

# Software Simulation for Optimization k-NN Based Indoor Localisation Technique Using Spearman's Rank Correlation Coefficient

Kameran Adil Ibrahim

College of Education-Tuzkhurmatu, Tikrit University, Tikrit, Iraq.

[Kamaran\\_zm@yahoo.com](mailto:Kamaran_zm@yahoo.com)

## ABSTRACT

The reuse of existing Wireless Fidelity (Wi-Fi) setup for indoor localization using Wi-Fi Received Signal Strength Indicator (RSSI) is nowadays an active research domain. Over the period these Wi-Fi setups show degradation in performance owing to signal attenuation caused by multipath, along with environmental changes adversely affecting the functional efficiency. To optimize the indoor localization precision in the presence of the issues as mentioned earlier, I propose Spearman's Rank based Correlation Coefficient approach which finds the minimum distances and provides these distances to the original K-Nearest-Neighbor (k-NN) classifier which uses Euclidean distance. After the complete indoor Wi-Fi environment is simulated in Matrix Laboratory (Mat-lab) tool, the results so obtained are promising and on the higher side as compare to the original k-NN classifier performance. In case of distribution of cumulative errors the proposed method achieved low amount of localized errors of 2.7m for 80% tested samples. And as for shadow fading increase in value of  $\sigma$  improves the effect of the proposed method significantly.

**Keywords:** Indoor localization, Mat-lab Simulation, Euclidean Distance, Spearman rank correlation coefficient, k-NN classifier.

## برنامج محاكاة لمفاضلة k-NN على اساس تقنية التوطين في الاماكن المغلقة مستخدما معامل ارتباط الرتب سبيرمان

كامران عادل ابراهيم

كلية التربية-طوزخورماتو، جامعة تكريت، العراق.

[Kamaran\\_zm@yahoo.com](mailto:Kamaran_zm@yahoo.com)

### الملخص

إعادة استخدام اعدادات الـ واي فاي الحالي للتوطين في الاماكن المغلقة مستخدما مؤشر قوة الاشارة المستقبلية واي فاي (RSSI) هو مجال فعال للبحث في الوقت الحاضر. وخلال هذه الفترة انظمة الـ واي فاي تظهر تدهورا في الاداء بسبب انخفاض الاشارة الناتج عن تعدد المسارات الى جانب التغييرات البيئية التي تؤثر سلبا على كفاءة ادائها. ولتحسين دقة التوطين في الاماكن المغلقة مع جود الاسباب التي سبق ذكرها. اقترح نهج معامل ارتباط الرتب سبيرمان الذي يجد الحدود الدنيا للمسافة ويزودها للمصنف (K-NN) الاصلي والذي يستخدم المسافة الاقليدية بعد محاكاة البيئة الداخلية للـ واي فاي كاملة في (المات لاب). ان النتائج التي تم الحصول عليها تعتبر واعدة وعالية مقارنة بالأداء الاصلي للمصنف (K-NN). في حالة توزيع الاخطاء التراكمية حققت الطريقة المقترحة كمية منخفضة من الاخطاء المترجمة  $m2.7$  لـ 80% للعينات التي تم اختبارها. اما بالنسبة لزيادة قيمة  $\sigma$  يحسن من تأثير الطريقة المقترحة بشكل كبير.

**الكلمات المفتاحية:** التوطين في الأماكن المغلقة ، محاكاة ( مات-لاب) ، المسافة التقليدية، معامل ارتباط الرتب سبيرمان، المصنف K-NN.

## 1. Introduction

There is a widespread need for Location-Based Services (LBSs) due to an exponential growth of wireless technology and mobile based computing. The current decade is witnessing large scale deployment of Wireless Fidelity (Wi-Fi) infrastructures in hotels, airports, educational institutes, and supermarkets, etc. Furthermore, portable Wi-Fi modules are commonly available either standalone or integrated into smart devices making it possible for localisation of signal strengths of Wi-Fi in the indoor scenario. Global Positioning System (GPS), and similar other Global Navigation Satellite Systems (GNSS) based technologies can be used for outdoor localisation. But the GNSS system doesn't efficiently perform for indoor localisations, therefore, making it necessary to find other means of indoor localisation like Received Signal Strength Indicator (RSSI). However, even the RSSI approach also faces challenges owing to three major causes. First and foremost cause is the difficulty is acquiring correct RSSI values since the variance of RSSI values obtained from a stationary receiver goes up to more than 4dB within 60 seconds. The second cause is RSSI multipath and Non-line-Of-Sight (NLOS) effect in indoors due to ceiling, walls, floor, people, and furniture. And thirdly, the variations in devices by different manufacturers also affect the accuracy of measurement.

