

IMPROVED INDOOR POSITIONING USING FINGERPRINT TECHNIQUE AND WEIGHTED K-NEAREST NEIGHBOUR

Sh. Naderi Salim^{1*}, M.M. Alizadeh^{1,2,3}, Sh. Chamankar¹, H. Schuh^{2,3}

¹ Faculty of Geodesy and Geomatics Engineering, K. N. Toosi University of Technology, Tehran, Iran

² German Research Center for Geosciences (GFZ), Potsdam, Germany

³ Institute for Geodesy and Geoinformation Science, Technische Universität Berlin, Berlin, Germany
shervin.naderi@email.kntu.ac.ir, alizadeh@kntu.ac.ir, chamankar@email.kntu.ac.ir, schuh@gfz-potsdam.de

Commission IV, WG IV/3

KEY WORDS: WLAN Positioning, Fingerprint Algorithm, Location-Based Services, Nearest Neighbour, Weighted K-Nearest Neighbour

ABSTRACT:

Global Navigation Satellite Systems are not effective when there is no direct line of sight between the user and the satellites, such as indoor environments and dense urban areas. Today, location-based services are used significantly due to their utility and ease of access. The fingerprint method is one of the common methods of determining the location in indoor environments. In this research, the indoor positioning system based on the fingerprint algorithm with a wireless network has been implemented. The positioning system based on the method of nearest neighbour and weighted K-nearest neighbour with two access points has been implemented in two different scenarios. The output accuracy of each technique has been compared to each other. The main goal of this article is to compare the accuracy of positioning with the fingerprint method using the mentioned algorithms and to find the most suitable mode and algorithm for determining the indoor position in most places. The improved weighted nearest neighbour method will have an almost acceptable result in all scenarios and also in the first scenario with dense and regular reference points the weighted K-nearest neighbour method with $RMSE=0.2812(m)$ has provided the best result. In the second scenario with scattered and irregular reference points the weighted K-nearest neighbour with $RMSE=0.6735(m)$ has given lower accuracy result.

1. INTRODUCTION

Global Navigation Satellite Systems (GNSS) can provide continuous position, velocity, and time, however, these systems are weakened or obstructed when there is no good sky sight, especially in indoor environments. In these conditions, GNSS cannot be very effective. With the continuous increase of demand for Location-based service (LBS), high accuracy and high-rate positioning has turned more important than ever. Indoor positioning with other systems, such as WLAN, Inertial Navigation System (INS), etc., have emerged as a promising solution to overcome the GNSS problem. The expansion WLAN based on radio signal technology has created a new opportunity for LBS. By using WLAN without additional equipment, it is possible to determine the position in indoor environments. The classification criteria of different positioning methods in indoor environments can be considered Beacon base, Device base, and device-free Base or Passive. Also, indoor positioning algorithms are divided into four general parts: proximity, triangulation, Scene analysis, and Dead-rocking. The common way to determine the position is to use the Received Signal Strength Indication (RSSI) from WLAN. (Madigan et al., 2005) used the RSS from fixed reference points. In (Aomumpai et al., 2014) and (Sánchez-Rodríguez et al., 2015) for the online step, researchers are focused on how to improve positioning accuracy. Scene analysis has advantages over other methods in terms of accuracy and does not require the position of WLAN receiver points. Also, due to the use of the RSS, the need for line of sight is eliminated. The data combination method has been used in positioning with Wi-Fi in (Rodionov et al., 2013). The fingerprint Algorithm is usually used in the Scene analysis method. Several positioning

methods can be used on this basis: probability methods, deterministic methods (Nearest Neighbour (NN), Weighted K-Nearest Neighbour (WKNN), etc.), Artificial Neural Networks, support vector machine are among these. (Güvenç and Chong, 2009)

