In Fig. 1(b), the GridPoint $G1$ and Midpoint M are displaced by 50%. The positions of the displaced point(s) and the remaining point(s) are all stored in List N. Lastly, in Line 14, the Centroid Q is computed and this centroid is considered the localized position.

The time complexity of FLBNN is O($y * d$) where y refers to the number of grid points and d refers to the number of Received Signal Strength Indicators (RSSIs) collected at the online stage. The FLBNN has the following property:

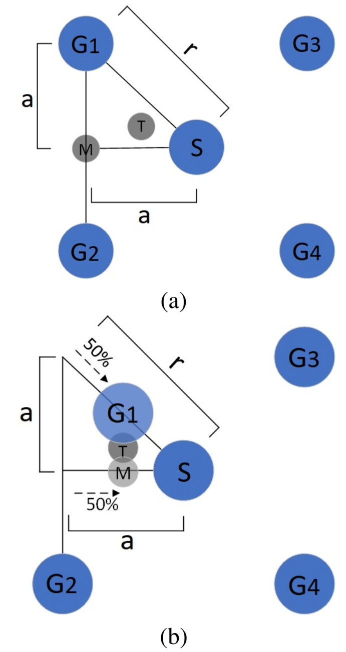

Fig. 1. (a) Construction of the Triangle; (b) The Shifting of G_1 and M by 50%.

Theorem 3..1. *Given a set of grid points that constructs a square of grid points that cover a square area (A), the upper bound for the distance errors of FLBNN (C) $\leq \sqrt{2}a$, where the term a refers to one of the sides of the triangle constructed by FLBNN and a $\neq$ r (hypotenuse of this triangle) .*

Proof. Assume that the square area consists of the following set of Grid Points $(G_p) = \{G_1, G_2, ..., G_n\}$, where the term n refers to the total number of grid points. As for one case, the number of grid points ($G^{'}$) per row (r_w) is even (i.e., $G^{'}$ = $2l$, where the term $l \in Z^{+}$). The worst case for the distance errors for the FLBNN algorithm is when the centroid (S) is computed from the Set $(S_e) = \left\{G_1, G_{2l}, G_{n-(2l-1)}, G_n\right\}$. Based on FLBNN, $G_i \in S_e$ is the nearest neighbour. Therefore, the Target node T is within the Triangle $\Delta G_i M S$ (see Fig. 1(a)). As part of the worst case, T is assumed to be located at one side of the hypotenuse (r). Furthermore, G_i (i.e., the closet grid point to T), M, and S are all located on the other side of r (i.e., the shifting percentage (sp) = 100% for M and for either S or G_i). Assume in Fig. 1(b), the node T is located at the original position of G_1 and M and G_1 are located at the position of S. The localized position $Q = (x_S + x_M + x_{G_i})/3, (y_S + y_M + y_{G_i})/3)$. Since $x_S = x_M = x_{G_i}$ and $y_S = y_M = y_{G_i}$, then $(\frac{3x_M}{3}, \frac{3y_M}{3}) = (x_M, y_M)$. In addition, $r = \sqrt{a^2 + a^2}$. Therefore, $r = \sqrt{2}a$. As for the other case, a similar proof applies when $G^{'}$ per r_w is odd (i.e., $G^{'}$ = $2l + 1$, where the term $l \in Z^{+}$). The worst case for the distance errors for the FLBNN algorithm is when the term S is computed from the Set $(S_e) = \{G_1, G_{2l+1}, G_{n-2l}, G_n\}$. □

There are a number of differences between FLBNN and K-NN and NNv3. First, once the midpoint is calculated between the nearest Grid Point G_i, Centroid S, and Midpoint M, the FLBNN algorithm has the ability to displace one or two of these points towards either G_i, S, or M by a specific percentage. For example, the algorithm can displace G_i towards S or the algorithm can displace M and S towards G_i. If Grid Point G_i is displaced by 100% towards Centroid S, Grid Point G_i is directly shifted to the same coordinates as Centroid S and Centroid Q is computed. Another example is that Grid Point G_i can be shifted halfway towards Midpoint M when the displacement percentage is equal to 50% and then Centroid Q is computed. Second, FLBNN computes two centroids (i.e., Centroid S and Centroid Q) where K-NN and NNv3 compute one and three centroids; respectively. Third, the distance errors of FLBNN have a theoretical upper bound of $(C) \leq \sqrt{2}a$, where the term a refers to one of the sides of the triangle constructed by FLBNN and $a \neq r$ (hypotenuse of this triangle).

