Application of NN-KNN Algorithm Combined with Cloud-based Indoor Positioning

Haitao Bai, Wenyao Yan

Xi'an Innovation College, Yan'an University,
Xi'an 710100, Shanxi, China

Abstract
Among many practical applications, such as environmental monitoring, health monitoring and battlefield surveillance, the position information plays an important part in understanding the application context. However, the localization accuracy and the calculation overhead are a pair of contradiction. For addressing the aforesaid problem, improving the localization accuracy and reducing the calculation overhead, a two-phase (offline and online) indoor positioning algorithm joint position fingerprint is proposed, which is combined cloud computing theory with NN-KNN. In the offline phase, the fingerprint database is constructed; and then, counting on that fingerprint database the object positioning is executed in the online phase, which based on NN and KNN. Simulation results have shown that compared with the NN and KNN positioning algorithms, the proposed algorithm can effectively improve the localization accuracy and mitigating the computation cost in the indoor environment.
Key words: INDOOR POSITIONING, LOCATION FINGERPRINT, CLOUD COMPUTING, NN, KNN

1. Introduction

In recent years, with the development of wireless local area network associating with positioning technology and application, the indoor positioning based on location fingerprint or Received Signal Strength (RSS) has been greatly concerned. The current main bright indoor location tracking technology, A-GPS positioning, radio waves in combination with ultrasound positioning and location-based, such as received signal strength. Refers to the location where the fingerprint characteristics based on radio waves, such as accepting the RSS positioning the fingerprint, which is used to indicate the basis of a unique location.

Although time of signal arrival (ToA) and time difference of signal arrival (TDoA) can perform well outdoors, they suffer great multipath interference and other signal attenuations indoors. A large number of indoor positioning methods adopt position fingerprint. Hereinto, position fingerprint is the special relationship between a specific position and some certain measurable physical stimulation, such as RSS. Position fingerprint-based localization algorithms are normally classified into two phases, the former is offline phase and the latter is online localization phase. What's more, the popular indoor positioning schemes can be divided into fingerprinting-based and machine learning-based methods. Most of these fingerprinting-based schemes leverage the RF signals such as LANDMARC [1], RADAR [2], improved upon RADAR, ActiveCampus [3], etc. Almost those schemes need site survey over the object region to create the fingerprint database, which results in the numerous manual efforts and cost and also brings the inaccuracy coming from manual process. Therefore, the machine learning-based positioning schemes are studied, such as depending on Nearest Neighbor (NN) [4] and K-Nearest Neighbor (KNN) [5] matching algorithm. In fact, NN and KNN algorithms consider the differences between the RSS at client side and the RSS in the related fingerprint database. Since NN algorithm take directly the nearest neighbor coordinates, it is faster than KNN. Nevertheless, KNN algorithm is more accurate than NN, and its calculation cost is also lower than NN [6, 7]. In this paper, their respective merit is combined to develop a novel joint scheme.

In this paper, a two-phase (offline and online) indoor positioning algorithm joint position fingerprint is proposed, which is combined cloud computing theory with NN-KNN. In the offline phase, the fingerprint database (i.e., the RSSs and coordinates) is constructed; and then, counting on that fingerprint database the object positioning is executed in the online phase, which based on NN and KNN. Simulation results have shown that compared with the NN and KNN positioning algorithms, the proposed algorithm can effectively improve the localization accuracy and mitigating the computation cost in the indoor environment.

The rest of the paper is organized as follows. In section 2, NN and KNN are described, and then NN-KNN algorithm is proposed. The experiment and analysis are presented in Section 3. Finally, in Section 4, the conclusion is listed.

2. NN-KNN algorithm

2.1NN algorithm and KNN algorithm

In the KNN algorithm, first acceptance signal strength calculated from the online phase (RSS) sample values and position in the fingerprint database, sample values. Assuming the reference point (RP) and an access point (AP) by the number of m and n, the distance is defined as follows:

$$\left\{L_{qi} = \left(\sum_{j=1}^{n}\left|s_j - s_{ij}\right|\right)^{\frac{1}{q}},\ i = 1,2,\dots,m \right. \quad (1)$$

Where s_{ij} is the location of the fingerprint database, the j AP RSS samples in the *i* value of RP, s_j is the *j* line positioning phase AP real RSS samples [8, 9].

According to the principle of KNN algorithm, the calculated K RP having a minimum distance, the position coordinates are (x_i, y_i) (i = 1,2, ..., k) , then the test point (TP) of the position coordinate estimates RP for which the K centroid. The following equation (2) is shown.

$$(\hat{x}, \hat{y}) = \frac{1}{k}\sum_{i=1}^{k}(x_i, y_i) \quad (2)$$

In NN algorithm, taken as the nearest RP TP of position estimates, i.e. (2) where K is taken as 1, to obtain the results NN algorithm.

