A WiFi RSSI Ranking Fingerprint Positioning System and Its Application to Indoor Activities of Daily Living Recognition

Zixiang Ma[1], Bang Wu[1], Stefan Poslad[1]

[1] IOT2US Lab, School of Electronic Engineering and Computer Science, Queen Mary University of London, London, United Kingdom

Corresponding author: Bang Wu (e-mail: bang.wu@qmul.ac.uk)

Abstract:

WiFi RSSI (Received Signal Strength Indicators) seem to be the basis of the most widely used method for Indoor Positioning Systems (IPS) driven by the growth of deployed WiFi Access Points (AP), especially within urban areas. However, there are still several challenges to be tackled: its accuracy is often 2-3m, it's prone to interference and attenuation effects, and the diversity of Radio Frequency (RF) receivers, e.g., smartphones, affects its accuracy. RSSI fingerprinting can be used to mitigate against interference and attenuation effects. In this paper, we present a novel, more accurate, RSSI ranking-based method that consists of three parts. First, an AP selection based on a Genetic Algorithm (GA) is applied to reduce the positioning computational cost and increase the positioning accuracy. Second, Kendall Tau Correlation Coefficient (KTCC) and a Convolutional Neural Network (CNN) are applied to extract the ranking features for estimating locations. Third, an Extended Kalman filter (EKF) is then used to smooth the estimated sequential locations before Multi-Dimensional Dynamic Time Warping (MD-DTW) is used to match similar trajectories or paths representing ADLs from different or the same users that vary in time and space In order to leverage and evaluate our IPS system, we also used it to recognise Activities of Daily Living (ADL) in an office like environment. It was able to achieve an average positioning accuracy of 1.42m and a 79.5% recognition accuracy for 9 location-driven activities.

Keywords:

WiFi; Indoor Positioning System (IPS); RSSI (Received Signal Strength Indicator); Genetic Algorithm (GA), Kendall Tau Correlation Coefficient (KTCC); Convolutional Neural Network (CNN); Kalman filter (KF); Dynamic Time Warping (DTW); Activities of Daily Living (ADL); Internet of Things (IoT)

Introduction

Indoor Positioning Systems (IPS) are increasingly needed as part of our daily life, as we increasing spend 87% of our time indoors [1] in increasingly more complex 3D spaces

[2]. IPS applications include: smart building services, mobile asset or people tracking [3], eHealth [4], location-enabled games, and customised advertisements. Although a range of IPS techniques exist, they face different limitations in different scenarios regarding ease of use, requiring the use of or not of dedicated signal transmitters and receivers, power consumption, cost and the heterogeneity of signal Radio Frequency (RF) receivers. This research targets using existing personal communication devices such as smartphones for indoor positioning rather than using additional specialised signal receivers. Under these requirements, WiFi or BLE-based positioning systems are the most widely adopted techniques. These RSSI techniques, from which position can be derived, can be divided into two types: First, the more straightforward to deploy, BLE or WiFi RSSI ranging, which uses path loss models can be used to calculate the user location. Even though these are easy to deploy, their accuracy can be severely decreased if there are obstructions, i.e., a non-free space, severely reducing the positioning accuracy. Second, WiFi and BLE RSSI location fingerprinting can be used, which requires collecting a fingerprint or map of sets of RSSI measurements at reference locations in an offline phase for latter comparison with RSSI signals in an unknown location in the operational phase. Unlike BLE-based systems, e.g., using iBeacon devices, WiFi location fingerprinting system does not need to deploy additional, specific, Access Points (AP). Moreover, scanning of barcodes and Quick Response (QR) Codes at known locations can be used as a crowdsourced way to collect and update RSSI fingerprints, i.e., at known payment spots, which can significantly reduce the fingerprint collection work. However, these types of RSSI use are also affected by the heterogeneity of the signal RF receiver in smartphones. Rather than use the raw RSSI measurements as the fingerprints for the radio map database, the ranking of RSSIs from different APs can be used instead, this mitigates against the heterogeneity of the RF receivers when constructing such fingerprints. `Another positioning method proposed in [5], applies unsupervised learning to handle the heterogeneity of RSSI receiver hardware to improve the location accuracy. However, over the growing numbers of deployed APs, an RSSI ranking of those APs can be applied, instead of applying additional unsupervised learning to process those measurements. KTCC can be used to measure the ordinal association between two measured quantities (concordant and discordant ranking pairs). In addition, CNN is good at feature extraction. RSSI ranking of different APs can be used as a feature to match patterns that can be applied in ranking-based fingerprinting systems. In addition, an AP selection algorithm to select a subset of all available APs, not only reduces the computational cost for online positioning but also increases the positioning accuracy, e.g., because some AP RSSIs are noisy or weak [6].

An IPS can be used to track the trajectories, paths or tracks of sequential locations of users, enabling indoor navigation services. Since the trajectory is based on sequential movement, motion models can be applied to help increase the tracking accuracy, e.g., through using a particle filter or Kalman filter. A further enhanced application of IPS navigation is that predefined hotspots or waypoints can be combined with the trajectories, and used to recognise location-driven human ADLs. There is a challenge in that different users can vary the trajectories in time and space. Multi-Dimensional Dynamic Time Wrapping can be to used to match paths from different or the same users that vary in time and space [7].

WiFi Channel State Information (CSI) can be used to recognise more fine-grained ADLs, e.g., Wi-Vi [8], WiSee [9] and WiTrack [10]. The principle of such CSI-based sensing methods is to make use of channel information in the time and frequency domain, e.g., the amplitude and phase of each subcarrier at each timestamp and to leverage how these changed features are affected by human activity between transceivers. By collecting and extracting these features at the training stage, test features can be matched with the training database to infer human motions and activities. However, such CSI-based methods rely heavily on a relatively stable RF environment, i.e., they often rely on a single occupant environment and no infrastructure or environment changes. Each occupancy or environment change would trigger an activity profile update or it may fail to work. Its recognition accuracy is reduced if there are more than two people in the same space. Although, vision-based methods can offer positioning and ADL recognition services, these face problems of being privacy invasive and are computationally intensive. Our purpose is to find the limitations of path matching to recognise location-driven ADLs to recognise coarse-grained activities such as to detect if someone went to the kitchen and dining area and stayed there. It's considered to be out of scope to consider more finely-grained activities such as taking a cold drink from a fridge or making tea.

The main contributions of this paper to build a novel RSSI ranking-based IPS that can also be used to recognise location-driven ADLs are three-fold:

1. Before constructing our radio map, our proposed AP selection based on a Genetic Algorithm (GA) is applied to reduce the positioning computational cost but also to increase the positioning accuracy.

2. Our novel RSSI ranking-based KTCC/CNN is used after our AP selection to improve positioning accuracy and to mitigate against signal receiver diversity. This technique could also be suitable for scenarios with dense Internets of Things, as transmitters embedded in smart objects can be treated as APs.

3. An Extended Kalman filter is then used to smooth the estimated sequential locations. Then, to the best of our knowledge, no one has in addition, used Multi-Dimensional Dynamic Time Warping (MD-DTW) to match similar trajectories or paths representing ADLs from different or the same users that vary in time and space. This can also be extended in multi-source signal scenarios, e.g. combined with magnetic field sensing, to upgrade the path or activity recognition accuracy.

Methodology

There are usually two phases (offline training phase and online positioning phase) used by location fingerprinting methods. In our IPS, we also have such two phases, but after we collected the fingerprints, we select APs to improve location accuracy as some are noisey. In our previous work [6], which we proposed an Interval Overlap Degree determination (IOD) method to do the AP selection. However, this method is based on the raw RSSI measurements, which is not suitable for ranking-based AP selection. Instead of using IOD, we propose another AP selection algorithm based upon a GA to

select optimal APs. Then the converted ranking fingerprints of the selected APs are be used to train our KTCC and CNN models to estimate the location of users.

In the online phase, the matching probability of using KTCC and CNN will be treated as weights of using Weighted K-Nearest Neighbour Classification Algorithm (WKNN) to derive the user location.