4. Findings (Analysis and Evaluation)

4.1. Experimental Design Model

The real-world design was used to demonstrate that the Fingerprinting Line-Based Nearest Neighbor (FLBNN) algorithm is considered precise and accurate. The design was based on a Google Pixel phone (i.e., target device) with four Cisco Wireless Access Points (APs) from the set of all deployed APs at Sheridan College (i.e., this is part of Sheridan College's Wi-Fi system). The phone received 5 GHz Wi-Fi signals from the APs. The design was also based on a 2.5 m $\times$ 12.5 m area with twelve grid points and five centroids in Room S144 at the S-Building. The distance between each pair of neighbouring grid points was 2.5 m. The distance between each pair of neighbouring centroids was also 2.5 m. Each AP was distanced at least 22 feet or 6.71 meters away from each other and their locations were unrevealed for confidentiality reasons. The design of the experiments was based on the Cisco's best practices for AP spacing for indoor localization [25].

There is a monotonic relationship between the distance (i.e., from the AP to the target) and the Receive Signal Strength Indicators (RSSIs) [25]. Therefore, based on our experiments, each grid point needs to be distanced 2.5 m away from each of the adjacent grid points in order to preserve the monotonic behavior of the RSSIs and create different fingerprints (i.e., RSSIs) for each of the grid points. In addition, the grid point space should be designed within a convex hull [25]. The convex hull is formed by the APs around the perimeter of the hull. Additionally, each AP should be distanced at 4.27 m away from the grid point space [25]. When the target is less than 4.27 m away from an AP, short range propagation anomalies occur [25]. Therefore, the monotonicity of the RSSI and the distance relationship degrades, and thus, the RSSIs cannot be used to accurately predict the distance from the Target node to an AP.

During the offline stage, the target device was stationed at each grid point or centroid and collected thirty-five RSSI measurements filtered via Kalman's filter. The collection of those measurements was done during daytime on September 7th, 2019. Moreover, the duration for the collection of the thirty-five RSSI measurements at each grid point for each AP was approximately five minutes. These thirty-five RSSI measurements for each of the four APs were used to calculate a single Path Loss exponent (PLe) value. Therefore, each grid point (or centroid) was fingerprinted with four PLe values [21]. The time to construct the entire radio map was approximately three hours (i.e., this time also included the time to move the phone between a grid point and another point and to log this information). Notice that the use of PLe values was done to deal with the challenge regarding the fluctuations of RSSIs introduced by different factors (e.g., multipath fading).

In the online stage, four testing locations (1.937, 3.5), (2.437, 9.5), (2.937, 6.5), and (3.437, 12) were chosen randomly in Room S144 where thirty-five filtered RSSI measurements were collected for each AP for the duration of five minutes and the total time for all these randomized positions was approximately thirty minutes. This time also includes the time to move the phone from one testing point to another point and also the log time for the RSSIs. Based on these measurements, the accuracy and precision for the following algorithms were evaluated: 1) K-Nearest Neighbor (KNN) [17], 2) Weighted KNN (WKNN) [20], 3) NN version 2 (NNv2) [17], 4) NN version 3 (NNv3) [17], 5) NN version 4 (NNv4) [17], 6) Path-Loss-Based Fingerprint Localization (PFL) [21], 7) Dual-Scanned Fingerprint Localization (DFL) [21] and, 8) Soft-Range-Limited KNN (SRL-KNN) [22].