2.2. Location fingerprint matching algorithm to calculate streamline

In the fingerprinting-based indoor location positioning process, due to the electromagnetic wave propagation loss, RSS exponentially with distance attenuation. When the short distance transmission, attenuation faster than the transmission distance decay more slowly, when the transmission distance close to about 8m, the signal power change the transmission distance is not obvious[10,11].

According to the above characteristic, in a certain location area, TP received from an AP than 8m RSS stabilized in the vicinity of a certain value fluctuations, so this part there is no contribution to the matching location data, which indicates that, in the position of the fingerprint matching process calculation, a large number of redundant calculations. The algorithm proposed in the matching calculation ignores other than 8m AP TP from the RSS, the location of the fingerprint matching process to reduce the amount of computation. The algorithm is as following.

Suppose there is n AP in a certain location area, is set when the time TP from AP8m RSS as the critical value a, a TP is received in real time from different RSS AP is, the elements of the vector x with each threshold value a subtraction, as shown in equation (3) below:

$$\begin{cases} x_1 - a \\ x_2 - a \\ \vdots \\ x_n - a \end{cases} \quad (3)$$

Calculated results of the judgment, When $x_i - a > 0$, remain x_i, when $x_i - a \leq 0$, discard x_i. The new vector $x' = (x_{i1}, x_{i2}, \cdots, x_{im}), (m \leq n)$ by the preserved xi composition.

In matching the calculation process, Suppose a RP at the location of the fingerprint database data as $y = (y_1, y_2, \cdots y_n)$, remove the counterpart x o form a new vector which is $y' = (y_{i1}, y_{i2}, \cdots y_{im})$ Under the rules of the algorithm. Then calculates the distance between vector x' and y' [12, 13].As shown in equation (4) below:

$$d = \sqrt{(x_{i1} - y_{i1})^2 + (x_{i2} - y_{i2})^2 + \cdots + (x_{im} - y_{im})^2} \quad (4)$$

As an alternative to the traditional distance calculation between vectors x and y. Other RP at the same calculation is carried out, and finally the positioning calculation based on position location algorithm used by the fingerprint.

Matching calculation process at each reference point, the algorithm computational savings compared to the conventional method as following: Addition: 3 (nm) times, multiplication: (nm) times, since each of the RP to the database screened, calculating the amount of actual savings: addition: (2n-3m) times, multiplication: (nm) times. Throughout the positioning process, due to the irregular and random TP position selected AP arrangement, the number of AP in each TP 8m distance outside varies, assuming the amount of the savings calculation is the average value of the total p a targeted area of RP, each positioning CPC savings calculation is: Addition: p (2n-3m) times, multiplication: p (nm) times.

2.3. NN-KNN algorithm

NN-KNN jointly proposed algorithm [14, 15], using a variety of algorithms for joint position, and according to the concept of cloud computing, the online phase of the real-time positioning tasks assigned to a different server for processing, the final position of each algorithm is a weighted sum of the results, to get the final positioning results.

Complete NN-KNN joint positioning algorithm shown in Figure 1.

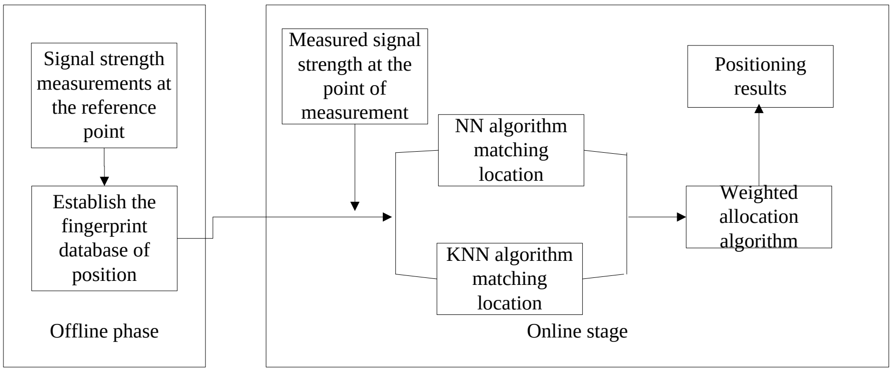

Figure 1. NN-KNN algorithm combined positioning map

First, the offline phase, the RSS measured from each AP at each of the RP and to establish the location of a fingerprint database according to the measured RSS.