2. PROPOSED METHODS

WLAN that widely spread uses the IEEE 802.11 standard and can easily be used in indoor locations to estimate the position of a receiving device such as a mobile phone. Each Wi-Fi can cover 300 meters of open space. The most popular method of WLAN positioning is the use of RSSI, which is easily extracted in 802.11 networks. Most wireless devices receive signal strength information received by receiver points such as strength measurement or signal-to-noise ratio (SNR) as communication quality. This signal strength usually covers ranges between -100 dBm (decibel-milliwatts) and -10 dBm. Because the signal strength shows a better correlation with the location than the SNR parameter, as a result, the WLAN positioning system uses more signal strength than the SNR (Mautz, 2012). In this article, we will implement the fingerprint method in three different scenarios, during this process, we will use the common selection method to select the best access points, and then we will filter the strength of the received signals by maximum and average filters; After that, we will estimate the position of the unknown points by using the Nearest Neighbour and Weighted K-Nearest Neighbour positioning methods that we will discuss each of these methods in the following.

* Corresponding author

2.1 Fingerprint Algorithm

The fingerprint algorithm is a method in which the coordinates of the points are obtained by comparing the observations made in the calibration stage by the operator and the observations made at the unknown point by the user, during a positioning algorithm. The fingerprint algorithm usually consists of two steps. In the first stage, which is also called the offline phase, RSSIs are taken from each transmitter at points known as Reference Points (RP). As a result, each RP has a set of RSSI from each transmitter. In the second stage, which is also called the online phase. First, the observations at the user's unknown point are sent to the reference system; then, with the help of positioning methods and based on the observations in the offline phase and the known coordinates of the RPs, the coordinates of the unknown point are estimated. Figure 1 and Figure 2 describe the offline and online phases. The main difference between positioning systems based on fingerprints is the use of different algorithms to determine positioning in the online section. (He and Chan, 2016)

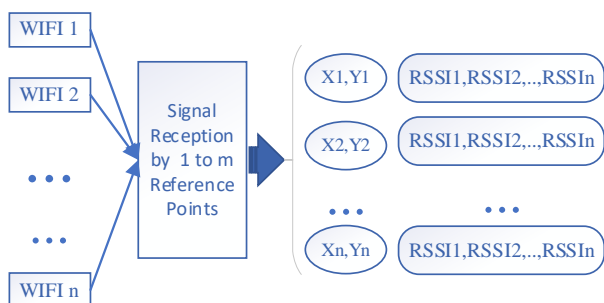


Figure 1. The offline phase includes signal reception by reference points from different transmitters and forming a database

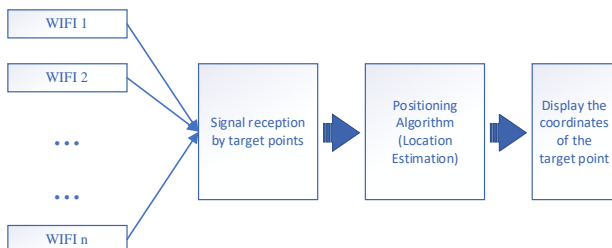


Figure 2. The online phase includes receiving the signal by the target point and determining the position with location estimation algorithms

2.2 NEAREST NEIGHBOUR

In this method, the distance between RPs and Access Points (AP) is calculated and stored with the help of a distance function in the offline phase. In the online phase, to obtain the coordinates of the unknown point, the distance between the unknown point and the transmitter is estimated by observing the signal strength, and the coordinates of the closest RPs are considered as the coordinates of the unknown point. We can use different distance functions to calculate the distance. p norm is one of the most common methods for distance estimation.

$$\|x\|_p = \left(\sum_{i=1}^{n_x} |\bar{x}_i|^p \right)^{\frac{1}{p}} \quad (1)$$

Calculating the distance with this method leads to the use of the NN method in the distance space.

If we use the strength of the received signal to check the similarity of the points, we have implemented the NN method in the signal space. (Bahl and Padmanabhan, 2000)

2.3 WEIGHTED K-NEAREST NEIGHBOUR

In this method, the coordinates of the unknown point can be estimated from the weighted average of K coordinates of the neighbour with the smallest distance. The easiest way to calculate the weight is to use the inverse of distance as the weight of each RP.

$$\bar{X} = \frac{1}{\sum_{j=1}^K w_j} \sum_{i=1}^K w_i P_i \quad (2)$$

Where K the number of effective reference points (nearest neighbours) and \bar{X} coordinate components are unknown. (Zhang et al., 2022)

2.4 Access point selection methods

Access point selection methods can be used in both offline and online stages. Choosing access points based only on the observations of the offline phase will not bring an appropriate result. Therefore, the methods of selecting access points are usually done explicitly or implicitly in the online phase and with the help of observations of this phase. In the explicit method, a set of access points are selected only according to the observations of the online phase, but in the implicit method, the selection of access points is done with the help of both online and offline phase observations. One of the methods that can be used in this section is to select a set of receivers based on online phase observations and apply access point selection techniques to them in the offline phase.