There are two approaches for localization methods that are based on RSSI - localization methods based on ranging and localization methods based on Received Signal Strength (RSS) fingerprint. Fingerprint based technique uses the model proposed by p. Bhal [1] and Y. Xu [2], which converts values obtained from RSSI to distances. These distances are used to perform localization based on lateration methods. RSSI based technique is a two-step procedure i.e. training and locating, though the training procedure is efforts intensive, time-consuming and susceptible to changing environment. Indoor localization of more than 10,000 locations around the world can be obtained, owing to the advent of Google maps.

As per reports, various WiFi devices exhibit difference and instability in terms the absolute RSSI values. For minimization of such outcomes, for identification of location, relative values of RSSI with a ranking are used in place of absolute values [3, 4 and 5] It apparently implies that through various Wi-Fi terminals, the obtained absolute values of RSSI from a group of Access Points of the included region might vary but they are expected to be more or less similar in a selected region.

Since the RSSI values decreases monotonically as the distance between Access Points (Aps) and source increases [6] therefore to determine the resemblance between various rankings of the same APs, a nonparametric statistical measure called Spearman rank correlation coefficient is used [7]. It uses monotonic function to describe the relationship between two variables; these variables can be continuous, discrete or even ordinals. Hence, a Spearman rank coefficient of correlation is proposed here for the determination of resemblance identification of the all gradings within the same set of APs. Lateration procedure based on RSS and KNN method respectively has been implemented for differentiation [8, 9]. A simulated real indoor environment in Mat-Lab software that employs segregated regions from attenuation factor propagation model [7] has been used for testing the proposed technique.

## 2. Proposed Methodology

For achieving optimum version of Spearman-distance-based K-NN [10] location technique, Spearman rank coefficient of correlation of RSSI measurements from varying APs is proposed in this article. The complete procedure for the proposed approach for localization procedure can be seen in Fig. (1). It is a three step procedure. At the outset, fingerprint database of the offline RSSI is created, followed by collection of the positional fingerprints of the randomly chosen entities and thirdly, based on equation (5) the Spearman distance is computed along with selection of all regions having minimum spearman distance. In the end, the native K-NN technique is implemented with respect to various regions having the least Spearman Distance and the final estimation of location is acquired.

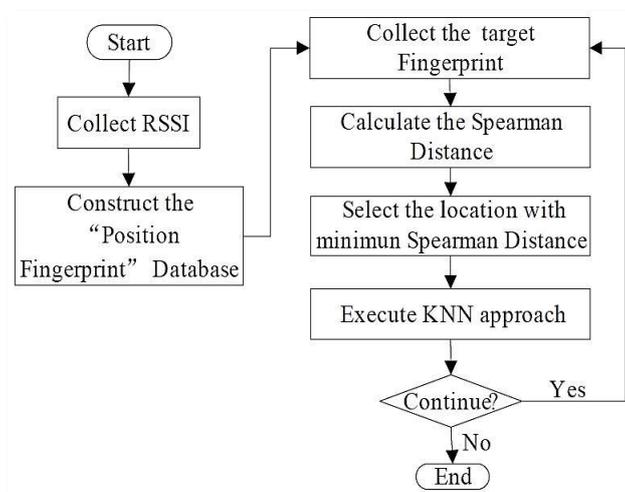


Fig. (1): Flow chart depicting all the steps involved in localization process.