The second step, the online phase, since the first measurement of TP at each AP, RSS, RSS, and then based on the location of the fingerprint database, using the above location fingerprint matching algorithm proposed streamlining the calculation using different algorithms on different server location fingerprint matching location.

The third step is to calculate the value obtained for the different positioning algorithms on different servers weighted sum TP position to get the final estimate.

3. Experiments and Analysis

3.1. experimental environments

The simulation environment of a laboratory building top of the layout chosen by respectively 44m * 32m and 14m * 12m rectangular mosaic, as shown in Figure 2. FIG gray for outdoor areas, size of 32m * 12m; white region by the separate rooms and corridors, forming region to be positioned in the region have been placed in nine AP, as shown in Figure 2, the red five-pointed star. The area to be positioned into the side length of 2m virtual grid, each grid is at the apex of RP.

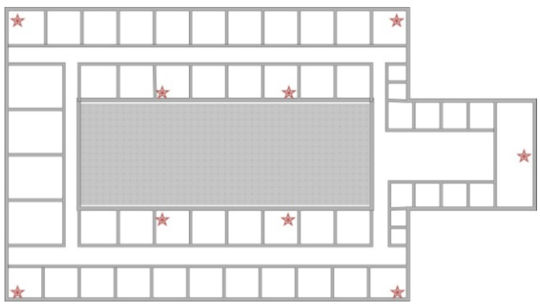

Figure 2. Experimental environment

Simulation, the location fingerprint database [16] and TP at the distance of each measuring RSS corresponding AP pitch is determined by the indoor transmission loss model, we choose the indoor transmission loss model for MK model, which belongs to the statistical model, the model considers the wall or the like of the signal attenuation, can more accurately reflect the wireless signal in the indoor environment the transmission loss, as shown in equation (5) below.

$$L_{pico} = L_0 + 10 \times n \times \log x + N_w \times L_w \quad (5)$$

In this formula, Lpico represents received signal strength, L0 represents the distance between transmitter propagation loss at the end of 1m, n is the path loss coefficient, x represents the distance transmitting end to the receiving end, NW indicates rays from the transmitter to the receiver through the wall of the number, LW representative wall loss coefficient. According to this article the simulation environment, Select L0 = 37dB, n = 2, LW = 3dB, simulation results used in this paper path loss is calculated as follows:

$$L_{pico} = 37 + 20 \times \log x + 3 \times N_w \quad (6)$$

3.2. Design of Experiment

In this paper, simulation experiments were chosen following equation (6) as the signal transmission loss model, and added white Gaussian noise 10dB SNR, the location area randomly selected 100 points as the target point. In equation (1), the paper orientation general guidelines for selecting a value of q 2, i.e. select Euclidean distance as the distance in accordance with the pattern matching indoor location positioning fingerprints.

In equation (2), KNN algorithm K will have a greater impact on the positioning accuracy, and therefore it is necessary to influence the value of K KNN algorithm to study the positioning accuracy. As the distance between the pattern matching positioning. This paper studies the change in the value of K varies between 1-10 when positioning the corresponding average error of KNN algorithm, the simulation results shown in Figure 3.

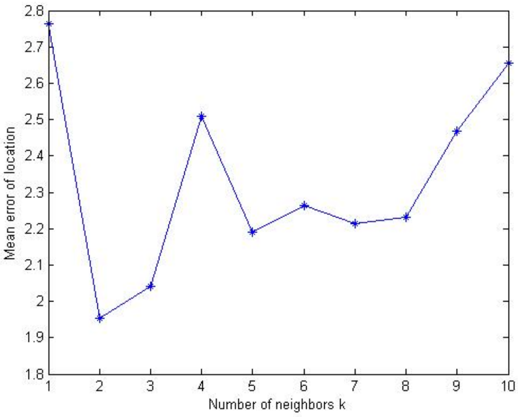

Figure 3. K different values impact on the average error of positioning KNN algorithm

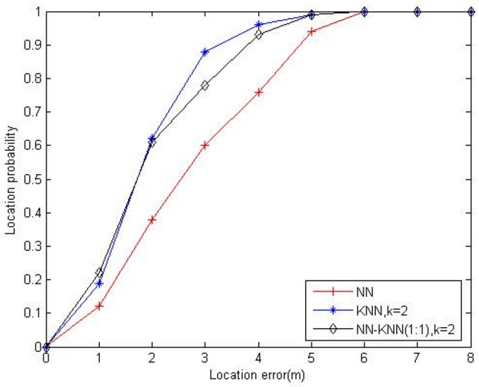

Figure 4. NN, KNN and NN-KNN algorithm positioning performance comparison (K = 2)

Simulation results show that the average error of positioning KNN algorithm with different values of K irregular changes, and K = 2 reaches a minimum. In this paper, the following experiments were selected so that the positioning error of the mean value of the smallest of the three K: K = 2, K = 3, K = 5 and compare the effects of different values of K joint positioning algorithm performance by simulation.