Figure 1 depicts all procedures of our location fingerprinting IPS, which we use AP appearance ratio before the AP selection algorithms and then, after estimating user's location, we applied filtering algorithms to help increase our positioning accuracy. Then, then combined path fingerprints will be used to recognise activities.
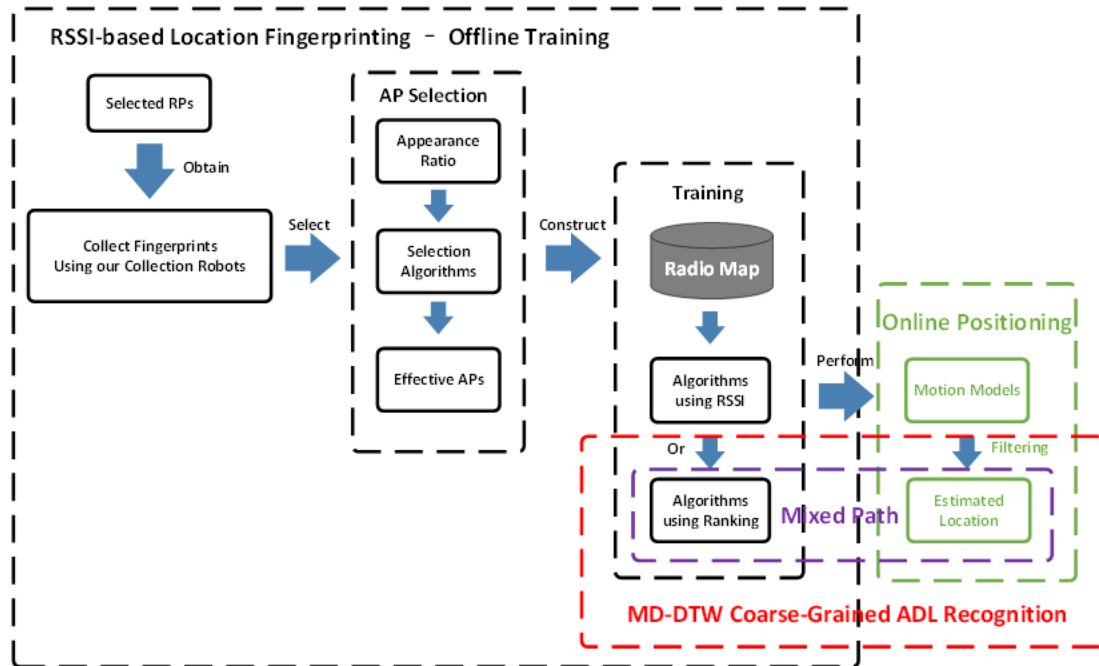

Fig. 1. Phases of our location fingerprinting method

Location Determination using Euclidean distance (ED)

In most positioning methods, ED is used to match the RSSI mean vector collected at an unknown location, with the RSSI mean vector at Reference Points (RPs) in radio map. The distance is expressed using the following equation:

$$ED_{im} = \sqrt{\sum_{j=1}^{n}\left(AVG\_RSSI_{i,j} - AVG\_RSSI_{m,j}\right)^2} \qquad (1)$$

where ED_i is the Euclidean Distance between a RP i and the unknown point m. AVG_RSSI_i is the average of the RSSI of RP i in the pre-recorded radio map. AVG_RSSI_m is the averaged RSSI vector at m, and n is the number of selected APs.

Then WKNN is applied to drive the user location, which is described as:

$$(\bar{x}, \bar{y}) = \frac{\tau_1}{\Sigma_K^1 \tau}(x_1, y_1) + \frac{\tau_2}{\Sigma_K^1 \tau}(x_2, y_2) + \ldots + \frac{\tau_k}{\Sigma_K^1 \tau}(x_k, y_k) \tag{2}$$

where we use $1/ED_k$ as the weight τ_k.

Location Determination using PCA

Instead of choosing a subset of APs, [11] replaces the elements with a subset of Principal components (PCs), using a statistical procedure, where PCs are obtained by a transformation of the measured RSSI. The theory of the paper is based on Principal Component Analysis (PCA) to find an effective transformation such that the retained information in the chosen PCs can be maximised, which means it will not only reduce the computation cost but also increase the positioning accuracy by directly calculating the ED distance using the transformed RSSI.

$$\begin{pmatrix} y_1 \\ \vdots \\ y_L \end{pmatrix}_{L*1} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1D} \\ a_{21} & a_{22} & \ldots & a_{2D} \\ \vdots & \vdots & \vdots & \vdots \\ a_{L1} & a_{L2} & \ldots & a_{LD} \end{pmatrix}_{L*D} \begin{pmatrix} x_1 \\ \vdots \\ x_D \end{pmatrix}_{D*1} \tag{3}$$

The concept of using PCA, which is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components, is to reduce the dimensions by combining APs. In other words, information reorganisation is adopted, rather than AP selection. As shown in equation 3, the principal components $[y_1, y_2 \ldots y_L]^T$ are produced by a transformation with real numbers. With appropriate transformation matrix A, the information transmitted into Y from X can be maximized. Matrix A is determined by using PCA, as it has the distinction of best presenting data in a least square sense, i.e., $Y = AX$, $\hat{X}$ is the reconstruction of X from Y, $\hat{X} = A^T Y$, PCA seeks to minimize the mean square reconstruction error:

$$J_\epsilon = E\left\{\left\|X - \hat{X}\right\|^2\right\} \tag{4}$$

The property guarantees that using PCA can retain the maximised information when dimensions are reduced. PCA can be done by using eigenvalue decomposition of a data covariance matrix or singular value decomposition of a data matrix, usually after mean centring (and normalising or using Z-scores) the data matrix for each attribute.

Location Determination using KTCC

Instead of using ED to match the RPs, we propose to use KTCC of RSSI ranking as the distance to match RPs. How to calculate KTCC can be expressed as (5) or (6):

$$\tau_1 = \frac{(\text{number of concordant pairs}) - (\text{number of discordant pairs})}{n(n-1)/2} \tag{5}$$

$$\tau_2 = 1 - \frac{2(\text{number of discordant pairs})}{n(n-1)} \tag{6}$$

where we consider the higher ordinal association τ between two quantities, the closer they are. Figure 2 visualised the concordant (blue) and discordant (yellow) pairs in a ranking fingerprint of five selected APs at two RPs. The blue parts represent the concordant pairs; the yellow parts are the discordant pairs.

For our KTCC method, we also use WKNN to derive the user's location. However, we use τ as the weight. τ_2 is also used as it can offer positive weight.
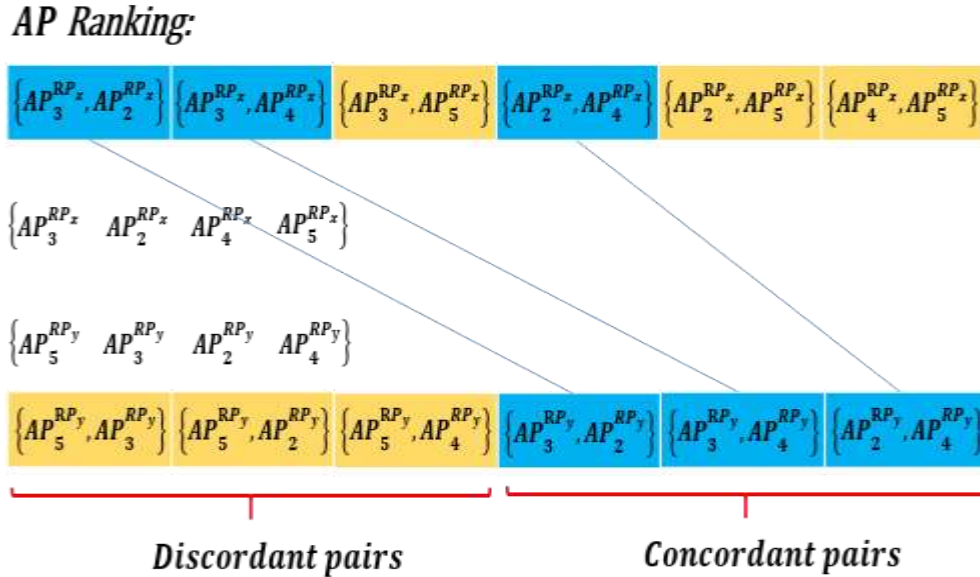

