
Dynamic route prediction with the magnetic field strength for indoor positioning

Khuong An Nguyen* and
Zhiyuan Luo

Department of Computer Science,
Royal Holloway, University of London,
Egham, Surrey, TW20 0EX, UK
Email: khuong.nguyen.2007@live.rhul.ac.uk
Email: zhiyuan.luo@rhul.ac.uk

*Corresponding author

Abstract: WiFi fingerprinting has been a popular approach for indoor positioning in the past decade. However, most existing fingerprint-based systems were designed as an on-demand service to guide the user to his wanted destination. This paper introduces a novel feature that allows the positioning system to predict in advance which walking route the user may use, and the potential destination. To achieve this goal, a new so-called routine database will be used to maintain the magnetic field strength in the form of the training sequences to represent the walking trajectories. The benefit of the system is that it does not adhere to a certain predicted trajectory. Instead, the system dynamically adjusts the prediction as more data are exposed throughout the user's journey. The proposed system was tested in a real indoor environment to demonstrate that the system not only successfully estimated the route and the destination, but also improved the single positioning prediction.

Keywords: fingerprinting; magnetic field strength; indoor positioning.

Reference to this paper should be made as follows: Nguyen, K.A. and Luo, Z. (2017) 'Dynamic route prediction with the magnetic field strength for indoor positioning', *Int. J. Wireless and Mobile Computing*, Vol. 12, No. 1, pp.16–35.

Biographical notes: Khuong An Nguyen holds a PhD in Computer Science at Royal Holloway, University of London, UK. Before that, he completed an MPhil in Advanced Computer Science at the University of Cambridge, where he implemented a Bluetooth tracking system in the Computer Lab. His research interests include indoor positioning and machine learning.

Zhiyuan Luo holds a PhD in Computer Science from Heriot-Watt University, Edinburgh, UK. Prior to joining Royal Holloway, University of London, UK, he worked for Cable and Wireless UK as a technical consultant, focussing on network modelling and performance evaluation. His expertise is on the development and application of high-dimensional data analysis and machine learning algorithms. He has been involved in multiple UK EPSRC and EU research grants.

1 Introduction

Indoor navigation is an important application to provide the positioning service to the indoor users, where the Global Navigation Satellite Systems such as GPS struggle to work reliably. Current indoor positioning systems either provide their own infrastructure, or rely on existing ones indoors. The former may offer highly accurate positions using expensive customised hardware (Want et al., 1992; Priyantha, 2005; Holm, 2009). The latter are affordable, yet their accuracies are not too great. Amongst the infrastructure-free category is location fingerprinting, which has been widely considered as one of the most efficient indoor tracking methods with good positioning accuracy to date (Bahl and Padmanabhan, 2000; Youssef and Agrawala, 2005;

Weber et al., 2010; Wang et al., 2015). Fingerprinting-based approaches make use of the ubiquitous indoor WiFi network to deliver the tracking service.

However, most previous fingerprint-based systems operated in an active tracking context, where the user submits a new WiFi signal sample for the system to discover his current position. This paper considers fingerprinting in a passive tracking scenario, where the system has the permission to monitor the user positions continuously to react in a timely manner when his position changes (i.e. to open the door, to switch off the light [Dey et al., 2000; Kim, 2011; Subrt and Pechac, 2012]). Crucially, with passive monitoring, the system is guaranteed to have access to a sequence of the signal data, which provides useful insights into the walking trajectory. In addition, these users normally have established

presence in the building, hence their personal routines should have been well-observed to help the system predict their intended travelling path in advance.

To achieve this goal, the paper will present a novel data set called the ‘routine database’ to reflect the user’s walking routine, based on the magnetic field data. The benefit of this approach and the details of the implementation will be discussed in the remaining sections of this paper. Overall, the proposed approach offers the following benefits.

- i Magnetic field strength is used instead of the WiFi readings for fingerprinting.
- ii The user’s personal travelling history is taken into account to predict his walking route.
- iii The system uses a sequence of real-time signal data to make the positioning prediction.

2 Extra information from the magnetic field

This section explores the use of the indoor magnetic field for fingerprinting. It only examines the features that are most relevant to fingerprinting, including the sensitivity and the uniqueness. More detailed experiments of the indoor magnetic field data can be found in Li et al. (2012), Kim et al. (2012) and Xie et al. (2014).

2.1 How to measure the magnetic field?

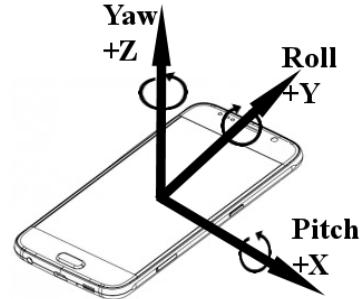
An Android phone, one of the most popular devices to perform fingerprinting, was used to collect the magnetic field for all experiments in this paper. The geomagnetic field sensor on the Android phone measures the strength and the direction of the magnetic field at a point in space, where the phone is held. This measure is presented as a three-dimensional vector $\vec{m} = (x, y, z)$, where x, y, z (measured in microTesla or μT) are the ambient strength of the magnetic field in the corresponding axis (see Figure 1). The fastest sampling frequency on the experimented Nexus 5 phone was about three samples per second.

It is important to note that these axes are relative to the position of the mobile phone in space. Therefore, the magnitude of each axis may differ as the device’s orientation changes, even when it stays in the same spot. The accelerometer included on the phone may be used to detect the direction of the gravity, which helps to deduct the 3D position of the phone. However, since the focus of this paper is mainly about the magnetic field, this idea is left out for future work. A simple solution to this problem is to combine the Magnetic Field Vector (MFV) into one scalar magnitude as follows (see equation (1)).

$$\|\vec{m}\| = \sqrt{x^2 + y^2 + z^2} \text{ } (\mu\text{T}) \quad (1)$$

However, with this approach, there is only measure left for each position. For the rest of this paper, this single scalar magnitude will be referred to as the magnetic field strength (MFS), and the whole magnetic field vector with all three magnitudes of the three axes will be referred to as the MFV.