4.2. Performance Metrics and Statistical Analysis

Each of these algorithms and our proposed algorithm were compared in terms of their accuracy and precision. During the online stage, each algorithm ran thirty-five times for every test point, and consequently, the metric values for each run were averaged. For the accuracy metric, the distance error between the estimated location and the actual location were computed. Moreover, precision refers to the distribution of the estimated locations and the relative distances between them. Regarding the cumulative probability distribution (i.e., Cumulative Distribution Function (CDF)), it represents the probabilities of locating a target within a range of distance values (e.g., the probability that the estimated location is within 1.5 m to 2 m of the actual target). The normal distribution of the data was verified via Shapiro–Wilk test with a 95% confidence level. Afterwards, Mann–Whitney U test was used to verify the statistical significance between the proposed and existing algorithms.

4.3. Accuracy and Precision Analysis

The existing fingerprinting-based algorithms were first compared to FLBNN in terms of the accuracy metric (see Fig. 2(a)). FLBNN with a displacement percentage of 10% - 90% achieved the highest accuracy compared to those for the existing fingerprinting algorithms. The K-NN, WKNN, NNv3, and NNv4 algorithms achieved similar accuracies and were considered the second highest in terms the accuracy compared to those for DFL, PFL, SRL-KNN, and NNv2 algorithms. In addition, the DFL algorithm achieved the lowest accuracy.

Fig. 3 supports the observations obtained in Fig. 2(a) in terms of the cumulative probability distribution for the accuracies. Fig. 3 demonstrates that the probability to achieve accuracy within one meter was significantly higher for the set of FLBNN compared to those for the other existing algorithms. The probability to achieve accuracy within 2.5 m for most algorithms reached 100%.

From Fig. 2(b), the FLBNN algorithm was the most precise algorithm when the shifting percentage for the grid point was specified between the range 30% and 40%. The precision was lowered when the shifting percentage was close to 90%. The second most precise algorithm was the NNv3 algorithm which was more precise compared to those for its other versions (i.e., NNv2 and

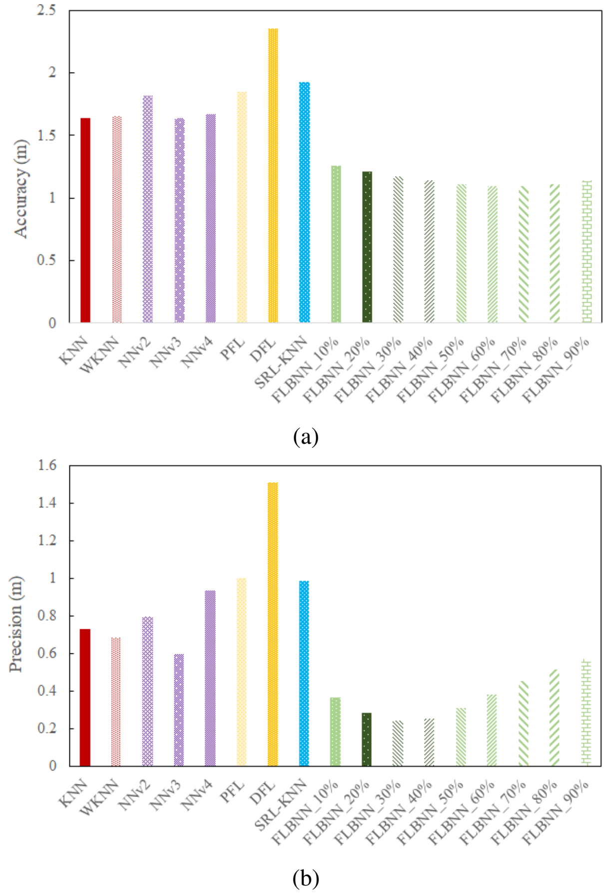

Fig. 2. (a) The Localization Algorithms in Terms of Accuracy; (b) The Localization Algorithms in Terms of Precision.