Fig.2. *Concordant* and *Discordant pairs in KTCC*

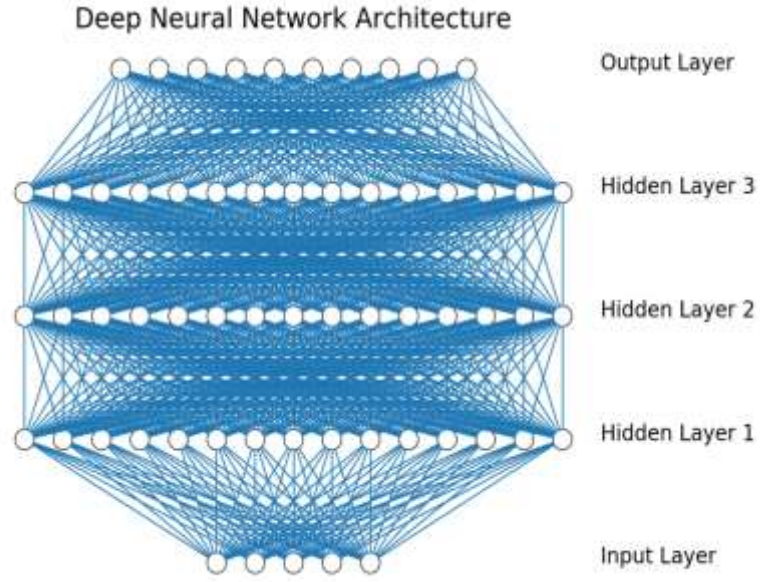# Location Determination based on Ranking-based CNN

Fig. 3. *An example of a Deep Neural Network Architecture (The input layer with 5 inputs, 15 neurons in each hidden layer, 10 outputs in the output layer)*

Deep neural network models have been applied successfully to solve complicated problems, e.g., facial recognition, driverless car, and machine translation. CNN is good at feature extraction. Our previous work [12] has tested that using KTCC can mitigate the heterogeneity issue, which means the ranking relation between APs can help to match mobility paths as patterns. So, we propose that CNN can also be used to extract such features. Furthermore, several IPS employ deep learning to estimate user location [13]. However, most of those methods just use raw RSSI measurements as the input data to train their positioning models, which does not consider the RF heterogeneity receiver impact and the consequent need to collect data from different smartphones. To mitigate against this hardware heterogeneity issue, RSSI ranking is used as our input.

For classification problems using deep neural network models, it is common to use a so-called softmax layer at the top of the network. For example, given 3 possible classes, the softmax layer has 3 nodes denoted by p_i, where $i = 1, 2, 3$. p_i specifies a discrete probability distribution, therefore, $\sum_i^3 p_i = 1$. Let h be the activation of the penultimate layer nodes (k), W is the weight connecting the penultimate layer to the softmax layer, the total input into a softmax layer, given by a, is

$$a_i = \sum_k h_k W_{ki} \tag{7}$$

$$p_i = \frac{\exp(a_i)}{\sum_j^3 \exp(a_j)} \tag{8}$$

Then, the outputs of softmax layer p_i will be treated as the weight in WKNN to estimate the user location.
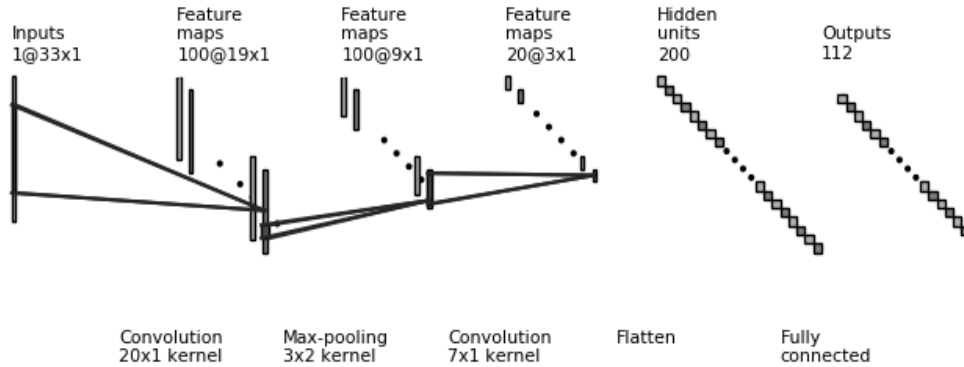
Fig. 4. *The architecture of our CNN Model (one example)*

Figure 4 shows the architecture of the CNN model. Keras[1] (Tensorflow as a backend) was used to implement the CNN model. The model consisted of two 1D convolutional layers and one max-pooling layer. After the convolutional layer, the model has a fully connected layer that is used to connect to the Softmax layer. The outputs of the Softmax layer become the weights in the location determining phase.

For example, after using the GA-based AP selection algorithm, 33 APs are selected. Then, the model inputs are a ranking of these APs. After the hyper-parameters (e.g., layer number, learning rate), were tuned, the following structure was found to be effective. The first convolutional layer has a depth of 100 and a filter size of 15. The filter size of the pooling layer is 3 with a stride of 2, which will mitigate against the overfitting issue of the model. The filter size of the convolution layer is 7, and the depth is 20. Next, the fully connected layer has 200 neurons, and tanh, as a popular activation choice, is used as the non-linearity function. Finally, the softmax layer is used to classify the test points.

Probability Comparison

In our IPS, we use WiFi RSSI ranks as our data input rather than using the raw RSSI measurements. Instead of using ED with RSSI values to compare which is stronger in the pairwise APs, we use a joint distribution. Let A represent the random variable for the RSSI measurements of AP A in the chosen time frame, and B the corresponding random variable for AP B. Assuming the APs are independent then we choose AP A over AP B if $P(A>B)>0.5$. More specifically, S1 is the set of RSSI values of AP A, and S2 is the set of RSS values of AP B in the time frame sampled.

$$P(A > B) = \sum_{r1 \in S1} \sum_{r2 \in S2} P(r1 > r2 \,|r1 \,\&\, r2) \qquad (9)$$

The reason why we chose this approach is that based on paper [14]. The RSSI measurements fluctuate over time, so it is better to use a probabilistic approach.

---

[1] For more details see: https://keras.io/.

AP Selection Based on Appearance Ratio

In our experimental space, a total of 106 APs is detected in our pre-constructed radio map; however, if we use all APs, this will not only increase the computational cost to estimate locations but also reduce the location estimation accuracy. the accuracy of AP selection algorithms, would also suffer from imbalanced RSSI measurements, as not all RSSI measurements of 106 APs can be recorded for each scan. For example, we collect fingerprints at each RP for 20 times, not all APs can be detected at each time. If AP_a is detected once, we will keep the measurement, if not, a value, e.g., -100 will be kept. In this case, if we directly use AP selection algorithms, those AP selection algorithms would select the missing APs with more -100, as it has a better performance and is more stable, which would otherwise decrease the positioning accuracy.

To solve this, we propose to select the APs using an appearance ratio first (when we set this ratio to 90%, 67 APs are selected), then using our AP selection algorithms on those selected APs. The idea of appearance ratio is simple, as at each RP, we recorded several scans, then we based on its appearance ratio (appearance times in every second divided by total scanned seconds) to select APs at each RP, then combine all selected APs then together, which is in order to mitigate the filling values.

AP Selection Based on Genetic Algorithm

Instead of selecting a specific number of AP, we use Genetic Algorithm (GA), an optimisation algorithm, to select the desired number of APs from all detected APs. The principle of GA is simple, which is based on natural selection, the process that drives biological evolution.

Fig. 5. The procedures of using the GA algorithm

The GA repeatedly modifies a population of individual solutions. At each step, the GA selects individuals at random from the current population to be parents and uses them to produce the children for the next generation. After consecutive generations, it would evolve toward an optimal solution. The primary procedures of our GA are as follows:

Selection

This procedure is to select the individuals (parents), which contribute the population to the next generation. In our case, we randomly generate a specific number of selected AP list, then convert them to a binary vector, 1 means selected, 0 means this AP is not selected.

Individual solutions are selected using a fitness function, where fitter solutions will have more chance to be selected. Here, we use the performance of our validation dataset to measure the fitness; it can also be combined with Information Gain (IG), mutual information (MI), which is mentioned in our previous paper [6].

num	AP_1	AP_2	AP_3	$\cdots$	$\cdots$	AP_{n-2}	AP_{n-1}	AP_n
1	0	1	1	$\cdots$	$\cdots$	0	0	1
2	1	1	1	$\cdots$	$\cdots$	1	1	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

m	0	0	0	...	...	0	1	1

This procedure combines two parents to form children for the next generation. Based on our assumption that the excellent performance of positioning accuracy using KTCC is correlated with the specific sub-groups of all APs. If in this case, it also makes GA perfectly fit for the ranking-based AP selection, as in the crossover procedure, they will change the sub-groups (DNA segment) to achieve a better positioning performance (new generation).