Figure 1 The three axes of the geomagnetic field sensor



2.2 The sensitivity of the magnetic field strength

The reliability of fingerprinting depends on the stability of the fingerprints. This is the ability to repeat the same measure at the same position at any time in the near future. This requirement is challenging, because the structure of the building may have changed, or because of other external factors that happened at the time of data measuring (e.g. moving people or furniture). The former may happen gradually which allows the system to cope with, although it may require the whole training data to be re-surveyed. The latter is normally caused by the indoor users’ movement, which may be avoided if the MFS has low sensitivity. Two experiments were performed to assess the sensitivity of the MFS in real time and over long periods of time.

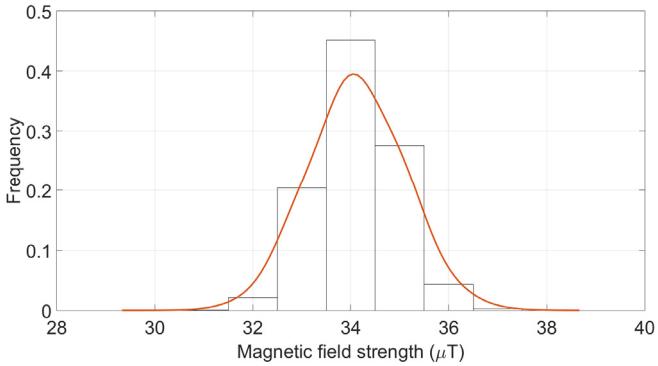
The first experiment measures the variation of the MFS at a fixed position. It is preferred that the variation is low to avoid the second challenge posed above. The experiment was performed over 6 hours during the working hours to reflect a busy environment, with many people walking in the building. Compared to the WiFi RSS, the MFS varied much less rapidly in its measurement unit in fixed positions. This experiment also suggested that the stronger the field strength was (the closer it is to the magnetic source), the more stable the MFS was, which was exactly the opposite of the WiFi RSS, where strong RSS was more sensitive than weak one (see Figure 2).

The second experiment assesses the repeatability of the MFS at the same positions. For this experiment, 15 positions in the building were marked with duck tape. They were revisited weekly, over one month to measure the MFS. Only five measures were taken at each observation point, and the average was used for this experiment. It was suggested that the variation was minimal, compared to the WiFi RSS observed at the same locations, with up to only 3 μT for the positions with strong MFS, and up to 4 μT for other positions with weak MFS (see Figure 3). Such difference was well-within the short-term variation range of the magnetic field reported in the last experiment. Also, there might be a slight displacement of the phone’s position when the experiment took place, which explains this small change.

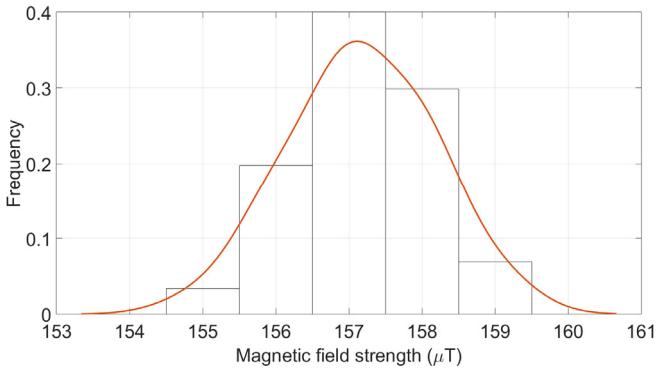
In summary, the magnetic field has low sensitivity and is highly repeatable with time, compared to the highly volatile WiFi RSS. The major difference of the WiFi RSS and MFS is that strong MFS was observed to be more robust than weak one, whereas strong WiFi RSS was more

sensitive than weak RSS. This is an important suggestion to apply the magnetic field to fingerprinting, since most positioning systems would favour the strong signals.

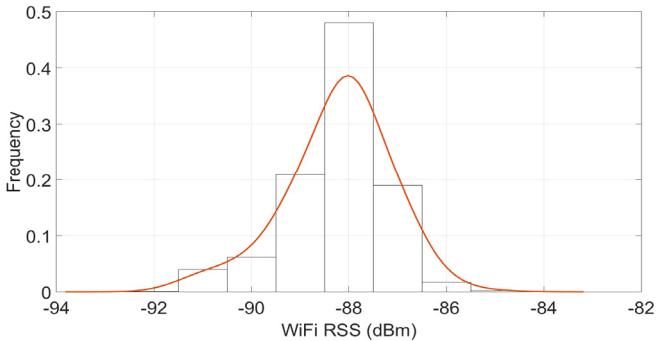
Figure 2 Histograms of the WiFi RSS and MFS in fixed positions from the same magnetic source and AP, over 6 hours during working hours. Strong MFS was more robust than weak MFS. Strong WiFi RSS was less robust than weak WiFi RSS (see online version for colours)



(a) Weak magnetic field strength from a central heater measured at a fixed position. It varied from 31 to 37 μT (7 μT variation). The normalised range is [0.84, 1]

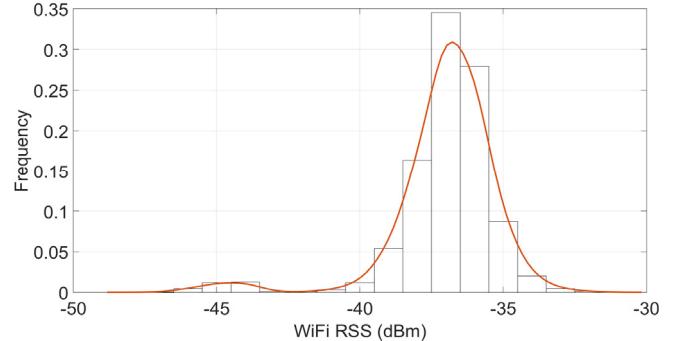


(b) Strong magnetic field strength measured from the same central heater at a different fixed spot. It varied less between 155 and 159 μT (5 μT variation). The normalised range is [0.97, 1]