### 3. Method

The correlation between target fingerprint  $T$  and the reference fingerprint  $R_{ij}$  can be obtained from Spearman rank correlation coefficient [11]. Though, similar amounts of APs may not be possessed by the fingerprints at target. Owing to this, few alterations are advisable prior to the computation of Spearman coefficient of correlation. Matrices  $V_R$  and  $V_T$   $N_t \times 2$  are generated, which are initialized as

$$V_T = V_R = \begin{bmatrix} ID_t & N_t \\ \cdot & \cdot \\ ID_t & N_t \end{bmatrix} \quad (1)$$

In the  $T$ 's ranking of RSSI, the second column of the corresponding  $V_T$  row must include the position of the APs such that

$$V_T(n_k, 2) = k \quad (2)$$

where  $V_T(n_k, 1) = T(k, 1)$ ,  $n_k \in [1, N_t]$  and  $k = 1, 2, \dots, N_t$ . Similarly,  $V_R$  can be renewed to be

$$V_R(n_k, 2) = k \quad (3)$$

Once  $V_R(n_k, 1)$  is equal to  $R_{i,j}(k, 1)$ ,  $n_k \in [1, N_t]$  and  $k$  is equal to  $1, 2, \dots, N_{i,j}$ .

Therefore, the target fingerprint and reference fingerprint at pixel  $(i, j)$ 's Spearman rank coefficient of correlation can be given as

$$\rho_{i,j} = \frac{\sum_{n=1}^{N_t} [(V_T(n, 2) - \bar{R}_T)(V_R(n, 2) - \bar{R}_R)]}{\sqrt{\sum_{n=1}^{N_t} [(V_T(n, 2) - \bar{R}_T)^2 (V_R(n, 2) - \bar{R}_R)^2]}} \quad (4)$$

Where

$$\bar{R}_T = \frac{1}{N_t} \sum_{n=1}^{N_t} V_T(N, 2),$$

$$\bar{R}_R = \frac{1}{N_t} \sum_{n=1}^{N_t} V_R(N, 2),$$

In the next step the Spearman distance  $d_{i,j}$  is given as

$$d_{i,j} = 1 - \rho_{i,j} \quad (5)$$

#### 4. Simulation

The simulation makes use of one hundred square meter field with 10 m sides with 400 predefined locations evenly distributed as shown in Fig. (2) as small blue circles. The generation of RSSI values of 400 already defined locations and in this area randomly selected objects are placed. The labels to the localized objects  $x$  and  $y$  are obtained as  $x = N_x \times \text{rand}$  and  $y = N_y \times \text{rand}$ . Matlab library function `rand` is used to generate random values in the range of 0 to 1. The highest value in the  $x$  and  $y$  directions are indicated by  $N_x$  and  $N_y$  and is selected to 10m shown in Fig. (2). At all the corners of the square area, located four APs are considered by the simulation. To simulate indoor settings, nine districts are formed by segregating the complete region, which represents nine independent regions with concrete walls for separating them from each other. The wireless signals from devices present in different rooms are deteriorated during propagation by various number of intercepting walls prior to reaching the APs. To the shadow fading propagation model a Partition Attenuation Factor (PAF) [2] is added for simulating the indoor signal propagation. Therefore, the propagation model showing impact of shadow fading can be obtained as [6]

$$P(d) = P(d_0) - 10\gamma \log\left(\frac{d}{d_0}\right) - W \times PAF + X\sigma \quad (6)$$

Where, the total loss (in decibel) is given by  $P(d)$ , loss at the reference distance  $d_0$  is given by  $P(d_0)$ , path loss exponent is given by  $\gamma$ , and a particular hindrance like an indoor wall is given by PAF. Here, this is used for penetration loss simulation under the condition of signals passing through the wall.

The number of walls between the APs and the object node are given by  $W$ , and the normal random variable is given by  $X\sigma$ . A clear interpretation can be attained if one considers the wireless device located in Room4 of Figure2, For  $AP_1$ ,  $AP_2$ ,  $AP_3$  and  $AP_4$ , the value for  $W$  is 1, 3, 3, 1, respectively. For simulating the indoor environment with better accuracy, it

becomes pertinent to consider the correlation. The shadow fading effect's spatial correlation can be computed as follows:

1. Covariance matrix K is generated such that:

$$K_{ij}(d_{ij}) = \sigma^2 \exp\left(-\frac{d_{ij}}{D_c}\right)$$

Where  $D_c$  is the decorrelation distance,  $d_{ij}$  represents the distance of  $i^{\text{th}}$  position with respective  $j^{\text{th}}$ .