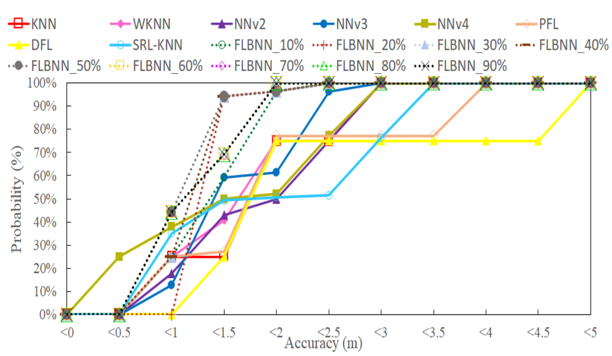

Fig. 3. Cumulative Probabilities in Terms of Accuracies for the Proposed and Existing Algorithms.

NNv4). The same applies to SRL-KNN and K-NN where the variant of K-NN (i.e., WKNN) achieved higher precision. For the DFL algorithm, it achieved the worst precision.

5. Discussion

A possible explanation for the behaviour in terms of the accuracy for FLBNN and some of the existing algorithms is mentioned as follows. The new algorithm shifts the closest grid point and the midpoint possibly closer to the target node. The assumption made here is that the target is not located at a grid point, centroid, or midpoint. The algorithm chooses four grid points that are the closest to the target and calculates the Centroid S. The algorithm then chooses the two closest grid points among the four selected grid points. In this case, the algorithm shifts the nearest one of the two grid points and the Midpoint M closer to Centroid S, where S is closer to the target than both the grid points and the midpoint. The target is generally closer to Centroid S than the midpoint or nearest grid point. As a result, the Midpoint M and the nearest grid point were shifted to move them closer to the target node. Therefore, the calculated Centroid Q in these experiments was closer to the target, and thus, the accuracy of the new algorithm was improved.

For the PFL and NNv2 algorithms, they only choose the grid points but they do not consider the areas between the grid points. For K-NN algorithm and its variants SRL-KNN and WKNN algorithms, they consider a general approximation of the grid points (i.e., centroids) as their main functionality. Lastly, NNv3 and NNv4 do not consider the size of the triangles constructed by the additional centroids and the grid points.

A possible explanation for the precision behaviour of FLBNN and some of the existing NN-based algorithms is mentioned as follows. The new algorithm calculates the estimated locations within a particular area of the space (i.e., triangle). The shifting of the grid points especially when the shifting is close to 40% constructs a small triangle where the localized points are within this small area. Hence, the distribution of those points in this experiment was more concentrated which lead to high precision.

The NNv3 algorithm achieved a higher precision compared to those for the existing algorithms because NNv3 is not influenced as much by Receive Signal Strength Indicators (RSSI) fluctuations since the algorithm chooses the four closest grid points instead of initially the closest centroid or closest grid point. Finally, SRL-KNN achieves a lower precision compared to that for WKNN because SRL-KNN chooses its four closest grid points based on a threshold surrounding the previous localized point calculated. Thus, a change in RSSI measurements significantly impacts the change in the four closest grid points.

This article considered several significant design factors for fingerprinting-based Localization to improve the accuracy and precision for indoor location services. The centroid method is one factor that is considered a coarse localization approach. However, with the integration of fingerprinting which is another factor, the coarse localization becomes fine localization. The integrated behaviours support the literature in terms of the integration between fingerprinting and geometrical methods to improve the performance of localization approaches [31]. In addition, the construction of confined areas (i.e., triangles) further supports the idea of the use of triangles in localization to improve the localization performance [17]. However, this paper enriches the literature with the points shifting to reduce the area of the constructed triangles. As a result, the distance errors and the size of their distributions are further reduced within the triangles. The implementation of the Path Loss exponent (PLe) further improve the results as the RSSIs are considered unreliable [21]. Because of only one square of four grid points and a triangle were considered, we focused on the implementation of a graph-based model instead of a tree-based model.

A drawback was captured when the experiments were conducted. Because of the adoption of the centroid idea, the chosen grid points may not be necessarily consecutive due to some similarities in the RSSI measurements across the grid points. As a consequence, some calculated centroids may have possibly long distance errors. However, the shifting of points still reduce the distance errors and the size of their distributions as there is a possibility that the computed locations are further shifted towards the target's location.