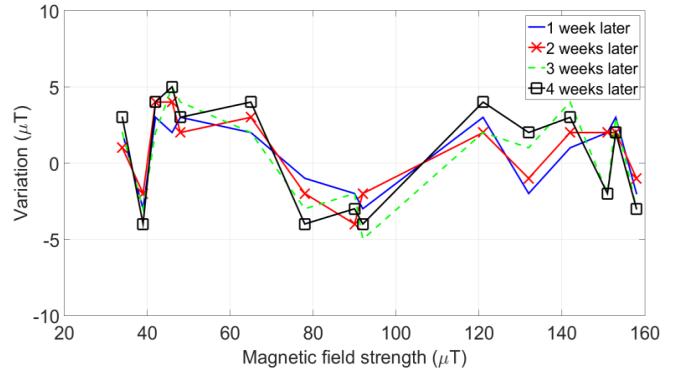
(c) Weak WiFi RSS from a single AP varied from -85 to -93 dBm (9 dBm variation). The normalised range is [0.91, 1]

Figure 2 Histograms of the WiFi RSS and MFS in fixed positions from the same magnetic source and AP, over 6 hours during working hours. Strong MFS was more robust than weak MFS. Strong WiFi RSS was less robust than weak WiFi RSS (see online version for colours) (continued)

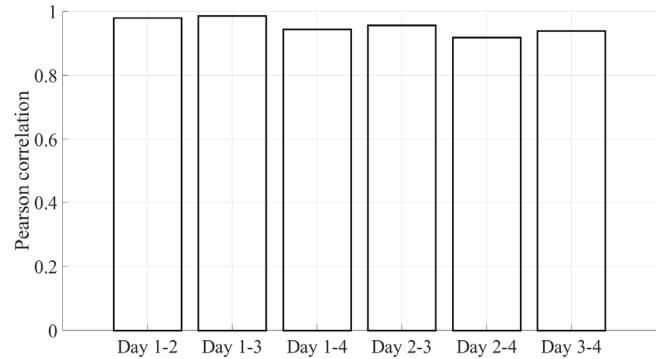


(d) Strong WiFi RSS from the same AP varied much more from -30 to -48 dBm (19 dBm variation). The normalised range is [0.62, 1]

Figure 3 The change of the MFS and WiFi RSS at 15 observation points over one month. Strong MFS was slightly less sensitive, while strong WiFi RSS was more sensitive. MFS sequences were highly correlated, while WiFi RSS ones were not

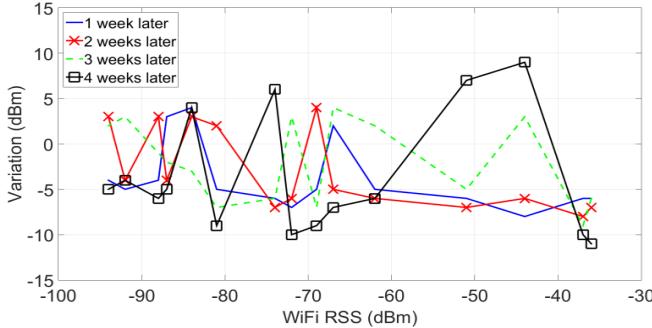


(a) MFS change over time

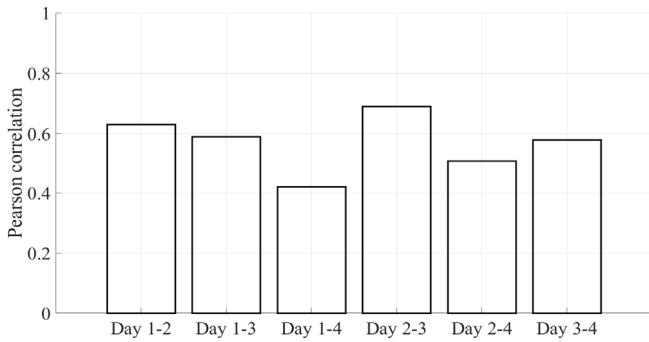


(b) Pairwise comparison of MFS sequences with Pearson correlation. The closer it is to 1, the more similar the two sequences are

Figure 3 The change of the MFS and WiFi RSS at 15 observation points over one month. Strong MFS was slightly less sensitive, while strong WiFi RSS was more sensitive. MFS sequences were highly correlated, while WiFi RSS ones were not (continued)



(c) WiFi RSS change over time



(d) Pairwise comparison of WiFi RSS sequences with Pearson correlation. The closer it is to 1, the more similar the two sequences are

2.3 The uniqueness of the magnetic fingerprints

The positioning accuracy of fingerprinting relies on how unique the surveyed fingerprints are. If all the training examples are similar, it will be challenging for any algorithm to use them to estimate the user position. With the magnetic field, the building is made of large ferrous metal structures (i.e. pipes, steel shells) which greatly distort the magnetic readings at many indoor positions (see Figure 5). However, their influence is mostly local, and the greater the anomalies, the more unique the magnetic fingerprint is.

The major challenge when using the magnetic data for fingerprinting is that it only contributes a maximum of three measurements corresponding to the magnitude of the three axes for each fingerprint, whereas the WiFi signal provides a big vector of many RSS readings from several nearby WiFi APs. A single walk through the ground floor of the test bed exposed several different positions with a similar MFV, in which the difference in the magnitude of each axis was just $2 \mu\text{T}$ (see Figure 4).

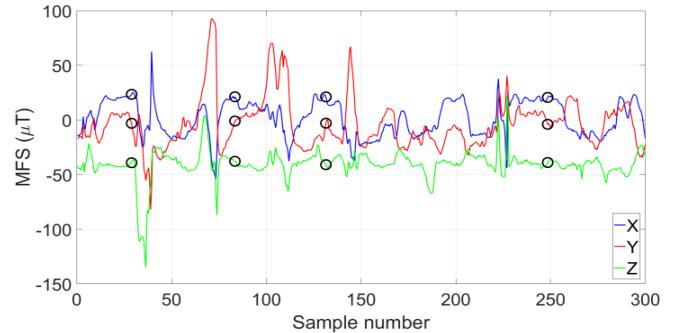
For this reason, the magnetic field alone should not be employed to represent a single position, as in the case with the WiFi RSS. To take advantage of the robust magnetic measure, however, this paper proposes to merge them into a long sequence to represent a walking trajectory, which is much more distinctive. Intuitively, it is less likely to have two different trajectories with the same magnetic field representation. These

trajectories form the user's routine database to be explored in the next section.