Mutation

In this procedure, random changes will be applied to those individual parents to form children. Namely, we randomly change the single segment from 1 to 0 or 0 to 1 based on the mutation rate.

Extended Kalman Filter

Extended Kalman Filter (EKF) is an extension of Kalman filter (KF). Although KFs [15] has been applied in a wide range of areas, in this case for navigation by acting as a filter that can smooth the point sequence to increase the indoor positioning. Some locations as points, because of noise, may be far away from a set of sequential locations or a path or trajectory. So, we can use the status (e.g. speed, direction) estimated from the first few points in this path to predict the status of remaining points. Then all points would be closer to the path. KF is used to smooth the possible low-accuracy points which are far away the path. KF has a constraint that the dynamic system must be linear. However, this condition is hard to satisfy since the dynamic systems in the real world are always complicated and cannot be summarised with a linear function. EKF is developed for such a nonlinear system.

EKF is able to estimate and update the states of a nonlinear system by linearising it with the help of first-order Taylor-expansion. In other words, when predicting the state in the next time step using state transition function, instead of finding the transfer matrix, we find the partial derivatives of the state transition function and observation function, represented by a Jacobian matrix. Different from the KF implementation, the state o(t) for EKF is defined as

$$o(t) = \begin{bmatrix} p(t) \\ vel(t) \\ \theta(t) \end{bmatrix} \tag{10}$$

where $p(t) = [x, y]^T$, $vel(t)$ is a scalar indicating the velocity of the moving target; $\theta(t)$ represent the orientation, defined as the angle between the orientation and one of the axes in the moving target's inertial frame. So, the state transition equation for each parameter in the state o(t) in the 2D scenario:

$$x(t) = x(t-1) + vel(t-1) * \sin(\theta(t-1)) * \delta t \tag{11}$$

$$y(t) = y(t-1) + vel(t-1) * \cos(\theta(t-1)) * \delta t \tag{12}$$

$$vel(t) = vel(t-1) \tag{13}$$

$$\theta(t) = \theta(t-1) \tag{14}$$

By finding a matrix of partial derivatives of the fours above, we can get the linearised model of the system and estimate or update the states of the moving target. After path smoothing using EKF, MD-DTW will next be used to detect and match similar paths.
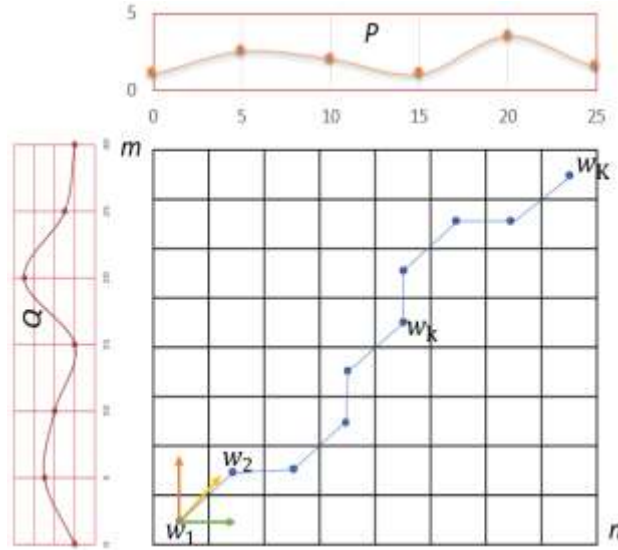
Path Matching-based ADL Recognition

Fig. 6. An example of warping path

The rationale for using path matching to recognise ADLs is threefold: 1) A trajectory is more distinctive than a position; 2) Most office corridors are long and narrow pathway; 3) Most users follow a straight walking direction [16]. Hence, Dynamic Time Warping (DTW) can be used to compare two-time series with different lengths to improve the recognition and classification our specific defined trajectories that represent activities [17]. Suppose two-time series P and Q, $(P = p_1, p_2, \dots p_i, \dots p_n, Q = q_1, q_2, \dots, q_j, q_m)$. To align, matrix can be stablished. Each element of the matrix stands for the distance (typically the Euclidean distance, but in our case, KTCC is used) between p_i and q_j. The main target of DTW is to find the warping path W ($W = w_1, w_2, \dots w_k, \dots w_K$). As shown in Figure 4, W starts from w_1 and end at w_K ($\max(n, m) < K < m + n - 1$). Then the path direction only has three cases at each w_k. The step length is only restricted in the warping path to adjacent cells. The distance between P and Q is $\frac{1}{K}\Sigma_1^K W_k$.

This ordinary DTW can effectively classify two time- series from a single source. However, in many cases, signals have multi-dimensions will be employed at the same time which can be used to better align two time-series compared to 1-dimension. Therefore, [7] proposed an improved MD-DTW method which was used in gesture

recognition conducted by the camera. Assume P^S and Q^S are two time- series S-dimension features, i.e. $P^S = p_1^S, p_2^S, \ldots p_i^S, \ldots p_n^S$, $Q^S = q_1^S, q_2^S, \ldots q_j^S, q_m^S$. Then the elements of distance matrix turn $Dis_{n,m} = \sum_{s=1}^{S}(p_i^S - q_j^S)^2$. Considering the range difference of S-dimension feature values, normalizing each dimension of feature values can effectively remove the difference. Usually, we normalize each dimension of feature values separately to a zero mean and unit variance.

[18] uses DTW with the magnetic field to predict the personalised route, [19] uses DTW to improve the positioning accuracy. However, most researchers focused on using DTW to improve the IPS accuracy; no one has used such combined measurements to recognise ADLs. One reason is that with current positioning accuracy, it may not achieve a good recognition accuracy. In our research, we aim to recognise coarse-grained ADLs using combined estimated locations and RSSI ranking based fingerprints. Since trajectories of human mobility contain the cue to infer ADLs, path matching can help us to find the closest trained path as well as its corresponding activity. Two kinds of WiFi fingerprint matching methods to localise are already introduced above. Compared to the location path (path of ground truth or estimated locations) matching scheme, ranking-based matching adds a matching process of RSSI ranking-based MAC address vectors from multiple points. Hence, fused matching can be implemented by using two different variables, estimated locations and sorted RSSI ranking MAC address vectors. Their feature matrix F_C and F_R are denoted as below.

The x_l and y_l are estimated locations, instead of the ground truth locations. As based on our experiments, by using estimated locations, we have achieved a higher (more than 20%) recognition accuracy using estimated locations than using ground truth locations. The reason we assume is that the bias of estimated location could also help to increase the recognition accuracy, namely, the designed path matching accuracy.

$$F_C = \begin{bmatrix} x_{l_1} & y_{l_1} \\ x_{l_2} & x_{l_2} \\ \vdots & \vdots \\ x_{l_q} & x_{l_q} \end{bmatrix} \quad F_R = \begin{bmatrix} MAC_{l_1}^1 & MAC_{l_1}^2 & \cdots & MAC_{l_1}^N \\ MAC_{l_2}^1 & MAC_{l_2}^2 & \cdots & MAC_{l_2}^N \\ \vdots & \vdots & \vdots & \vdots \\ MAC_{l_q}^1 & MAC_{l_q}^2 & \cdots & MAC_{l_q}^N \end{bmatrix} \tag{15}$$

However, past work only focused on path matching using one of these two feature vectors, F_c or F_R. Fusing different signals can offer improve the location accuracy. What's more, it also can be extended into scenarios with multi-source information. For example, smartphones can not only collect RSSI measurements but also other measurements, e.g., magnetic field measurements F_{MF}. Therefore, the multi-source fusion feature matrix can be denoted by F_F. In this case, we only fuse F_c and F_R which can be viewed as a good test for multi-source feature path matching.

$$F_F = \begin{bmatrix} F_{Pos}^1 & F_R^1 & F_{MF}^1 & \cdots \\ F_{Pos}^2 & F_R^2 & F_{MF}^2 & \cdots \\ \vdots & \vdots & \vdots & \cdots \\ F_{Pos}^q & F_R^q & F_{MF}^q & \cdots \end{bmatrix} \tag{16}$$

The fusion process is as follows. Firstly, as the range of KTCC results is from 0 to 1, instead of standardising each type measurement, we normalised the distance between the

estimated coordinates. Min-max normalisation can map the distance values to the range of 0 to 1 to and make each type of result be within the same scale. Equation (12) shows the normalised the distance ($\tilde{z}$). Then, in the phase of path matching, MD-DTW (uses the sum of the distances of each type measurements as a new distance) is utilised to find the best match. In our case, to recognise the trajectory of human mobility, it needs to find the path which minimises the distance between the current path and the pre-trained path (i.e., $MD - DTW = \text{argmin}\{\frac{1}{K}\sum_{1}^{K} W_k^S\}$).

$$\tilde{z} = \frac{z - \min(z)}{\max(z) - \min(z)} \tag{17}$$

where z equals to $sqrt\left(\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2\right)$. i and j represent different index of estimated coordinates.