2.4.1 Fisher's criterion: Fisher's criterion is a metric for the degree of differentiation between different transmitters in all control points in terms of stability. This metric checks the efficiency of each access point in the reference points, in the offline phase, by using the statistical elements of the signal intensity of the reference points.

The score assigned to each access point is calculated from the following equation. (Khalajmehrabadi et al., 2017)

$$\zeta^i = \frac{\sum_{j=1}^N (\psi_j^i - \bar{\psi}^i)^2}{\frac{1}{M-1} \sum_{m=1}^M \sum_{j=1}^N (r_j^i(t_m) - \psi_j^i)^2} \quad (3)$$

In the equation (3), i number of access points, N number of reference points, $\bar{\psi}^i = \frac{1}{N} \sum_{j=1}^N \psi_j^i$ average time of observations

from the transmitter i at the location of the reference point j , ψ_j^i duration of observation of the transmitter i at point j , M number of observations made in time, and $r_j^i(t_m)$ is signal strength observation at the location of point j of the transmitter i at the moment m .

After calculating the score of all senders, a set of senders with the highest score is selected. This criterion is based on the fact that transmitters with higher variance should have a lower score because these transmitters are less reliable. (Kushki et al., 2017)

Among the disadvantages of this method, we can point out that only offline phase observations are considered in this method, and if the selected transmitters are not visible in the online phase, or if the offline phase observations are accompanied by errors, this method will not be effective. (Khalajmehrabadi et al., 2017)

2.4.2 Common selection method: This method is similar to Fisher's criterion method, with the difference that instead of using the average time of observations in the fractional difference, the mutual distinction between the observations of two reference points is used.

$$\xi^i = \frac{\sum_{1 < j < j' < N} (\psi_j^i - \psi_{j'}^i)^2}{\frac{1}{M-1} \sum_{m=1}^M \sum_{j=1}^N (r_j^i(t_m) - \psi)^2} \quad (4)$$

In equation (4), values of ψ , N , M , i , j are defined similar to equation (3) and ξ^i is score of the transmitter i .

In this method, a set of transmitters with the highest score are selected. (Zhou et al., 2013)

2.5 Signal Strength Filter

The strength of the received signal interferes with different types of noise, so there are three different filtering methods which are maximum filter, mean (average) filter, and limited filter to improve the positioning accuracy.

2.6 Data storage methods

RSSIs at the RPs and target points can be stored raw or the average of RSSIs can be stored to reduce the amount of information, the latter is more common than other methods. To know how the signal strength changes, mean and variance and histogram methods are used.

In this paper we stored average of RSSIs.

2.7 Accuracy Assessment

In this research, Root Mean Square Error (RMSE) has been used to validate and compare the results of methods and scenarios (Equations 5).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n ((y_i - \hat{y}_i))^2} \quad (5)$$

Where, n , y_i , and \hat{y}_i are the total number of observations, biomass observed, and estimated at i point, respectively.

3. PROCESS AND IMPLEMENTATION

In this section, three scenarios are considered for implementation, and the specifications of each of these scenarios are listed in Table 1. Two access points that obtained by the maximum entropy method and having the same position in all scenarios have been used. The approximate size of the studied environment are 10×5 meters². The map of the first scenario is shown in Figure 4. In the first scenario, 27 reference points are considered densely and regularly with a distance of one meter from each other.