6. Conclusion

6.1. Summary

The quality of life for several fields such as security and health care is significantly improved by location-based services [4]. Among the several applications that benefit from these services, the localization of seniors is of great importance. Especially, seniors who suffer from Dementia may have difficulty in remembering how to navigate within a facility. Therefore, localization enables doctors and nurses to assist those seniors [33].

Several geometrical localization techniques [9] [10] are considered insufficient for accurate and precise indoor localization [10]. In order to solve these issues, fingerprinting-based techniques [22] are not as susceptible to angle multipath and do not need additional hardware. However, the performance of these techniques are always need to be improved.

This article demonstrates that the shifting capability of the points significantly improves the accuracy and precision for indoor localization. This goal was achieved by a novel set of fingerprinting-based algorithms for indoor localization. The new set of algorithms is referred to as the Fingerprinting Line-Based Nearest Neighbour (FLBNN) and it has the ability to shift grid points to compute centroids close to target nodes.

An analysis of the algorithms revealed that they have a theoretical upper bound for their distance errors. In addition, the real-world results based on Sheridan College's Wi-Fi system demonstrated the following: 1) the FLBNN algorithms outperformed some existing Nearest Neighbour (NN)-based algorithms [17, 20–22] in terms of the accuracy and precision (e.g., by 68% for the accuracy and by 64% for the precision), and 2) the new algorithms had lower probabilities in terms of the distance errors compared to those for the same studied NN algorithms. Lastly, FLBNN ran in O($y * d$) (where, y refers to the number of grid points and d refers to the number of Receive Signal Strength Indicator (RSSI) measurements collected at the online stage. Therefore, the proposed algorithms are considered reliable for indoor location services. When mobility is considered for target nodes, the reliability is still possibly achieved due to the points shifting capability of the new algorithm.

6.2. Limitations

The first limitation of this study is that the deployed grid points space was relatively small and only few centroids existed within the space. However, the grid points space was designed in such a way to follow the Cisco's best practices for Access Point (AP) spacing, where the space should ideally be designed within the convex boundary formed by all four APs [25]. In addition, If more grid points were deployed beyond the specified number of grid points, the similarities between the RSSI features will increase and they will not be useful anymore.

The second limitation is that only four test points were used during the experiments. The third limitation is that there was a small set of APs that were considered in the experiments as the other surrounding APs did not follow the Cisco's best practises. The fourth limitation is that we did not explore the accuracy and precision when varying Wi-Fi environment settings such as channel selection, radio frequency selection, and physical ambience as those APs were only configured by Sheridan College's network architect. Finally, we only experimented with Wi-Fi technologies and did not explore with other technologies such as Bluetooth and Zigbee.

6.3. Future Work

The limitations of the conducted experiments indicate several recommendations as follows. One recommendation is to work with a larger grid points space to further explore the performance of the proposed and existing algorithms. Another recommendation is to consider different Wi-Fi environment settings (e.g., channel selection). This will also include the deployment of Bluetooth beacons. Lastly, new variations of the FLBNN algorithms will need to be considered. Especially, the integration of the new set of algorithms with Machine Learning algorithms (e.g., Support Vector Machine (SVM)) to determine the best shifting displacements of the points.

Acknowledgements

We would like to thank TELUS and the Natural Sciences and Engineering Research Council of Canada (NSERC) for their generous funding. In addition, we extend our thanks to Mark Orlando and Simon Hood for their previous contributions.

References

[1] V. Patela, D. Shaha, and N. Doshia, "Emerging Technologies and Applications for Smart Cities," *Journal of Ubiquitous Systems & Pervasive Networks*, vol. **15**, no. 2, pp. 19-24, December 2021.

[2] K. Harichaa, A. Khiata, Y. Issaouia, A. Bahnasseb, and H. Ouajjia, "Towards smart manufacturing: Implementation and benefits," *Journal of Ubiquitous Systems & Pervasive Networks*, vol. **15**, no. 2, pp. 25-31, December 2021.