First, in 1: 1 ratio of the weighted distribution were studied in the K = 2 and K = 3 when comparing NN algorithms, KNN algorithm proposed joint positioning performance NN-KNN algorithm, shown in Figure 4, Figure 5.

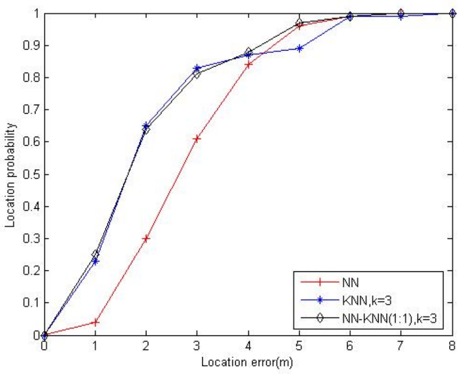

Figure 5. NN, KNN and NN-KNN algorithm positioning performance comparison (K =3)

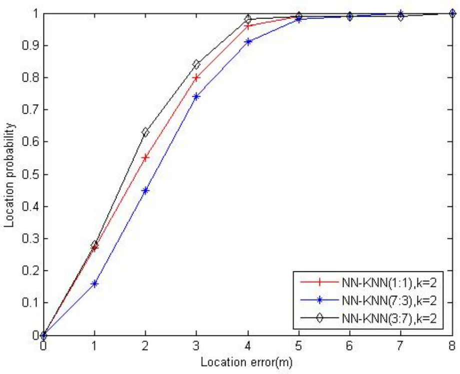

Figure 6. NN-KNN algorithm positioning performance comparison of different weight

(K = 2)

Secondly, a comparison of the weighted distribution in different proportions NN-KNN

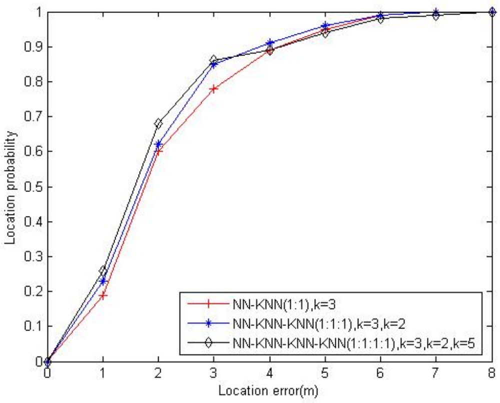

Figure 7. NN-KNN algorithm positioning performance comparison of different weights
(K = 3)

Finally, many K value for different values of K joint algorithm is studied, and choose K = 2, K = 3 and K = 5 simulation analysis, the results shown in Figure 8.

As can be seen from Figure 4 and Figure 5, in a 1: 1 ratio of the weighted distribution, NN-KNN jointly proposed algorithm performance in positioning accuracy of less than 2m or less than NN algorithm and KNN algorithm, and 2m is the most concerned about indoor positioning accuracy. In Figure 6 and 7, NN algorithm and weighted KNN algorithm, 3: 7-weighted algorithm locating the optimal performance, 1: 1, followed by a weighted algorithm positioning performance, and 7: 3 weighted algorithm positioning the worst performance in the Figure 4 and 5 above, they have been verified KNN algorithm outperforms NN positioning algorithm, which algorithm is described with a single performance of the weighted proportion joint algorithm, algorithm performance of the algorithm given a higher weighting values larger, can be improved positioning accuracy. Figure 8 illustrates the increase in the number of joint algorithm can improve the positioning performance of the algorithm.

Under different K value, it can be seen from the above simulation diagram, K = 2 when the joint positioning algorithm outperforms the joint when K = 3 algorithm positioning performance, which the aforementioned K= 2 positioning KNN algorithm outperforms K = 3:00 KNN algorithm consistent positioning performance, indicating that the value of K KNN algorithm and combined algorithms have the same impact on the trend.

algorithm localization performance, still select K = 2 and K = 3 will be described in two different situation Figure 6, Figure 7.