2. Obtaining cholesky factorization of covariance matrix K.

3. Generation of non-correlated normal random variables w, such that:

$$W = [W_1, \dots, W_m]^T$$

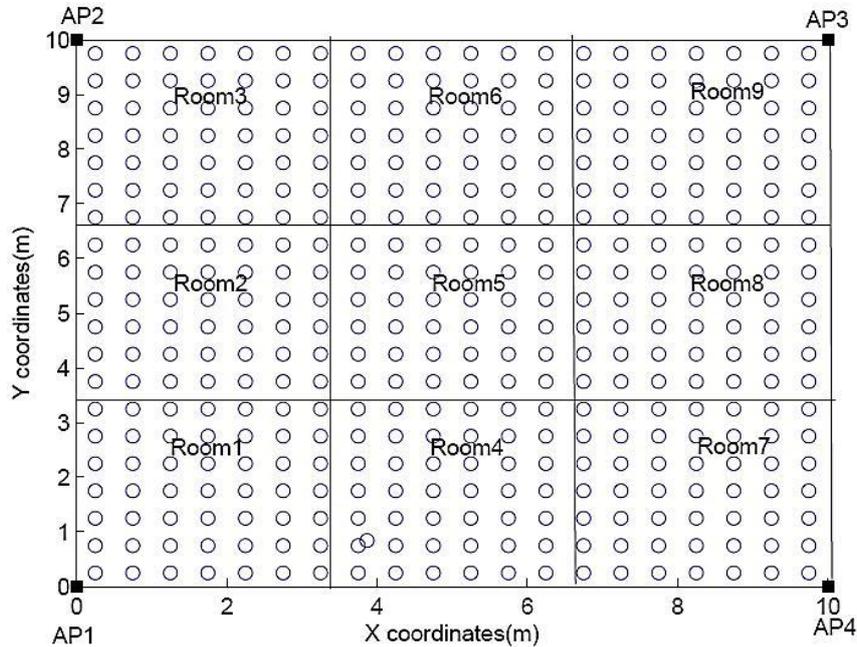
5. Correlation of shadow fading at location i and j is expressed as :

$$E[X_\sigma(i)X_\sigma(j)] = K_{ij}(d_{ij}) = \sigma^2 \exp\left(\frac{-d_{ij}}{D_c}\right)$$

With these calculations, the environment of NLOS (Non-line-of-sight) is considered and the values for parameters like  $d_0$ ,  $P(d_0)$ ,  $\gamma$ , PAF,  $D_c$  are further shown is [Table \(1\)](#).

**Table (1): Parameter setting**

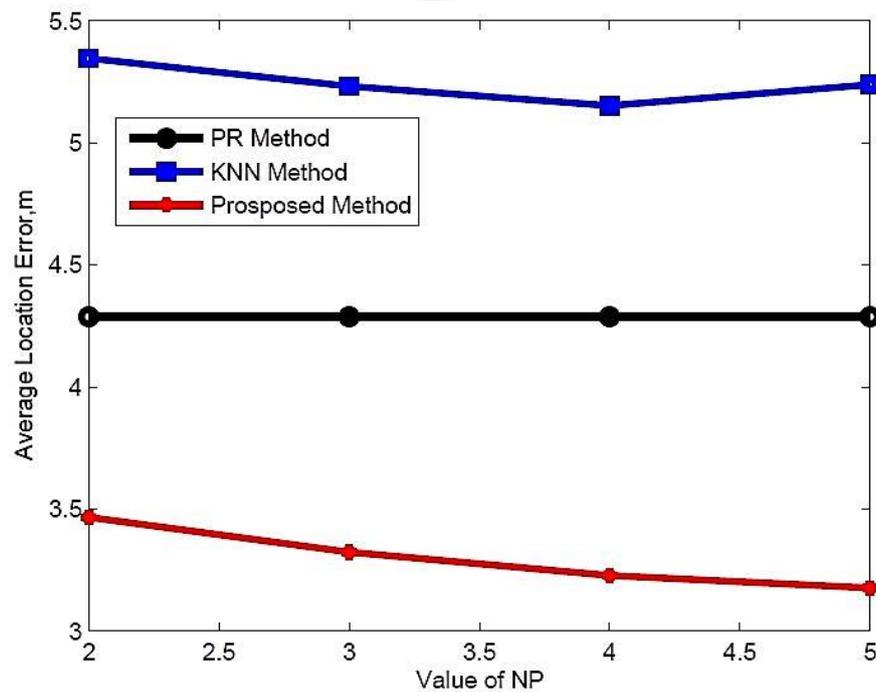
Parameters	$d_0$	$P(d_0)$	$\gamma$	PAF	$D_c$
Values	1m	-37.3dBm	3.3	5dB	