Figure 4 The magnitude of each of the three axes measured from a single walk through non-repeated positions on the ground floor of the test bed. The black circle pinpoints the positions with a similar reading

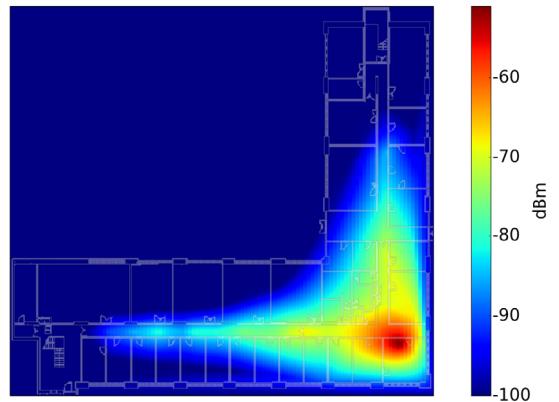


(a) The walking trajectory



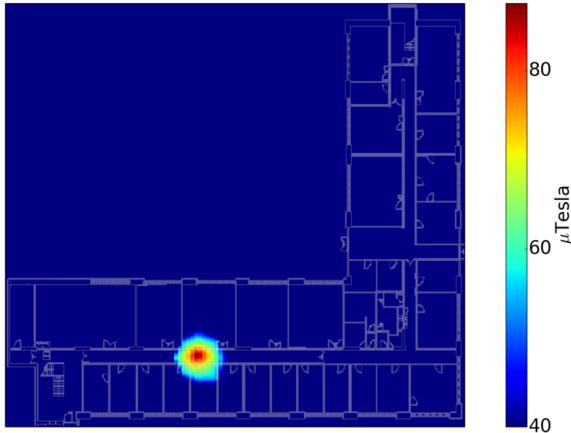
(b) The magnetic readings

Figure 5 Magnetic field strength is more resilient than WiFi measurement from its transmitting shape. Distance-wise, WiFi coverage is far greater than magnetic field, from a single power source



(a) The propagation of WiFi RSS from a single AP in an office. The non-uniform signal coverage was caused by signal passing through doors and walls. Positions further down the corridor may still see the AP, although the signal was relatively weak

Figure 5 Magnetic field strength is more resilient than WiFi measurement from its transmitting shape. Distance-wise, WiFi coverage is far greater than magnetic field, from a single power source (continued)



(b) Magnetic field strength from a central heater attached on the wall in the same indoor corridor. It provides very fine scale reading over short distance

3 The personal routine database

In addition to the fingerprinting database, a routine database is introduced for each user. This new database describes the user's travelling history in the form of the frequently visited places, and the routes amongst them. Since each user normally has different preferred trajectories, he should have his own routine database. Each training example of the routine database represents a walking trajectory, whereas each training example in the fingerprinting database represents a single position. The training trajectory is a sequence of continuous MFV readings, mapped to their Cartesian coordinates. It is essential to clarify that the main purpose of this routine database is not to substitute the fingerprinting counterpart, but to provide prior knowledge to predict where the user may go next. The fingerprinting database will still be needed when the user takes a completely new route, which has not been recorded in the routine database. This routine database was introduced based on three observations.

- i The indoor user often takes the same route to travel between familiar places.
- ii The user often walks in a straight path.
- iii For a moving user, there is often not enough time for the tracking system to collect multiples readings at every single position in real time.

This section outlines the process of generating the routine database, and the formal model of such database.

3.1 The offline phase to generate the routine database

Initially, the routine database is generated manually in the same manner as the fingerprinting database. The two differences are

that the magnetic field is collected instead of the WiFi RSS, and each training example in the routine database is a sequence of MFV. The expert first identifies the route he wants to cover. He then uses an Android phone (with the app described in Section 4.1) to record the magnetic data at different positions along the route, and provides them with the corresponding Cartesian labels. At each position, the magnetic data should also be measured several times and an averaged measure is computed, although the number of repeated measures need not be as high as in the case with the WiFi RSS, because of the low variation of the magnetic field as previously discussed in Section 2.2. The starting and ending positions of the trajectory are also labelled by the room or office number, so that the system can provide a human-readable response when needed. The only criterion the expert should consider at this phase is how long the gap between the two consecutive positions in the trajectory is. Ideally, it is preferred that this gap is small. The granularity of the fingerprinting database can be used as a reference to decide this gap (e.g. 1 metre between consecutive positions). A method to compare two trajectories with different gaps will be discussed in the next section. The expert will attempt to cover as many trajectories as possible, which is similar in the sense that he also tries to cover as many individual training positions for the fingerprinting database as possible.

It may also be useful to incorporate the user's personal timetable into his routine database, by looking for the starting and ending positions of the events on the calendar. However, such approach is beyond the scope of this paper. For the experiments described here, the routine database is manually generated by the user to reflect his personal routine.

3.2 Modelling the routine database

Without any loss of generality, the routine database RD is formally modelled as a set of M examples $RD = \{T_1, \dots, T_M\}$. Each training example $T_i = (R_1, \dots, R_K)$ ($1 \leq i \leq M$) represents a walking route, where R_j ($1 \leq j \leq K$) is the data of the position j th along the route, with K is the total number positions on the route. Each position $R_j = (\overrightarrow{MFV}_j, \overrightarrow{L}_j)$ contains the magnetic field vector $\overrightarrow{MFV}_j = (x, y, z)$ recorded at that position, along with its Cartesian label $\overrightarrow{L}_j = (L_x^j, L_y^j)$, with x, y, z is the magnitude (μ T) of the three axes.