Experiment Setup

Location fingerprinting

A field experiment was conducted in a 13.0 m × 30.0 m PhD office of the Queen Mary University of London. The training dataset was collected using smartphone (Nexus 5). Validation and test datasets were collected manually by holding the Nexus 5 at mid-body height in front.
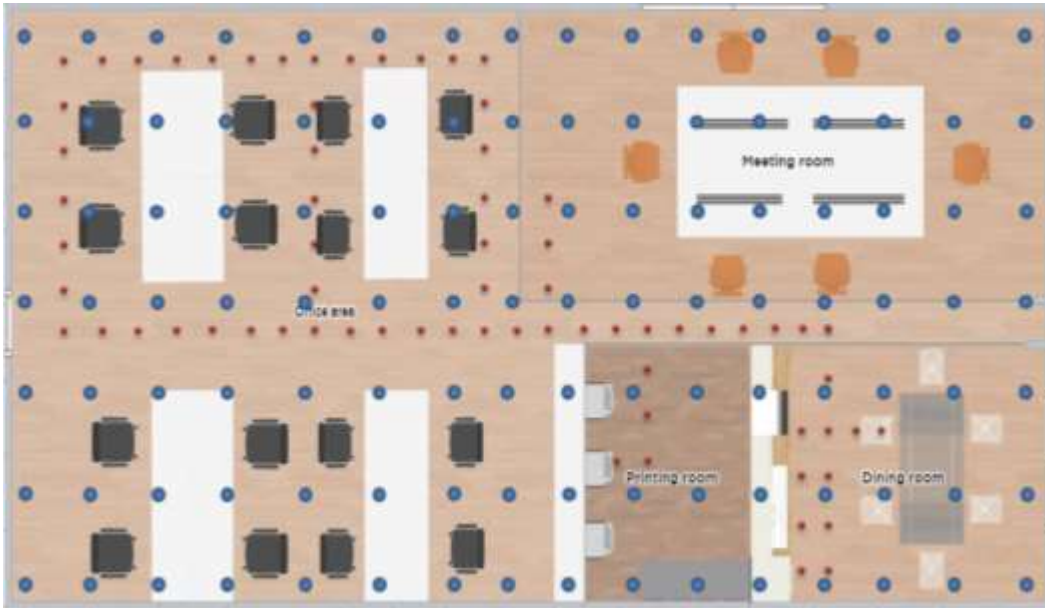
Fig.7. *The ground floor of office*

Over 106 APs (106 different MAC addresses) were detected during the whole training data collection procedure (we believe some of them are dummy ones), hence this is also one reason why we need AP selection algorithm).

At each RP, the varying RSSI measurements (and hence varying ranking) was collected for 40 seconds. We use the ranking of collected RSSIs in two second as each training data for our model. Hence each training RP has 20 RSSI ranking vectors.

The 113 Reference Points (RPs) shown in Figure 7 are our training data, which is used to train our models; Our validation dataset is the rest 70 Test Points (TPs), which is used to validate our system (the layout fits for the ADL recognition). We collected data from those TPs 20 times, half of them will be treated as test dataset, the rest will be treated as validation dataset. Figure 7 also shows the layout of the office. The blue dots are RPs, which are 1m apart. The red dots are TPs, the distance between each TP is based on the length of each step.

Fig. 8. Our fingerprints collection application

Fig. 9. Our fingerprints collection drone (prototype)

Figure 8 and 9 show our fingerprints collection application and drone, as another key issue of location fingerprinting is that the time-consuming and laborious collection work. To solve this, we developed the fingerprints collection drone. However, it is not fully autonomous, e.g., to keep it fly correctly within the designed path, which we are still improving it. The fingerprints collection application will collect raw RSSI measurements and simultaneously convert them to a ranking ordinal vector using probability comparison. As we also mentioned that it is onerous to collect and update the radio map. Currently, we set our drone as a carrier to carry the smartphones to collect fingerprints follow the lines at a fixed 1m height.

ADL recognition

Fig. 10. *Overview of class 1 activity*

For the ADL recognition part, we designed 9 specific activities which are listed in Table 1. Those 9 activities are divided into two classes relying on the length of the paths.

The activities among the Class 1 with long paths are shown in Figure 10, and the Class 2 activities with short paths are shown in Figure 11 (in the kitchen). Since the activity recognition accuracy is related to the path length and dimension of feature matrix, these two classes activities can constitute a comparative test for the effect of path length on recognition accuracy. As shown in Figure 8, there are different directions (colours) from user desk to other places (exit, meeting room, printing room and kitchen). Therefore, there are 3 paths for each activity among class 1, 12 paths in total. For activities among

class 2, each of them has 1 path, 5 paths in all. During the experiment, we repeatedly collected activities for 20 times so that there were 17*20=340 paths.

Fig. 11. *Overview of class 2 activities*

Table 1. List *of* activities

Classes	ID	Steps	Activities	Path
Class A	1-3	15/13/25	Leave the office	From user desk to exit
	4-6	16/18/32	Have a meeting	From user desk to meeting room
	7-9	20/22/36	Print documents	From user desk to printing room
	10-12	27/29/43	Go to kitchen	From user desk to kitchen
Class B	13	4	Eat food	From microwave to dinner table
	14	3	Make tea	From fridge to kettle
	15	5	Drink tea	From kettle to dinner table
	16	4	Heat food	From firdge to microwave
	17	7	Have a drink	From fridge to dinner table

Evaluation

Location fingerprinting

In our pre-experiment, we have used two different types of radio map; one is using averaged RSSI measurements at each RP, which makes a relatively stable feature of each RP, another type is just to use the multiple scans measurements, which maintains data diversity. Then, we undertook repeated estimating locations by changing the parameter K

in WKNN, which is shown in Figure 12. It shows that the positioning accuracy is related to the k, and we will choose the best performance k for each method using the validation dataset. The best performances of these four methods are 2.57m, 3.74m, 2.44m, and 2.18m, respectively. It also shows ED using averaged RSSI measurements performs better than using multiple measurements, and KTCC using ranking based on a total probability comparison performs worse than KTCC using multiple measurements. This maybe because KTCC may be sensitive to varying RSSI, which causes a varying RSSI ranking. Our focus is on ranking-based methods. So, in the following experiments, we use the multiple measurements radio map for ranking-based methods.