**Fig. (2):** Layout of Simulation in Matlab

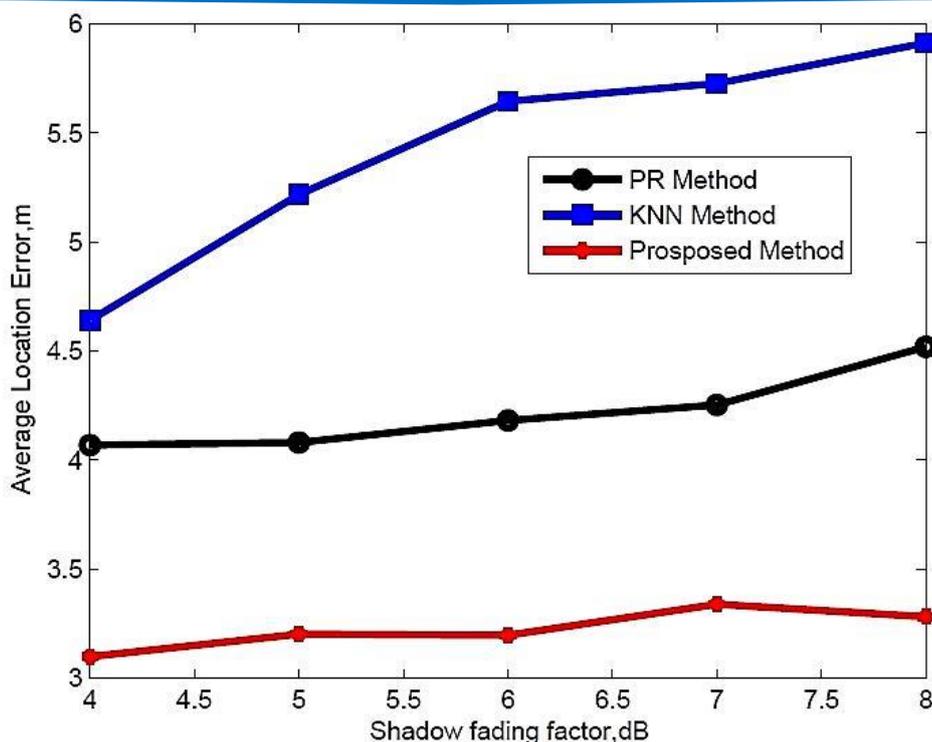
## 5. Results and Discussions

One can conveniently assess the change in the number of neighbour points that are closest to each other by choosing various values from the range of 2 to 5. The mean of location errors (ALE) of the Polynomial Regression-based method (PR method), the KNN method, and the Proposed method is given in Fig. (3). The fact that the value of ALE for the PR method is independent from the value of Neighbor Points (NP) as NP remains unaffected, is quite evident Fig. (3) depicts the curve corresponding to the KNN method showcasing a steady decline in Average Location Errors (ALE) as observed in afore rising which reaches the lowest point when NP has a value 4. However, ALE exhibits a slow and steady fall, ranging from 3.5m to just below 3.2 m at the time NP increases to 5 from the low point 2 in case of the proposed method, NP is stationed at value 4 in subsequent subsections for sound and a reasonable comparative analysis.