The task is, given a magnetic sequence without the Cartesian label $(\overrightarrow{MFV}_1, \dots, \overrightarrow{MFV}_L)$, the system finds the best training trajectory that matches this sequence.

3.3 Estimating the user trajectory with the routine database

The user will have an app running in the background to collect the magnetic data automatically. This is also the common assumption for most passive monitoring systems. The recording process is triggered when the accelerometer

readings are above a certain threshold for a window period of a few seconds which indicates that the user is moving (see Figure 6). This same process will stop when the user no longer moves, for which the accelerometer readings are within the threshold for a window period of time. It is worth noting that the use of accelerometer in this paper is basic, although these accelerometer measures can tell more information about the user movements as in WiFi Simultaneous Localisation and Mapping (SLAM) research (Martin et al., 2010; Kim et al., 2012; Wang et al., 2012; Faragher and Harle, 2013).

When triggered, the magnetic data are recorded continuously to generate a real-time trajectory. It may be possible that some false recordings will get through when the phone is momentarily picked up (e.g. to make a call). This issue may be avoided by setting an initial threshold for the trajectory's length, so that the prediction module will only engage when the trajectory has accumulated at least a certain amount of data. This approach is left for future research.

As the user travels, the system assumes that the training trajectory with the closest match for the current real-time sequence is the correct route. This assumption, in turn, tells the system what the potential destination that the user may arrive at. However, the system does not stick to a single predicted trajectory. Throughout the journey, the user's magnetic sequence gets larger, and also becomes more unique. Every time a new reading arrives, the system recalculates the measure to the training examples to find a better trajectory, if any. Therefore, the potential destination is dynamically adjusted as the user navigates the building.

To compare the difference of two magnetic sequences, this paper employs Dynamic Time Warping (DTW). The benefits of DTW are to mitigate the difference in the measuring time of the training and real-time sequences, the ability to compare signal sequences of different lengths, and most crucially is to eliminate the temporal variance in walking speed (Müller, 2007; Subbu et al., 2011). The smaller the DTW measure is, the closer the two trajectories are. The variant of DTW employed in this paper is called open-end DTW (OE-DTW) (Giorgino, 2009). Normal DTW will stretch the shorter sequence to match the longer one up until the end, for which this OE-DTW version relaxes the endpoint constraint. It serves the purpose of this paper better, because in the early stage of the journey, the user's MFV sequence is short, and may introduce certain bias when compared to the full training trajectories, even though they start from the same position. In principle, OE-DTW achieves its target by constructing different incomplete versions of the longer sequence, and picks the best match. In order to compare the DTW measure at different points of the journey, these measures will be normalised by dividing with the length of DTW aligned sequence. For the rest of the paper, the term OE-DTW will be simply referred to as DTW, and DTW measure means normalised OE-DTW measure.

Figure 7 demonstrates an example of this estimation process in real time. A video demonstrating this process can

be viewed at <http://www.cs.rhul.ac.uk/~wruf265/dtw.mp4>. The route prediction algorithm with the routine database is summarised below (see Algorithm 1).

If the DTW measures between the real-time sequence and all training trajectories are low, in other words, no training trajectory matches the current sequence, the system will simply make a single position-by-position prediction with the fingerprinting database. This situation may happen because the user is relatively new to the building, or because his walking routines have not been well-recorded.

Algorithm 1: Estimating the trajectory with the routine database

Data: Routine database with M training trajectories $\{T_1, \dots, T_M\}$.

Result: Route prediction.

while user is still moving **do**

- read Magnetic field vector \overline{MFV}_k at current location k ;
- /* append it into the magnetic sequence */
- $(mag_{cur} = mag_{cur} + \overline{MFV}_k);$
- /* Find the closest training trajectory */
- $*/$
- $min = \infty;$
- $T_{ref} = T_1;$
- for** $i = 2 \rightarrow M$ **do**
- if** $DTW(T_i, mag_{cur}) < min$ **then**
- $min = DTW(T_i, mag_{cur});$
- $T_{ref} = T_i;$
- end**
- end**
- return** T_{ref}

Figure 6 Acceleration reading with the phone in the trouser pocket. When the device is static, the reading is around 9.8 m/s^2 , which is equivalent to the earth's gravitational force (Heiskanen and Meinesz, 1958) (see online version for colours)

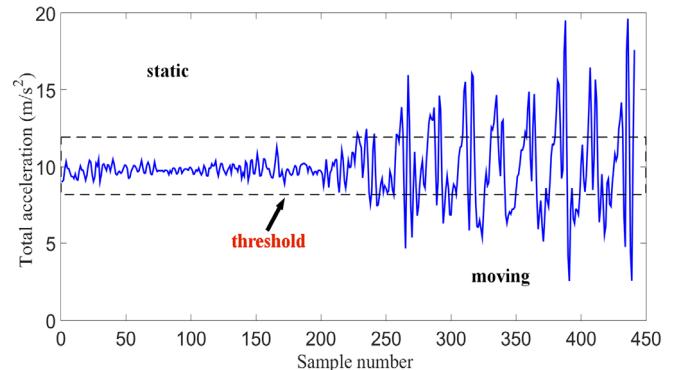
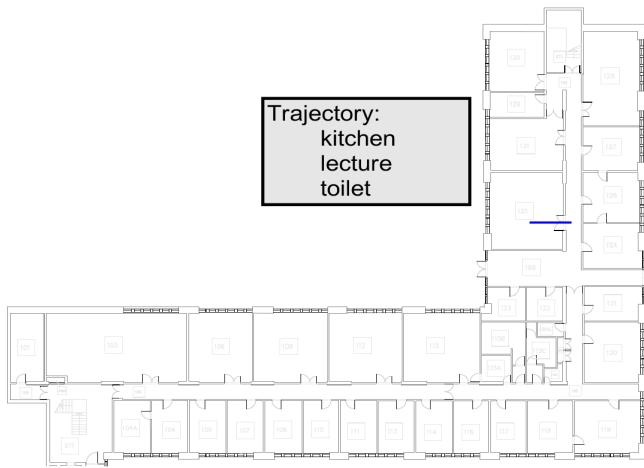
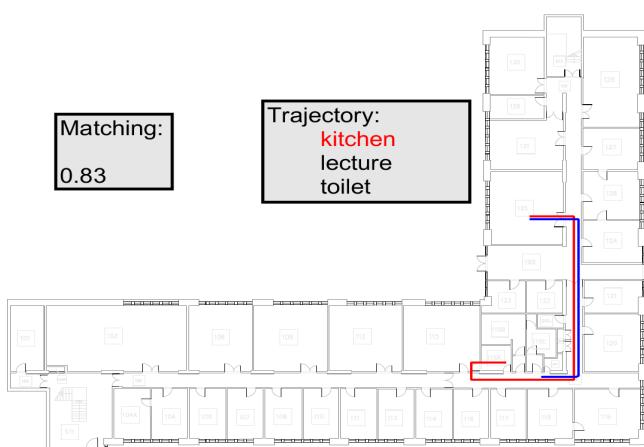