[3] F. Xhafa, "Special issue "Internet of Things (IoT)-based services"," *Computing*, vol. **101**, no. 7, pp. 725–727, July 2019.

[4] F. Zafari, A. Gkelias, and K. K. Leung, "A survey of indoor localization systems and technologies," *IEEE Communications Surveys & Tutorials*, vol. **21**, no. 3, pp. 2568–2599, April 2019.

[5] F. Dalkılıç, U. C. Çabuk, E. Arıkan, and A. Gürkan, "An analysis of the positioning accuracy of ibeacon technology in indoor environments," in *International Conference on Computer Science and Engineering (UBMK)*, Antalya, Turkey, Oct. 2017, pp. 549–553.

[6] X. Zhou, T. Chen, D. Guo, X. Teng, and B. Yuan, "From one to crowd: a survey on crowdsourcing-based wireless indoor localization," *Frontiers of Computer Science*, vol. **12**, no. 3, pp. 423–450, June 2018.

[7] S. Xia, Y. Liu, G. Yuan, M. Zhu, and Z. Wang, "Indoor fingerprint positioning based on wi-fi: An overview," *ISPRS International Journal of Geo-Information*, vol. **6**, no. 5, p. 135, April 2017.

[8] T. Xie, H. Jiang, X. Zhao, and C. Zhang, "A wi-fi-based wireless indoor position sensing system with multipath interference mitigation," *Sensors*, vol. **19**, no. 18, p. 3983, September 2019.

[9] C. Laoudias, A. Moreira, S. Kim, S. Lee, L. Wirola, and C. Fischione, "A survey of enabling technologies for network localization, tracking, and navigation," *IEEE Communications Surveys & Tutorials*, vol. **20**, no. 4, pp. 3607–3644, July 2018.

[10] A. Yassin, Y. Nasser, M. Awad, A. Al-Dubai, R. Liu, C. Yuen, R. Raulefs, and E. Aboutanios, "Recent advances in indoor localization: A survey on theoretical approaches and applications," *IEEE Communications Surveys & Tutorials*, vol. **19**, no. 2, pp. 1327–1346, Nov. 2016.

[11] F. Li, M. Al-qaness, Y. Zhang, B. Zhao, and X. Luan, "A Robust and Device-Free System for the Recognition and Classification of Elderly Activities," *Sensors*, vol. **16**, no. 12, p. 2043, Dec. 2016.

[12] D. Li, Y. Lei, and H. Zhang, "A Novel Outdoor Positioning Technique Using LTE Network Fingerprints," *Sensors*, vol. **20**, no. 6, p. 1691, Mar. 2020.

[13] J. Svečko, M. Malajner, and D. Gleich, "Distance estimation using RSSI and particle filter," *ISA Transactions*, vol. **55**, pp. 275–285, Mar. 2015.

[14] Y. Wang, M. Li, L. Bai, T. Fu, and F. Gao, "Normalized amplitude of channel state information: The robust parameter for indoor localization in wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. **14**, no. 2, p. 155014771875942, Feb. 2018.

[15] G. Li, E. Geng, Z. Ye, Y. Xu, J. Lin, and Y. Pang, "Indoor positioning algorithm based on the improved rssi distance model," *Sensors*, vol. **18**, no. 9, p. 2820, Aug. 2018.

[16] X. Fang and L. Chen, "An optimal multi-channel trilateration localization algorithm by radio-multipath multi-objective evolution in rssranging-based wireless sensor networks," *Sensors*, vol. **20**, no. 6, p. 1798, Mar. 2020.

[17] T. E. Salti, M. Orlando, S. Hood, G. Knelsen, M. Iarocci, Z. Lazzara, Y. Xie, J. C. Cheung, and I. Woungang, "A new set of bluetooth-based fingerprinting algorithms for indoor location services," in *IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, Vancouver, Canada, Oct. 2018, pp. 1–7.