Scenarios	Number of Reference Points	Distance of adjacent points
First	27	1 meter
Second	29	Random and Irregular

Table 1. Specifications of each scenario

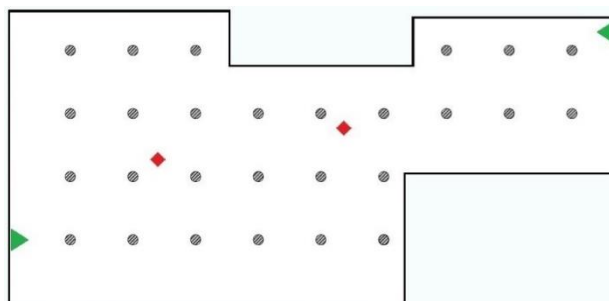


Figure 3. The map of the building and the distribution of reference points in scenario 1 (Gray circles are reference points, green triangles are access points and red diamonds are target points)

In the second scenario, 29 points with different dispersion and distances are considered irregularly. The map of the second scenario is shown in Figure 5.

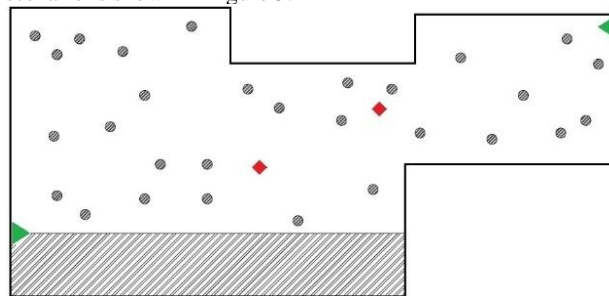


Figure 4. The map of the building and the distribution of reference points in scenario 2 (Gray circles are reference points, green triangles are access points and red diamonds are target points)

In all scenarios, the signal strength of reference points has been measured from two access points in 3 minutes. The obtained signal strength was filtered by maximum and mean filters and stored.

Then, in each scenario, the unknown points were estimated by NN and WKNN positioning methods, and their RMSE was calculated.

An MSI GP75 Leopard laptop with Windows 10 was used to measure the signal strength. The software used to collect signal strength was the Wirelessmon app (Passmark, 2021), which is free and open-source software.

4. RESULT AND DISCUSSION

In this article, to improve the accuracy of indoor positioning and reduce the cost and time of positioning, the methods of selecting the best access points and signal strength filter are used. The numerical results of the algorithms mentioned in the previous sections were obtained for two scenarios and then compared and studied.

4.1 First Scenario

In the first scenario, all the reference points are dense and regular, with a distance of one meter from each other.

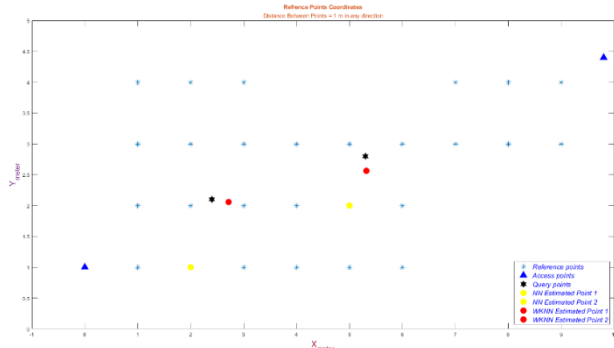


Figure 5. A view of determining the position of the target points by different methods (NN and WKNN) in the first scenario (Blue stars are reference points, blue triangles are access points, black stars are target points, yellow circles are estimated points from NN method and red circles are estimated points from WKNN method)

The target points were obtained by positioning methods, and RMSE was calculated for each method as described in table 2.

Methods	RMSE (Meters)
NN	1.0247
WKNN	0.2812

Table 2. RMSE for positioning methods in the first scenario

4.2 Second scenario

In this scenario, all the reference points are densely scattered, with a random distance from each other. Access points have the same location in both scenarios.