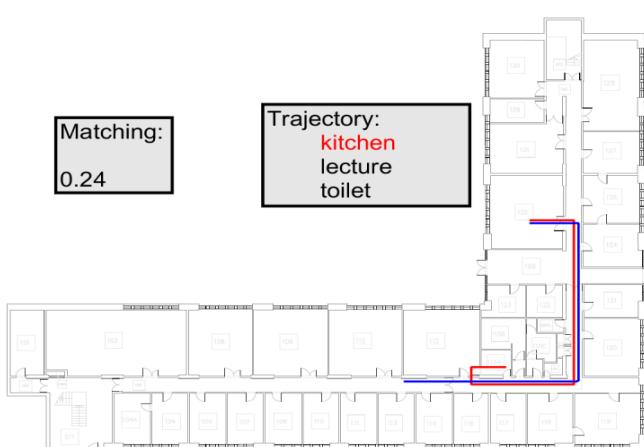
Figure 7 An example of the dynamic route prediction with the routine database. The three training trajectories were Kitchen-bound, Lecture room-bound, and Toilet-bound. The matching number on the left is calculated by DTW and is scaled to [0, 1] for demonstration purpose



(a) The user starts moving from his office. The blue line demonstrates the user path

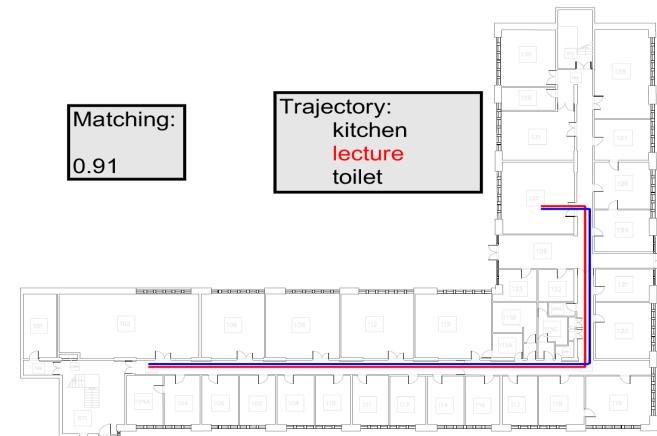


(b) The red line shows the predicted route



(c) However, the user turns to a different route in midway. The matching measure dropped as he navigates away from the predicted route (kitchen)

Figure 7 An example of the dynamic route prediction with the routine database. The three training trajectories were Kitchen-bound, Lecture room-bound, and Toilet-bound. The matching number on the left is calculated by DTW and is scaled to [0, 1] for demonstration purpose (continued)



(d) The system switches to the next matching route (lecture room)

4 Evaluation of performance

This section describes the real-world experiments of the proposed technique. The *R* implementation of OE-DTW was used to conduct the experiments (<http://cran.r-project.org/web/packages/dtw/> – last accessed in September 2016).

4.1 Test bed

Our office building has a 48.1 metres \times 45.7 metres floor plan, where the magnetometer data were collected (see Figure 8). There are nine WiFi APs inside the building, and more than 30 APs in nearby buildings. The smartphone used to collect data was the LG Nexus 5 (Quad-core 2.3 GHz). The PC server has an Octa-core Intel i7 3.9 GHz CPU. The Android app we designed can run silently in the phone's background to collect WiFi and magnetometer data (see Figure 9). It uses the accelerometer to determine if the user is static, and automatically slows down the scanning speed to prolong battery power (see Table 1).

Table 1 Comparison of the battery and memory consumption of our app and other popular apps. For one charging cycle, our app ranks fourth on the least battery consumption list, and second on the least memory consumption list

Application	Power use (%)	Mem. use (MB)
Skype	2	94
Hangouts (SMS)	4	65
Chrome	7	83
Our app	8	29
Gmail	14	94
WhatsApp	19	21
Google Maps	24	98

Figure 8 The interiors of the test bed. All measurements were carried out during daytime



(a) The corridor in the South wing



(b) The pathway in the East wing

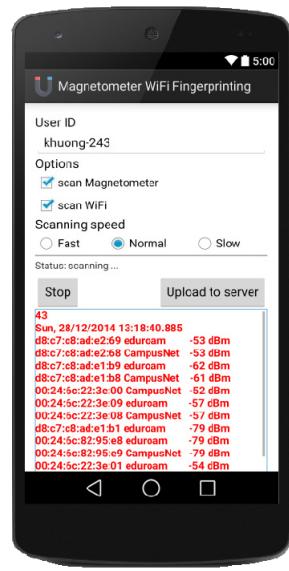


(c) The central heater which emits strong magnetic reading up to $80 \mu\text{T}$