Furthermore, the best positioning accuracy for the above methods (same order) using a total of 106 APs are 2.91 m, 3.84 m, 2.77 m, and 2.51 m, which by using an appearance ratio method only (67 APs are selected from all 106 APs using the appearance ratio), the positioning accuracy increases by 11.6%, 3.3%, and 11.9%, respectively. However, ED (using averaged RSSI measurements) using whole APs performs 15.1% better than using appearance ratio selected APs, which means ED using averaged RSSI measurements can be more robust to use.
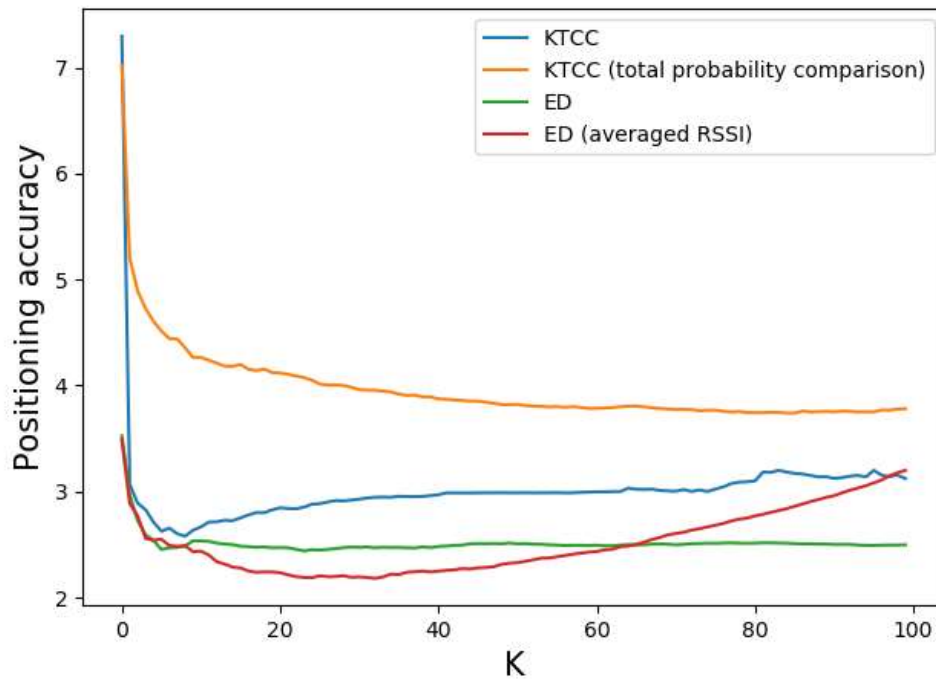

Fig.12. *Performance with using different K in WKNN*

Using an AP selection algorithm not only reduces the online positioning computational cost but also increases the positioning accuracy. We have proposed our IOD AP selection algorithm which performs better than using IG and MI [6]. However, we found that PCA performs better than our IOD algorithm based on the results shown in Figure 13 (dash-lines show the best corresponding positioning accuracies). However, it is not a 'real' AP selection algorithm; it needs to calculate PCs each online scan, which has a higher computational cost than using AP selection algorithms when doing online positioning. In

this paper, we choose to deploy PCA as our positioning baseline method to compare with our proposed ranking-based methods, as it offers a higher positioning accuracy, which the best accuracy is 2.01m when using only 2-dimensional coordinates.
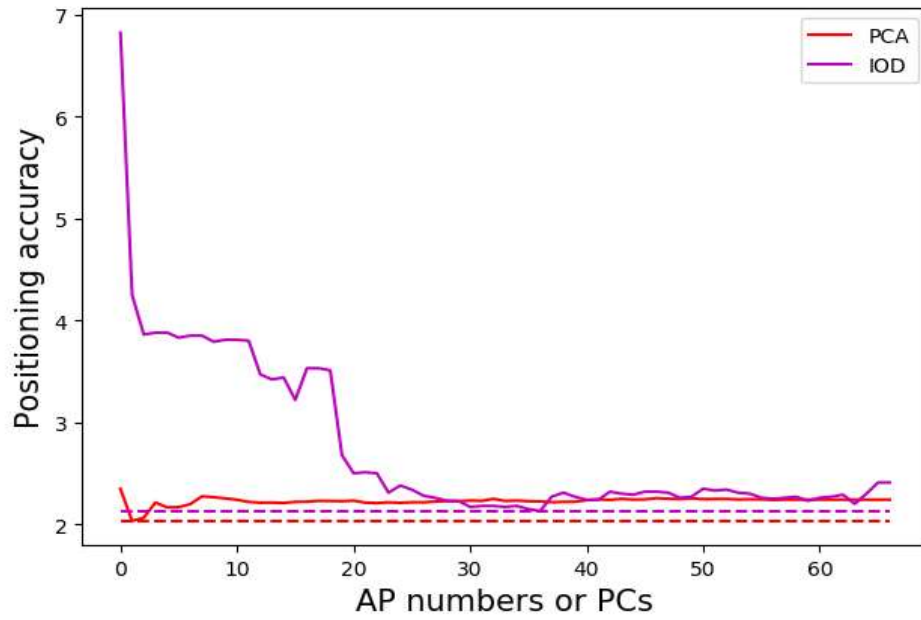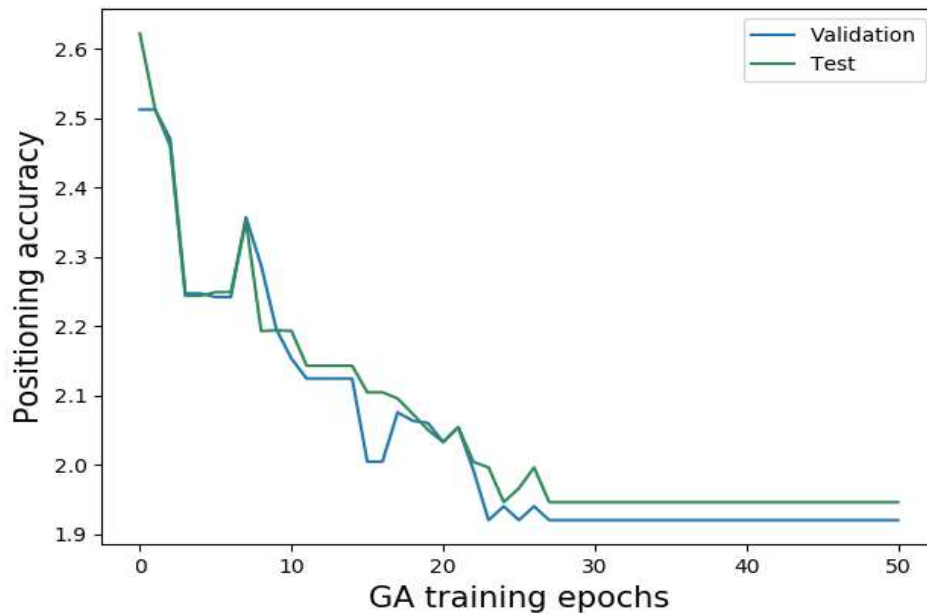

Fig.13. *Performances of using IOD and PCA*

Fig.14. *Improved positioning accuracy by using GA*

However, those AP selection algorithms are not designed for ranking-based methods. To solve this, we have proposed our GA algorithm to select APs, which can directly find the optimal AP list, instead of considering all APs.

Figure 14 shows the improved positioning accuracy (test dataset) by using our proposed GA, which its positioning accuracy converges after having 28 training epochs (validation:1.92m, test: 1.94m with using 33 selected APs). Then, the selected AP list will be used for our KTCC and CNN methods.
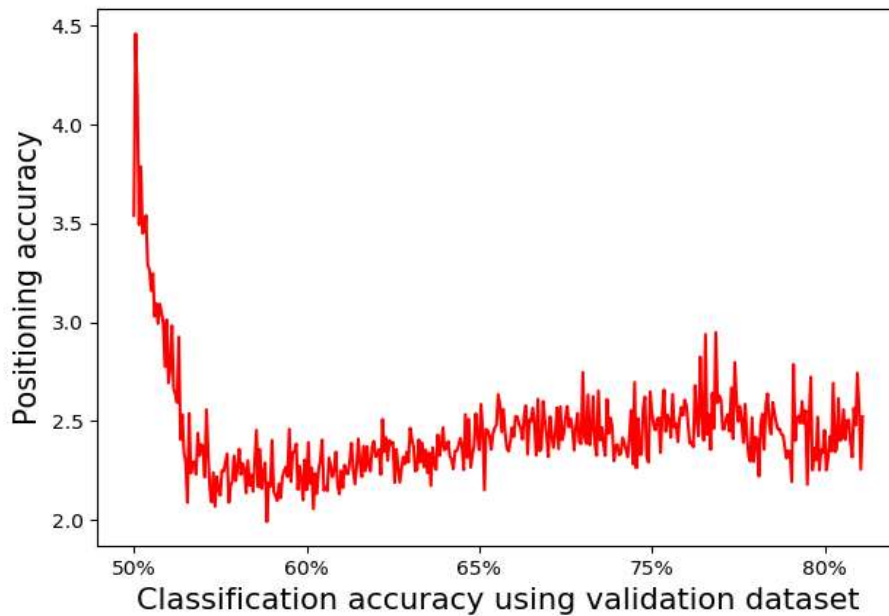
Fig.15. *Improved positioning accuracy by using CNN*

Figure 15 shows the performance of using our CNN model with GA selected APs, it also shows that over the training, the positioning accuracy increased, then decreased, which means that it relates to how well we trained the model. The best positioning accuracy achieved is 1.88m, where the corresponding classification validation accuracy is 57%. N.B. it is not necessarily the higher the classification accuracy, the higher the positioning accuracy. An overtrained model will also cause a decreased positioning accuracy, which means if we use models like CNN, we need to keep the balance between classification accuracy and positioning accuracy. A similar positioning accuracy also means that less hyperparameters tuning is needed and has the ability to compete with the state-of-art neural networks used for IPS. However, it is possible to achieve a higher positioning accuracy by tuning the hyperparameters of neural networks. So, we can depend on the requirements to choose the methods. In our experiment, we used KTCC to do the ADL recognition, because of its simplicity. Both KTCC and CNN-based methods perform 3.5% and 5.9% better than our PCA baseline method, which the accuracy is 2.01m.