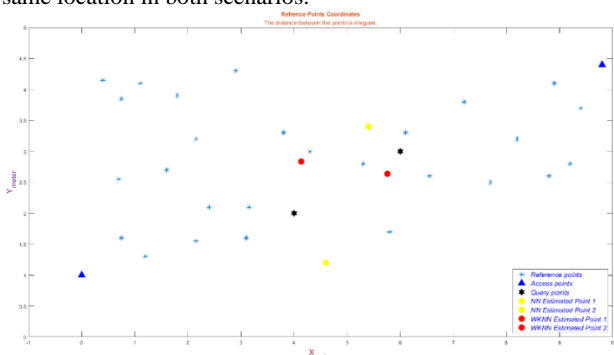


Figure 6. A view of determining the position of the target points by different methods in the second scenario (Blue stars are reference points, triangles are access points, black stars are target points, yellow circles are estimated points from NN method and red circles are estimated points from WKNN method)

Methods	RMSE (Meters)
NN	2.1354
WKNN	0.6735

Table 3. RMSE for positioning methods in the second scenario

4.3 Accuracy Assessment

For a better view of the results obtained from each method in all scenarios and to assess the accuracy more easily, all the results are listed in Table 5.

Also, in Figure 10, you can see a view of the RMSE diagram for all methods in each scenario.

Methods	RMSE (Meters)	
	First Scenario	Second Scenario
NN	1.0247	2.1354
WKNN	0.2812	0.6735

Table 4. RMSE for positioning methods in all scenario

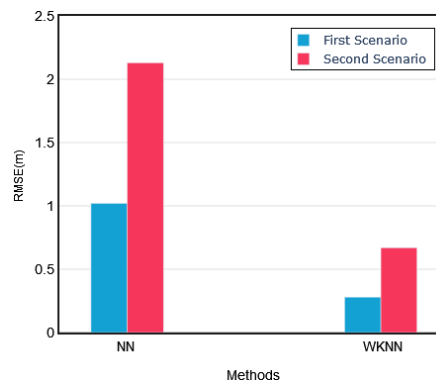


Figure 7. Methods accuracy assessment with RMSE in all scenarios

5. CONCLUSION

Considering the importance of positioning in indoor environments, in this article, we tried to obtain the best algorithm and the best conditions of selecting reference points. In the first scenario, the reference points are considered at a certain distance from each other and located regularly. In this scenario the accuracy of the weighted K-nearest neighbour method is improved. The nearest neighbour method results in an accuracy slightly higher than the maximum distances considered between adjacent reference points, and the weighted nearest neighbour with $RMSE=0.2812(m)$ provides the best result. In the second scenario, the reference points are scattered and irregular, which causes a great decrease in the accuracy of the nearest neighbour method, so the RMSE has increased to $2.1354(m)$. In this scenario, the weighted K-nearest neighbour has the lower accuracy result with $RMSE=0.6735(m)$ compared to first scenario. In this way, it is clear from the results that the improved weighted nearest neighbour method provides an almost acceptable result in all scenarios. In the first scenario and especially in environments where the number of reference points are increased and regularly located, the weighted K-nearest neighbour method have provided the best results compared to second scenario and environments with scattered and irregular reference. This result shows that in environments with average dispersion and an average number of points, the improved weighted nearest neighbour method has an acceptable result and in environments, with a higher number of reference points and with the same dispersion the accuracy of determining the location in the improved weighted K-nearest neighbour method has the best result compared with the environments where reference points have less number and not a uniform distribution.

REFERENCES

AOMUMPAI, S., KONDEE, K., PROMMAK, C. , KAEMARUNGSI, K. Optimal placement of reference nodes for wireless indoor positioning systems. 2014 11th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 14-17 May 2014 2014. 1-6.

BAHL, P., PADMANABHAN, V. N. RADAR: an in-building RF-based user location and tracking system. *Proceedings IEEE INFOCOM*, 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064), 26-30 March 2000 2000. 775-784 vol.2.

GÜVENÇ, I., CHONG, C.-C., 2009. A Survey on TOA Based Wireless Localization and NLOS Mitigation Techniques. *IEEE Communications Surveys & Tutorials*, 11.

HE, S., CHAN, S.-H. G., 2016. Wi-Fi Fingerprint-Based Indoor Positioning: Recent Advances and Comparisons. *Commun. Surveys Tuts.*, 18, 466–490.

KHALAJMEHRABADI, A., GATSIS, N., AKOPIAN, D., 2017. Modern WLAN Fingerprinting Indoor Positioning Methods and Deployment Challenges. *IEEE Communications Surveys & Tutorials*, 19, 1974-2002.