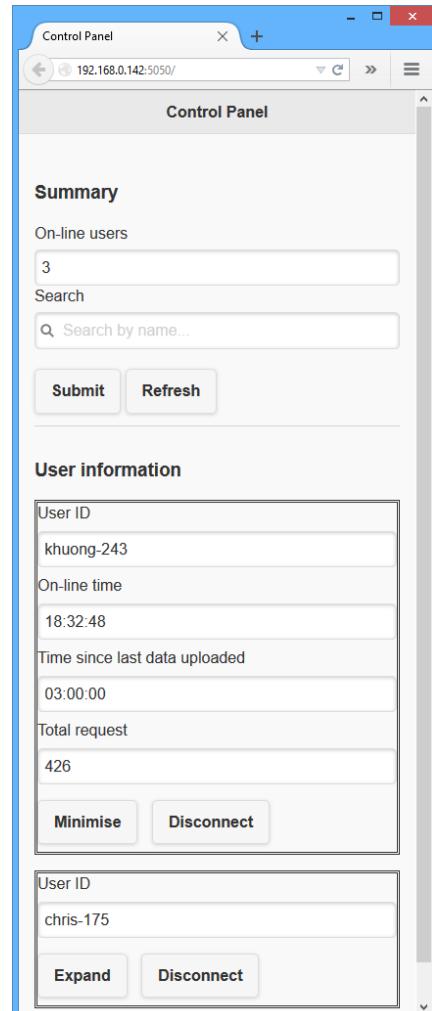


(d) The pillar which emits strong magnetic reading up to $160 \mu\text{T}$

Figure 9 The client and server's interface



(a) The Android app used to collect the magnetic field strength. By default, it uploads the data onto the server every 24 hours. The normal scanning speed is 1 second, while the fastest option aims to collect data continuously as soon as the previous scan finishes



(b) The main control server can see who is connected, and their information. It stores the routine database. Each user has a unique ID

4.2 Criteria of evaluation for the routine database

Owing to the difficulties in recruiting more experts to generate the routine database for many users, this paper only has one routine database with 50 training trajectories, covering all three corridors on the ground floor of the test bed, with the starting and ending points in different offices (see Figure 10). In particular, the longest trajectory has 62 MFV measures, and the average length is 34. The magnetic field data were collected with the phone held in the user's hand at chest level, and the impact of the device's orientation is assumed to be negligible.

The test set contains 23 test trajectories measuring separately at different times from the training set, for which ten of them are fully covered in the training set (see Figure 11). For every fully observed test trajectory, there are at least three training trajectories that start from the same position.

The other ten test trajectories are partly covered so that only the beginning portion of the route was recorded in the training set. These routes either finished in unknown positions or the remaining half of their trajectories was not observed (see Figure 12).

The remaining three test routes are completely new trajectories, which were recorded two floors above the training space in the same building (see Figure 13). There is no information of these routes in the training database.

Using this routine database, the paper aims to address the following questions.

- 1 *How well does the routine database predict the route and the potential destination?* The main purpose of the routine database is to estimate the user's trajectory and destination. Therefore, it is natural to assess how well it performs, in terms of how quickly the system picks the correct trajectory and how often the system switches trajectory during the journey.
- 2 *Can the routine database improve the accuracy of the single positioning prediction?* If the routine database identifies the correct route the user is taking, can it also improve the positioning prediction, compared to the result produced by the fingerprinting database?

For the experiments in this paper, the test trajectory is broken down into ordered individual MFV which represents a single position on the trajectory. Each MFV will be accumulated one-by-one into a sequence to simulate a walking user in real time. At each step, the system computes the DTW measure of the current sequence to every training trajectory. The training trajectory with the smallest measure will be chosen as the predicted route.

Figure 10 Coverage of all training routes (see online version for colours)

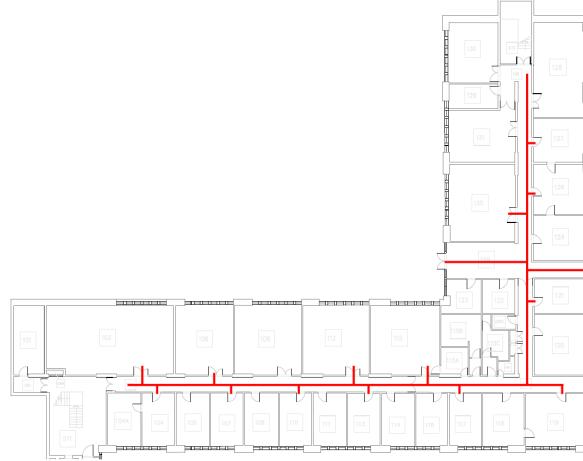


Figure 11 The ten fully observed test trajectories. The blue circle denotes the starting point (see online version for colours)



Figure 11 The ten fully observed test trajectories. The blue circle denotes the starting point (see online version for colours) (continued)