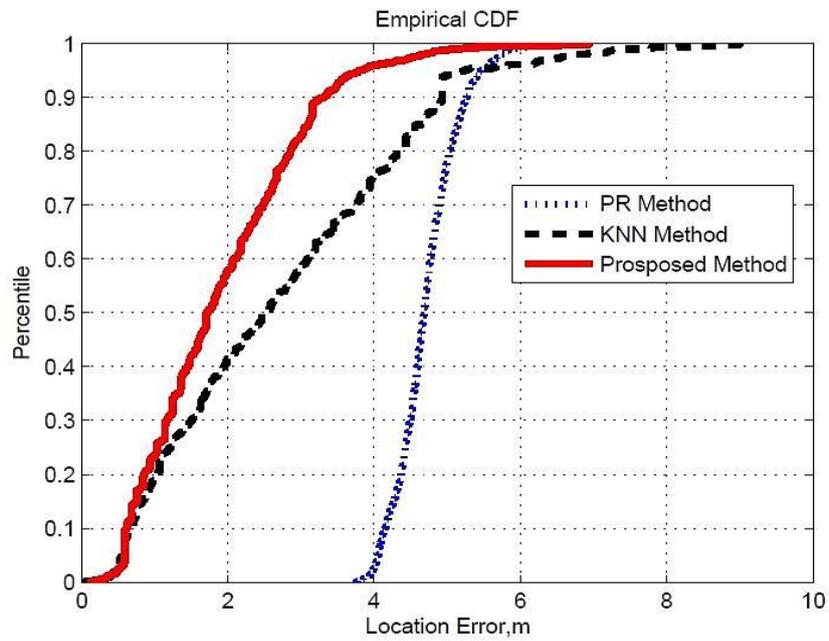
**Fig. (3):** Comparing average Location Errors (ALE) for varying values of NP.

For a 1000 times, the same quantifications are obtained in the range 4dB-8dB to understand the effect of shadow fading factor. The alterations in ALE corresponding to the variations in the shadow fading factor within the above given scope can be seen in Fig. (4). The simulation has also been performed on the native KNN technique and the polynomial-regression-based method (PR Method) [2] for comparative purposes. It is apparent that there is significant increase in positioning error for the native KNN technique as the shadow fading factor grows. It is obvious to get such outstanding outcomes since a big shadow fading leads to formation of position fingerprint that is more unreliable and which consequently increases the effect of location error to a particular limit. In contrast, there is no vital effect on the rest of the two methods by the value of shadow fading. There is a homogenous performance in terms of localization for the native KNN method and the PR technique when value of  $\sigma$  is small and that is not superior to the proposed technique. In comparison with the other two methods, with the increase in the value of  $\sigma$ , the efficacy of our proposed method rises considerably, which thereby is indicative of the superiority of the proposed method in awful indoor locations.

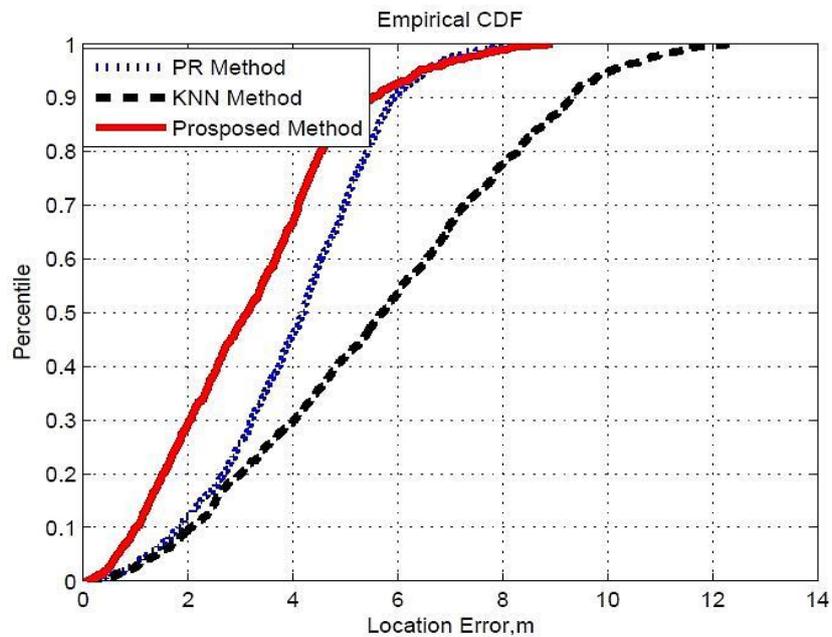


**Fig. (4): Comparing ALE of three different location methods**

For errors that are localized, the Cumulative distribution function(CDF) available in indoor environment that is simulated, when the shadow fading factor is 5dB and 7dB can be seen Fig. (5) and Fig. (6) respectively. For 80% testing samples, it can be seen that the error that are localized, amount to less than 2.7% and this is quite less as compared to the one obtained for KNN technique which is 4.5m and the PR technique which is 4.9m. This can be seen in Fig. (5). For 80% of testing samples, the errors that are localized amount to 4.6m in a situation when shadow fading factor is increased to 7dB starting from 5dB. The native KNN technique could not perform this better when the errors due to localization were 8.2m and those due to PR technique were 5.8m. From the above observations, it can be deduced that the proposed scheme showcases superiority in the indoor regions when there is instability in RSSI values along with changes of temporal nature. This is due to the fact that the proposed scheme considers the RSSI gradings which are useful for obtaining precision in location of fingerprints.