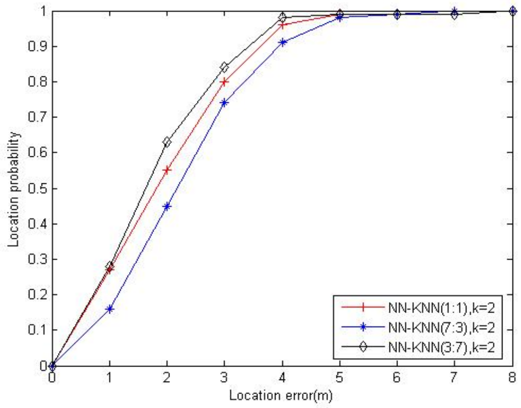

Figure 8. K value more joint positioning algorithm performance comparison

4. Conclusions

This paper proposes a joint NN-KNN algorithm based on the concept of cloud computing, and propose a method of reducing the amount of calculation of the position of fingerprint matching algorithms to streamline, the algorithm is applied to the joint. The joint algorithm, the paper comprehensive analysis of its performance compared with NN algorithm and KNN algorithms, as well as the relationship between the joint performance of the algorithm at different weights with a single algorithm performance between simulation results show that the proposed algorithm can effectively improving the performance of indoor positioning calculation process and reducing the amount of computation. In future work, the algorithm performance analysis will be carried out in other environments, as well as the design of a weighted allocation algorithm based on adaptive cloud environment.

References

1. L. M. Ni, Y. Liu, Y. C. Lau, and A. P. Patil (2004) LANDMARC: indoor location sensing using active RFID. *Wireless Networks*, 10(6):701–710.
2. P. Bahl and V. N. Padmanabhan (2000) RADAR: an in-building RF-based user location and tracking system. *Proceedings of IEEE INFOCOM*, Vol.2, p.p. 775–784.
3. W. G. Griswold, P. Shanahan, S. W. Brown, R. Boyer, M. Ratto, R. B. Shapiro, and T. M. Truong (2004) ActiveCampus: Experiments in community-oriented ubiquitous computing. *Computer*, 37(10), p.p.73–81.

4. Dasarathy, B. V. (Ed.) (1991) *Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques*, ISBN 0-8186-8930-7. IEEE Computer Society.
5. Chan, C. L., Baciu, G., and Mak, S. C. (2009). Using Wi-Fi signal strength to localize in wireless sensor networks. *The 2nd International Conference on Communications and Mobile Computing*, p.p. 538–542.
6. G. E. Hinton, R. J. Williams, et al. (1986) Representation by back propagation errors. *Nature*, 3(6), p.p. 899-905.
7. D. Sengupta, R. Iltis (1989) A neural solution to the multitarget tracking data association problem. *IEEE Transaction on Aerospace and Electronic System*, 25(1), p.p. 96-108.
8. Ahmad Hatami and Kaveh Pahlavan (2006) Comparative Statistical Analysis of Indoor Positioning Using Empirical Data and Indoor Radio Channel Models. *IEEE CCNC 2006 proceedings*, p.p.1018-1022 .
9. Yong liang Sun,Yubin Xu,Lin Ma et al. (2009) KNN-FCM Hybrid Algorithm for Indoor Location in WLAN. *Proc. of 2nd International Conference on Power Electronics and Intelligent Transportation System*, p.p.251-254.
10. Sun Peigang,Zhao Hai,Luo Dingding,Zhang Xiaodan,Yi Zhenyu(2007) Research Orientation smart space RSSI. *ACTA ELECTRONICA SINICA*, 35(7),p.p. 37-49.
11. Zhang Xiaoliang, Zhao Ping, Xu Guangqing (2013) On the application of an optimization algorithm of KNN indoor Positioning. *Electronic Design Engineering*, 21(5), p.p.110-115.
12. Tang Li, Xu Yubi, Zhou Mu (2009) Based on K-nearest neighbor algorithm WLAN indoor positioning technology research. *Computer Science*, 2009(4), p.p.51-55.
13. Zhou Kun,Xiu Chundi,Yang Dongkai (2013) Perceived probability of indoor positioning algorithm. *GPS*, 38(6), p.p.7-11.
14. ROOS T. MYLLYMAKI P, HENRY TIRRI H , et al. (2002) A Probabilistic Approach to W I AN User Location Estimation. *International Journal of wire 1ess Information Networks*, 9(3), p.p.110-112.
15. LEE D J. Campbell, M E. (2012) Iterative smoothing approach using? Gaussian mixture models for nonlinear estimation. *IEEE／RSJ International Conference on Intelligent Robots and Systems(IR0S)*, p.p. 2498-2503.
16. Tang Wensheng, Kuang Wangqiu (2008) RF indoor location fingerprint database generated spatial correlation algorithm. *Compute Engineering and Application*, 44(33), p.p.226-230.