KUSHKI, A., PLATANIOTIS, K. N., VENETSANOPOULOS, A. N., 2017. Indoor Positioning with Wireless Local Area Networks (WLAN). In: SHEKHAR, S., XIONG, H. & ZHOU, X. (eds.) *Encyclopedia of GIS*. Cham: Springer International Publishing.

MADIGAN, D., EINAHRAWY, E., MARTIN, R., JU, W.-H., KRISHNAN, P., KRISHNAKUMAR, A. S., 2005. *Bayesian indoor positioning systems*.

MAUTZ, R., 2012. Indoor positioning technologies.

RODIONOV, D., KOLEV, G., BUSHMINKIN, K., 2013. A hybrid localization technique for patient tracking. *Annu Int Conf IEEE Eng Med Biol Soc*, 2013, 6728-31.

SÁNCHEZ-RODRÍGUEZ, D., HERNÁNDEZ-MORERA, P., QUINTEIRO, J. M., ALONSO-GONZÁLEZ, I., 2015. A Low Complexity System Based on Multiple Weighted Decision Trees for Indoor Localization. *Sensors*, 15, 14809-14829.

ZHANG, H., WANG, Z., XIA, W., NI, Y., ZHAO, H., 2022. Weighted Adaptive KNN Algorithm With Historical Information Fusion for Fingerprint Positioning. *IEEE Wireless Communications Letters*, 11, 1002-1006.

ZHOU, Y., CHEN, X., ZENG, S., LIU, J., LIANG, D., 2013. AP Selection Algorithm in WLAN Indoor Localization. *Information Technology Journal*, 12, 3773-3776.

PASSMARK., 2021. Passmark software. Available: <https://www.passmark.com/products/wirelessmonitor/> [Accessed].

APPENDIX

Figures 8 and 9 are bigger and higher-resolution images of the first and second scenarios in the result and discussion section.