**Fig. (5):** Location Error when cumulative distribution function (CDF) is set to  $\sigma = 5\text{dB}$ .



**Fig. (6):** Location Error when cumulative distribution function (CDF) is set to  $\sigma = 7\text{dB}$

## 6. Conclusion

This work introduces a novel indoor location system based on Spearman-distance relying on the fingerprint of RSSI values previously provide by the APs. From the training procedure, a radio map is formed using the values of RSSI, which are called as “Fingerprints.” The spearman rank correlation coefficient is computed after acquisition of the position fingerprint which is unidentified. In the later step, the native KNN method is integrated along with the spearman distance based on the spearman rank coefficient of correlation. The outcome of the experimentation indicates that the proposed technique of using eanking of spearman correlation coiffients in comparison with the other two techniques i.e. PR method and KNN is capable of attaining superior performance.

Performing received signal strength (RSS) based indoor localization is particularly challenging but using the ranking of Spearman’s coefficient to improve the performance of k-NN classifier was attempted in this article and successfully tested. This implementation can be directly used in indoor localization of Wi-fi routers to decide the place of the router where it is used to its optimum.

## References

- [1] K. Wu, J. Xiao, Y. Yi and D. Chen, “*CSI-Based Indoor Localization*”, IEEE Transactions on Parallel and Distributed System, 24(7), 1300, (2013).
- [2] P. Bahl and V.N. Padmanabhan, “*Radar: An In-building RF-Based User Location and Tracking System*”, Proceedings of the IEEE INFOCOM, (2000).
- [3] Deora, S. and Krishnamachari, B., “*Harnessing Non-Uniform Transmit Power Levels for Improved Sequence Based Localization*”, IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS), Marina Del Rey, CA, (2014).
- [4] Yu Y., Jiang C., Zhao X. etc., “*Sequence-based localization algorithmwith improved correlation metric and dynamic centroid*”, Science China Information Sciences, 54(11), 2349, (2011).
- [5] J. Machaj, P. Brida, and R. Piche, “*Rank based fingerprinting algorithmfor indoor positioning*”, in 2011 International Conference on Indoor Positioning and Indoor Navigation (IPIN), IEEE, 1, (2011).

- 
- [6] J. Krumm, G. Cermak, and E. Horvitz, “*Rightspot: a novel sense of location for a smart personal object*”, in Proceedings of Ubicomp 2003, Seattle, USA, 36, (2003).
- [7] The MathWorks, Inc., “*Pairwise distance between pairs of objects*,” <http://www.mathworks.de/access/helpdesk/help/toolbox/stats/pdist.html>, (May 2017).
- [8] J. Yang, Y. Chen, “*Indoor Localization Using Improved RSS-Based Lateration Methods*”, Proceedings of the 28<sup>th</sup> IEEE Conference on Global Telecommunications, GLOBECOM’09, Piscataway, NJ, USA, (2009).
- [9] Xu X., Deng Z., “*A novel sequence based localization approach for wireless sensor networks*,” In: Proceedings of the 5<sup>th</sup> International Conference on Wireless Communications, Networking and Mobile Computing, Beijing, China, (2009).
- [10] P. Torteeka and X. Chundi, “*Indoor positioning based on Wi-Fi fingerprint technique using fuzzy K-nearest neighbor*”, in Proceedings of the 11<sup>th</sup> International Bhurban Conference on Applied Sciences and Technology (IBCAST14), Islamabad, (2014).
- [11] S. A. Zekavat and R. M. Buehrer, “*Handbook of Position Location: Theory, Practice and Advances*”, John Wiley and Sons, (2011)
- [12] Mautz, R., “*Indoor positioning technologies*”, PhD thesis, (2012).