[18] A. F. G. Gonçalves Ferreira, D. M. A. Fernandes, A. P. Catarino, and J. L. Monteiro, "Localization and positioning systems for emergency responders: A survey," *IEEE Communications Surveys Tutorials*, vol. **19**, no. 4, pp. 2836–2870, May 2017.

[19] P. Bahl and V. N. Padmanabhan, "Radar: an in-building rf-based user location and tracking system," in *Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064)*, vol. **2**, Tel Aviv, Israel, Mar. 2000, pp. 775–784.

[20] M. Brunato and R. Battiti, "Statistical learning theory for location fingerprinting in wireless lans," *Computer Networks*, vol. **47**, no. 6, pp. 825–845, April 2005.

[21] J. Zhang, G. Han, N. Sun, and L. Shu, "Path-loss-based fingerprint localization approach for location-based services in indoor environments," *IEEE Access*, vol. **5**, pp. 13756–13769, July 2017.

[22] M. T. Hoang, Y. Zhu, B. Yuen, T. Reese, X. Dong, T. Lu, R. Westendorp, and M. Xie, "A soft range limited k-nearest neighbors algorithm for indoor localization enhancement," *IEEE Sensors Journal*, vol. **18**, no. 24, pp. 10208–10216, Oct. 2018.

[23] S. Subedi, D.-H. Kim, B.-H. Kim, and J.-Y. Pyun, "Improved Smartphone-Based Indoor Localization System Using Lightweight Fingerprinting and Inertial Sensors," *IEEE Access*, vol. **9**, pp. 53343–53357, 2021.

[24] N. Shaw, T. El Salti, J. C. - C. Cheung, and E. Sykes, "A New Set of Wi-Fi Dynamic Line-Based Localization Algorithms for Indoor Environments," *Procedia Computer Science*, vol. **191**, pp. 76-84, 2021.

[25] "Location Based Services Design," March 2018, Accessed on: January 27 2022. [Online]. Available: https://www.ciscolive.com/c/dam/r/ciscolive/apjc/docs/2018/pdf/BRKEWN-2650.pdf

[26] J. H. Lee, B. Shin, D. Shin, J. Kim, J. Park and T. Lee, "Precise Indoor Localization: Rapidly-Converging 2D Surface Correlation-Based Fingerprinting Technology Using LTE Signal," *IEEE Access*, vol. **8**, pp. 172829-172838, 2020.

[27] J. Bardwell. Converting signal strength percentage to dbm values. Technical report, WildPackets, Inc., 2002.

[28] S. J. Julier and J. K. Uhlmann, "Unscented Filtering and Nonlinear Estimation," *Proceedings of the IEEE*, vol. **92**, no. 3, pp. 401–422, Mar. 2004.

[29] I. Bisio, F. Lavagetto, M. Marchese, and A. Sciarrone, "Smart probabilistic fingerprinting for wifi-based indoor positioning with mobile devices," *Pervasive and Mobile Computing*, vol. **31**, pp. 107–123, Sept. 2016.

[30] R. Ma, Q. Guo, C. Hu, and J. Xue, "An improved wifi indoor positioning algorithm by weighted fusion," *Sensors*, vol. **15**, no. 9, pp. 21824–21843, Aug. 2015.

[31] J. Choi, W. Lee, J. Lee, J. Lee and S. Kim, "Deep Learning Based NLOS Identification With Commodity WLAN Devices," *IEEE Transactions on Vehicular Technology*, vol. **67**, no. 4, pp. 3295-3303, April 2018.

[32] Y. Shi, W. Zhang, Z. Yao, M. Li, Z. Liang, Z. Cao, H. Zhang, and Q. Huang, "Design of a Hybrid Indoor Location System Based on Multi-Sensor Fusion for Robot Navigation," *Sensors*, vol. **18**, no. 10, p. 3581, October 2018.

[33] R.- S. Hsiao, C. - H. Kao, T. - X. Chen, J. - L. Chen, "A Passive RFID-Based Location System for Personnel and Asset Monitoring," *Technology and Health Care*, vol. **26**, no. 1, pp. 11-16, 2018. Accessed on: October 14, 2021.