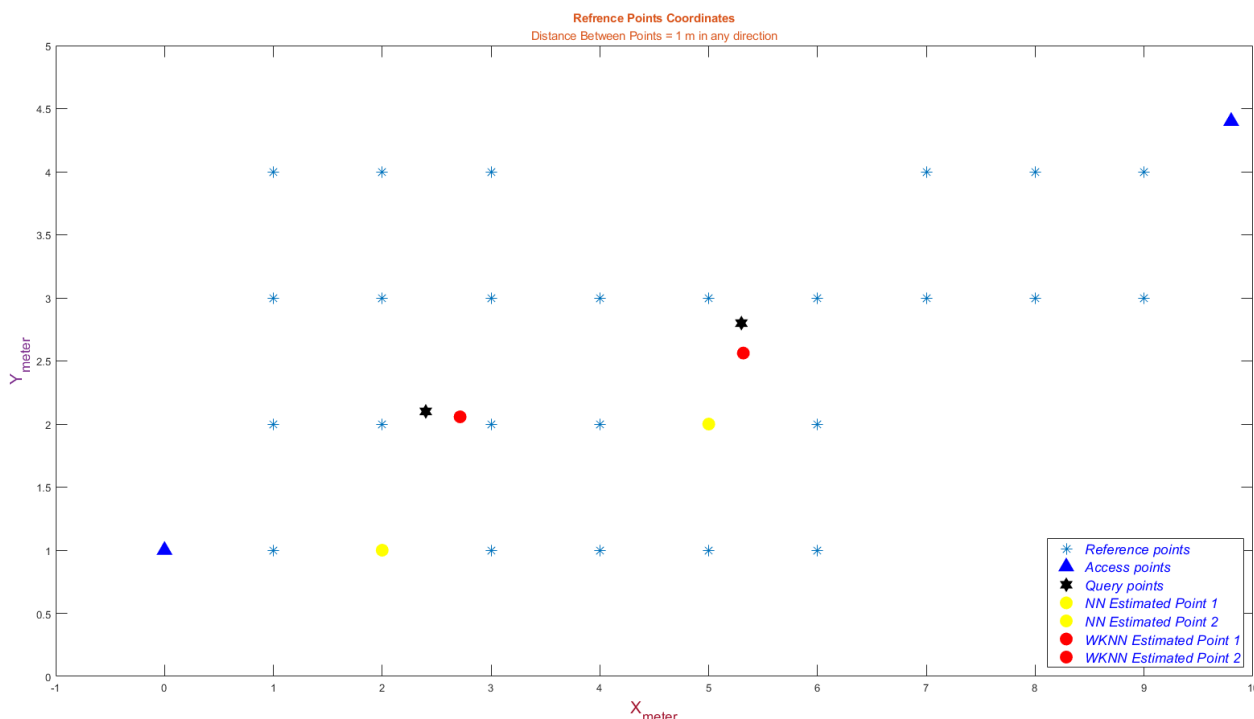


Figure 8. A view of determining the position of the Target points by different methods (NN and WKNN) in the first Scenario (Blue stars are reference points, triangles are access points, black stars are target points, yellow circles are estimated points from NN method and red circles are estimated points from WKNN method)

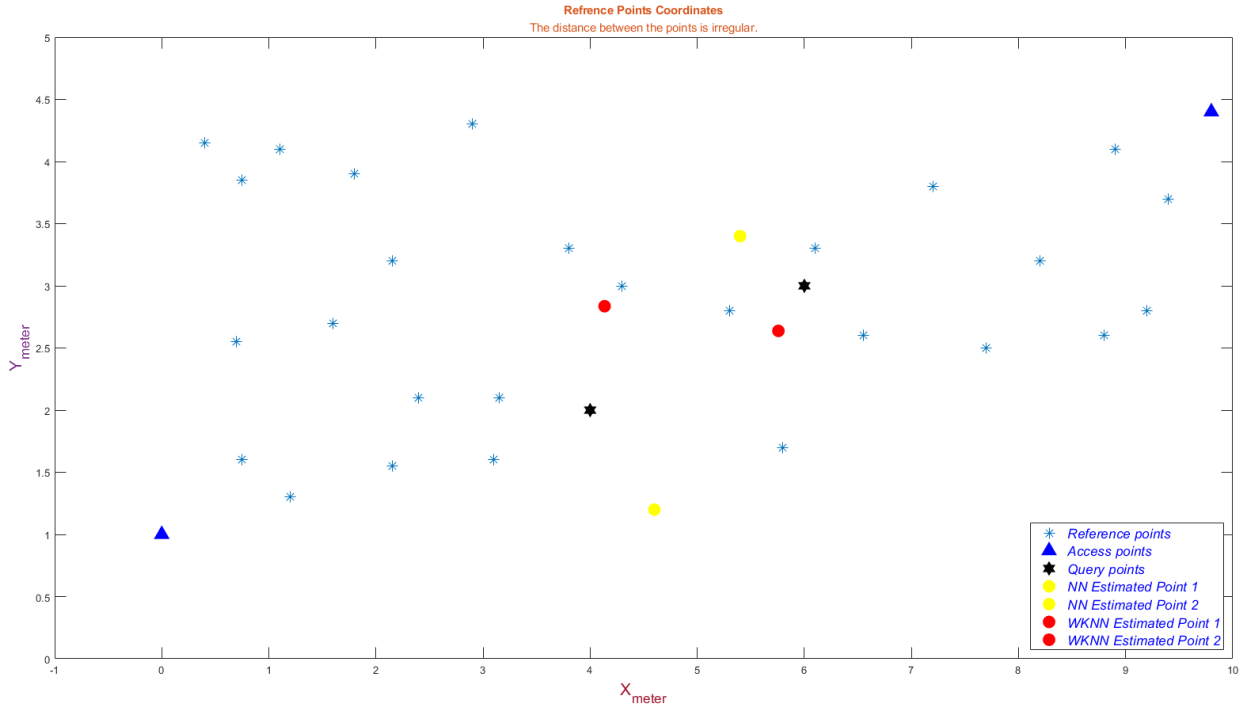


Figure 9. A view of determining the position of the Target points by different methods in the second scenario (Blue stars are reference points, triangles are access points, black stars are target points, yellow circles are estimated points from NN method and red circles are estimated points from WKNN)