As the TPs are collected from several designed paths, we can deploy the EKF model to increase the positioning accuracy with using a uniform motion model. The positioning accuracy increased from 1.94m (KTCC), 1.88m (CNN) to 1.42m, 1.54m, respectively, when we use the EKF. Figure 16 shows the improved positioning accuracy (KTCC).
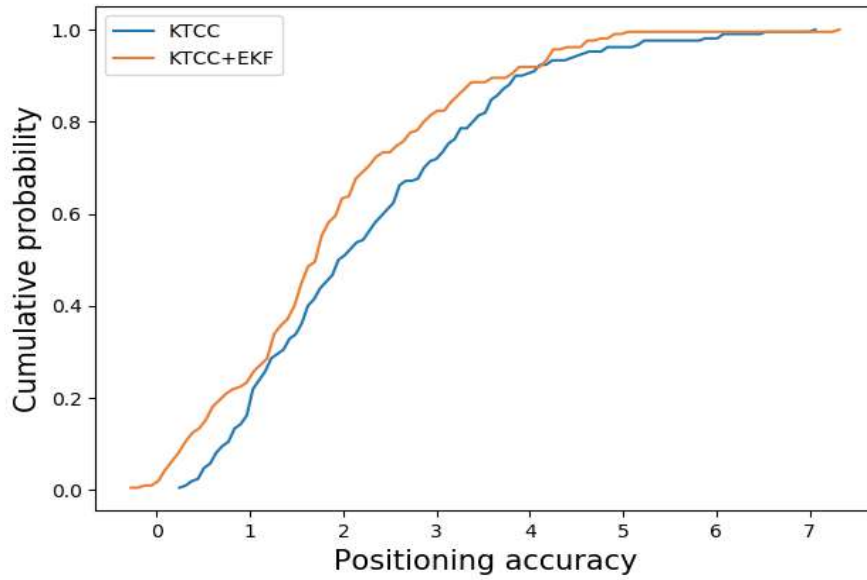
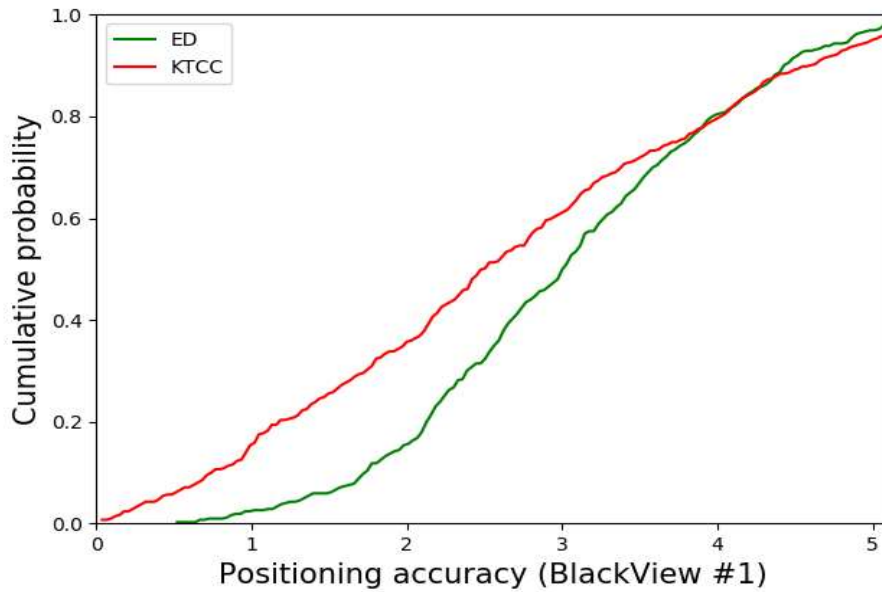**Fig.16.** *Improved accuracy using EKF*

Fig.17. *Accuracy comparison using a different phone*

To test our proposed RSSI ranking location fingerprinting method using KTCC again [12], a repeat experiment was carried out in the Lab testbed. Figure 17 shows that for the using a different phone (BlackView #1), the proposed method (averaged positioning error 2.42 m, 90% in 4.52 m) performs better than using the ED method (3.18 m, 90% in 4.51 m), which also proves that our ranking method is more robust and mitigates the RF receiver heterogeneity issue.

ADL recognition

Figure 18 and Table 2 shows the activity recognition accuracy of using estimated locations, ranking vectors and fusion-based methods. C1 is the set of class 1 activities, C2 is the set of class 2 activities and mix represents C1 plus C2 activities. From the result, we can see the recognition accuracy of C1 is better than that of C2 which indicates that the activities with short paths are harder to be recognised than activities with long paths which also can explain the degradation of recognition accuracy for the mix set compared to that of C1, which means the longer path (but with less common points), the higher accuracy. Moreover, the proposed fusion MD-DTW performs the best recognition accuracy, it also means ranking-based fingerprints can do more than just localisation, it can also be explored to recognise ADLs.
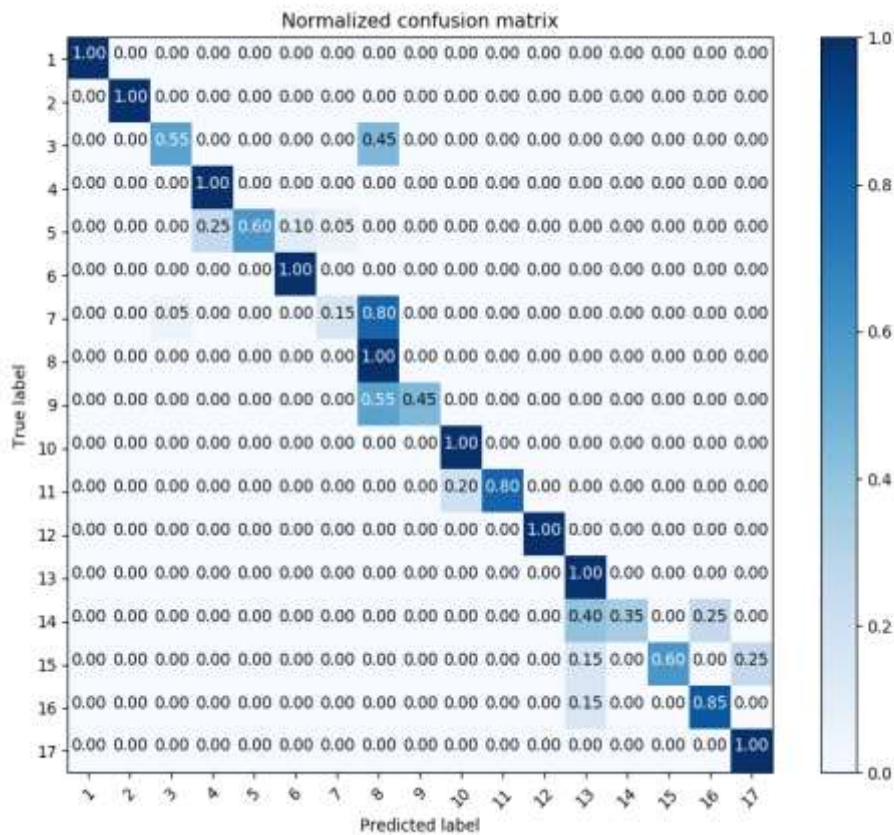

Fig.18. *Confusion matirx of our recognition results*

Table 2. WiFi location-based activity recognition result

Percent	Single Dimension DTW	Fusion MD DTW

	C1	C2	mix	C1	C2	mix
Estimated KTCC	44.6	30.0	40.3	--	--	--
Ranking Vectors	77.9	62.0	73.2	--	--	--
Fused Vectors	--	--	--	79.8	76.0	79.5

In Figure 18, the Confusion matrix of the recognition results give the confusion matrix for using the proposed method to recognise the designed activities. The worst performance (15%) is to recognise the print document activity (80% data are wrongly recognised for activity 8 – the same type activity as 7, but uses a different path). However, it is acceptable as the final output still will be the print document activity. The reason for this is the path 7 and 8 share a lot of common points (step or distance), which makes it harder to discriminate between them. This also happens for path 3 and path 7, path 14 and path 16, and path 15 and path 17. However, it is found that if the paths have more than 4 different steps (2 m), the most recognition accuracies become 100%, which means it has a high probability to discriminate paths with a 2 m (distance) difference.

However, it is not simply related to the distance, a more detailed relation between steps (distance) and the number of common steps and recognition accuracy will be considered in the future.

Conclusions

RSSI measurements for different smartphones and at different times were collected. The heterogeneity of hardware impact decreases the positioning accuracy of conventional localisation methods. We have proposed to use KTCC and CNN with WiFi ranking to address this issue. We also introduced a ranking-based AP selection algorithm using GA, which not only reduces computational cost but also increases the positioning accuracy.

Our location fingerprinting validation results show that our IPS can get an average real-time positioning error of 1.42m, which would not be severely affected by the heterogeneity impact. It improves the positioning accuracy compared to existing state of the art IPS systems (our PCA baseline system). However, we do not know the exactly APs number and their positions, so topology of APs is not taken into account.

This paper also focused on the ADLs recognition using WiFi ranking fingerprint-based location awareness. We designed 9 activities, 17 routes and validated our strategy in an office at the Queen Mary University of London. The best performance of our proposed fusion method is 79.5%, By extending our multi-source information fusion model, other signals like magnetic field [20] can also be fused to provide more eigenvalues for path matching to lift the recognition accuracy.

There are Four directions for our future work. The first is that Recurrent Neural Network (RNN) can be investigated to derive user's locations and to recognise their activities, as it is a time-series data, which RNN has a good capability to deal with [21]. Second, we can fuse more information, e.g., GPS measurements, accelerometer measurements, to expand

the data dimension, then to increase recognition performance when the size of the AP set is not large. Third, we aim to improve on ADL recognition [22] [23] using stay points (where a moving object stays for a period of time) and to combine other sensors sch as inertial sensors to improve the separation of different spatial overlapping activities, which is one of the weak points in this study. Forth, we need to improve user's trust in the system by safeguarding the privacy of sensitive users' location tracking information [24], [25].

Acknowledgements

References

[1] N.E Klepeis, W.C, Nelson, W.R Ott, et al., "The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants," J. of Exposure Analysis and Environmental Epidemiology volume 11, pages 231–252 (2001) T. Platts-Mills, "How environment affects patients with allergic disease: indoor allergens and asthma," Annals of allergy, vol. 72, pp. 381-384, 1994.

[2] M. Worboy,. "Modeling indoor space," Proc. 3rd ACM SIGSPATIAL Int. Workshop on Indoor Spatial Awareness, pp. 1-6, 2011.

[3] Z. Liang, I. Barakos and S. Poslad, "Indoor Location and Orientation Determination for Wireless Personal Area Networks." Lecture Notes in Computer Science, vol. 5801, pp 91-105, 2009.

[4] W. Leister, M. Hamdi, H. Abie, et al., "An Evaluation Scenario for Adaptive Security in eHealth," 4th Int. Conf. on Performance, Safety and Robustness in Complex Systems and Applications, PESARO, Nice, France., pp. 6 -11, 2014.

[5] W. Tsui, Y. H. Chuang, and H. H. Chu, "Unsupervised learning for solving RSS hardware variance problem in WiFi localization," Mobile Networks and Applications, vol. 14, pp. 677-691, 2009.

[6] Wu, Z. Ma, S. Poslad, and W. Zhang [is it correct?], "An efficient wireless access point selection algorithm for location determination based on RSSI interval overlap degree determination," in 2018 Wireless Telecommunications Symposium (WTS), Phoenix, USA, Apr. 2018.

[7] G.A. ten Holt, M.J.T Reinders and E.A. Hendriks, "Multi-dimensional dynamic time warping for gesture recognition." 13th annual conference of the Advanced School for Computing and Imaging. Vol. 300. 2007.

[8] F. Adib and D. Katabi, "See through walls with wifi!," In ACM SIGCOMM, 2013.

[9] Q. Pu and et al. "Whole-home gesture recognition using wireless signals," ACM MobCom, 2013.

[10]   F. Adib, Z. Kabelac, D. Katabi, and R. C. Miller, "3D tracking via body radio reflections," Usenix NSDI, 2014.

[11]   Y. Xu, M. Zhou, and L. Ma, "WiFi indoor location determination via ANFIS with PCA methods," IEEE Int. Conf. on Network Infrastructure and Digital Content, 2009. IC-NIDC, 2009, pp. 647-651

[12]   Z. Ma, S. Poslad, J. Bigham, X. Zhang, and L. Men, "A ble rssi ranking based indoor positioning system for generic smartphones," Wireless Telecommunications Symposium (WTS), 2017, pp. 1–8.

[13]   H. Salamah, M. Tamazin, M. A. Sharkas, and M. Khedr, "An enhanced wifi indoor localization system based on machine learning," Int. Conf. on. IEEE Indoor Positioning and Indoor Navigation (IPIN), 2016, pp. 1–8.

[14]   R. Elbakly and M. Youssef, "A robust zero-calibration rf-based localization system for realistic environments," 13th Annual IEEE Int. Conf. on. IEEE Sensing, Communication, and Networking (SECON), 2016, pp. 1–9.

[15]   Z. Chen, H. Zou, H. Jiang, Q. Zhu, Y. C. Soh, and L. Xie, "Fusion of WiFi, smartphone sensors and landmarks using the Kalman filter for indoor localization," Sensors, vol. 15, no. 1, pp. 715–732, 2015.

[16]   K. A. Nguyen and Z. Luo, "On assessing the positioning accuracy of Google Tango in challenging indoor environments," Int. Conf. on Indoor Positioning and Indoor Navigation (IPIN), 2017 pp. 1-8

[17]   D. Berndt, J. Clifford. "Using dynamic time warping to find patterns," time series," AAAI-94 workshop on knowledge discovery in databases, pp 229–248.

[18]   K. A. Nguyen and Z. Luo, "Dynamic route prediction with the magnetic field strength for indoor positioning," International Journal of Wireless and Mobile Computing, vol. 12, pp. 16-35, 2017.

[19]   Z. Shi, Z. Zhang, Y. Shu, P. Cheng, and J. Chen, "Indoor Navigation Leveraging Gradient WiFi Signals," in Proc. of the 15th ACM Conference on Embedded Network Sensor Systems, 2017, p. 48.

[20]   Z. Ma, S. Poslad, S. Hu, & X. Zhang, "A fast path matching algorithm for indoor positioning systems using magnetic field measurements," IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), 2017 (pp. 1-5). IEEE.

[21]   Ma, Z., J. Bigham, S. Poslad, B. Wu, X. Zhang, E. Bodanese, "Device-Free Daily Life (ADL) Recognition for Smart Home Healthcare using a low-cost (2D) Lidar," Proc. IEEE Globecom, 2018.

[22]   J. Yang, H. Zou, H. Jiang, and L. Xie, "CareFi: Sedentary Behavior Monitoring System via Commodity WiFi Infrastructures," IEEE Trans. Veh. Technol., 2018.

[23]   H. Zou, Y. Zhou, J. Yang, W. Gu, L. Xie, and C. Spanos, "WiFi-based Device-Free Human Activity Recognition via Automatic Representation Learning," in Proc. of the 23rd Annual International Conference on Mobile Computing and Networking, 2017, pp. 606–608.

[24]   C. Jensen, S. Poslad, T. Dimitrakos. (Eds), "Trust Management," Proc. 2nd Int. Conf. on Trust Management, iTrust, Lecture Notes in Computer Science (LNCS), Springer-Verlag, vol. 2995, 2004.

[25]   S. Poslad, M. Hamdi, H. Abie, "Adaptive security and privacy management for the internet of things". Proc. ACM conference on Pervasive and ubiquitous computing adjunct publication, 2013, pp. 373-378.