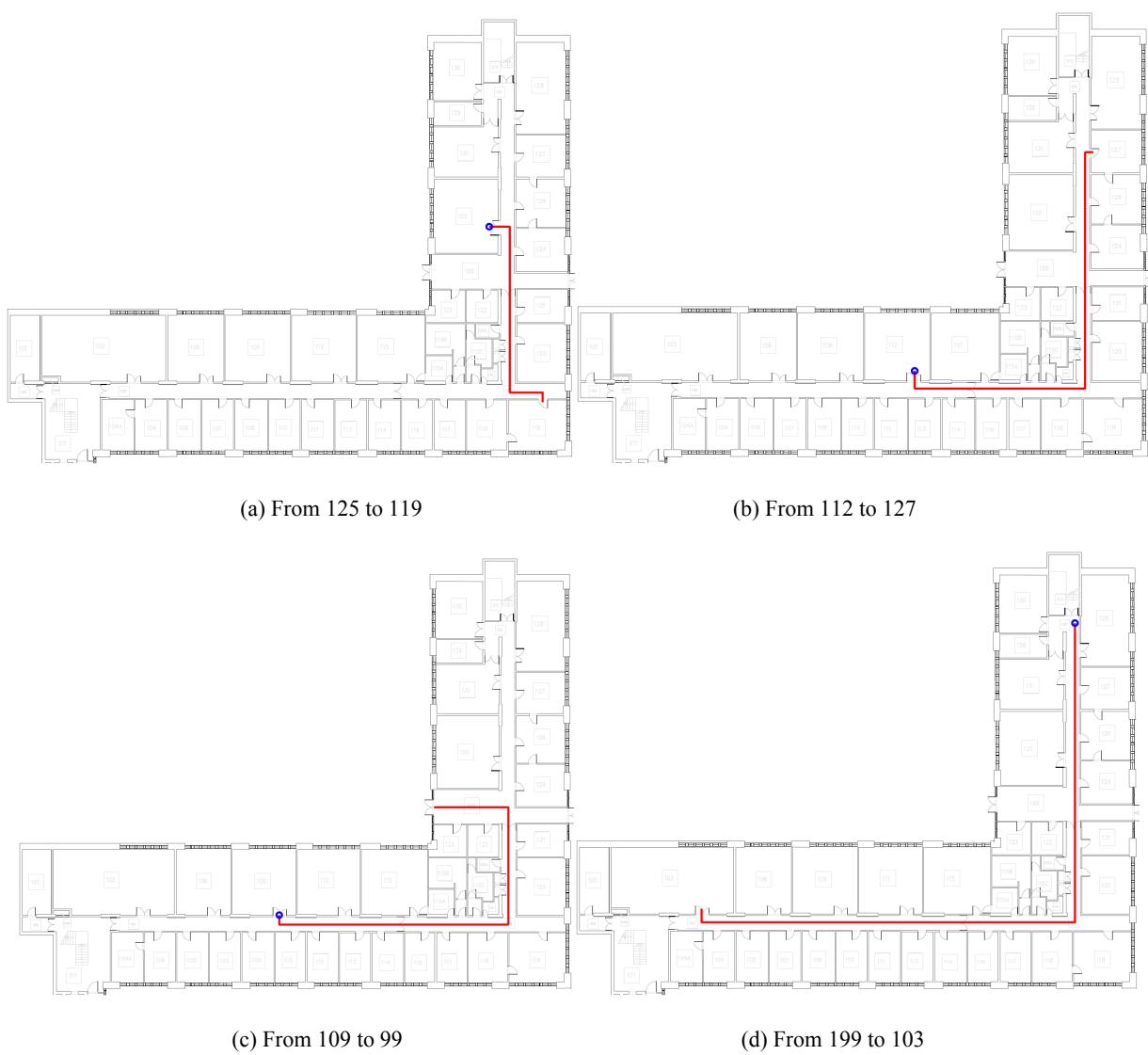
Figure 11 The ten fully observed test trajectories. The blue circle denotes the starting point (see online version for colours) (continued)**Figure 12** The ten partly observed test trajectories. The blue circle denotes the starting point (see online version for colours)

Figure 12 The ten partly observed test trajectories. The blue circle denotes the starting point (see online version for colours) (continued)



Figure 13 The three completely new test trajectories measured on a different floor. The blue circle denotes the starting point (see online version for colours)



4.3 The destination and route prediction evaluation

The first experiment evaluates the accuracy of the route and the destination prediction. Two observations stand out from this experiment. Firstly, in terms of destination prediction, the accuracy was not particularly high. At all 317 positions on the ten fully observed test trajectories, the system managed to suggest the correct destination for only 156 positions, which was less than 50% (see Figure 14). This result is understandable, because there are several training trajectories that start from the same position for each test trajectory. Thus, the system was not always able to pick the training trajectory that leads to the correct destination. Unsurprisingly, none of the remaining 13 partly observed and new test trajectories suggested the correct destination.

However, in terms of route prediction, each position is deemed correctly predicted as long as the predicted trajectory overlaps with the correct trajectory from the beginning up until the current position. For example, considering test trajectory 106-125, which goes from office 106 in the lower-left to office 125 in the top-right (see Figure 11 (b)), there are three training trajectories that start from the same office with this test trajectory but end in different positions (see Figure 15).

In this experiment, when the user's sequence was short, the system matched it to the shortest training route 106-115 (see Figure 15 (a)). Although the user's actual route was 106-125, this taken route overlaps the predicted route 106-115. In other words, the user was actually following the same trajectory up until this point. In terms of positioning prediction, this predicted route did not make much difference, which will be evaluated in the next section. When the user's magnetic sequence grew larger as he went past room 115 – the destination of the current predicted route, the system recalculated the DTW measure and found the correct matching route 106-125. Overall, at all 317 positions on the ten fully observed test trajectories, the system managed to suggested the correct trajectory for 286 positions, which was almost 90% (see Figure 16). This result also illustrated that all the wrong route predictions happened at early positions when the sequence was short. For this routine database, all test positions with at least nine MFVs found their correct routes.

The second experiment evaluates the value of the DTW measure between the test trajectories and the training trajectories. For all ten fully observed test trajectories, the DTW measure at each position on the test trajectory was

relatively low, with the highest measure recorded at just 12.6 (μT) (see Figure 17 (a)). This low measure is understandable because there are several training trajectories that match these fully observed test trajectories. For the ten partly observed test trajectories, the DTW measures were also low, up until the point where the user strayed away from all training trajectories, from which the DTW measure quickly grew up (see Figure 17 (b)). For the three completely new test trajectories, the DTW measures went up rapidly right from the very beginning with the biggest measure recorded at 116.7 (μT), because there was no training trajectories that match these new test trajectories (see Figure 17 (c)). Table 2 summaries the DTW measures for the three categories of the test trajectories.

This experiment also exposed how often the system changed trajectory during the journey, since it always picked the training trajectory with the lowest DTW measure to the current sequence at each position. Figure 18 pointed out that for this routine database, when the test trajectory had less than ten MFVs, the system switched routes about five times on average across all 23 test trajectories. In

contrast, when the length of the test trajectory was within [10, 20] MFV, the system switched route only two times on average. For all ten fully observed test trajectories, there were four routes being used on average throughout each test trajectory. Overall, the longer the test trajectory was, the more successful the system could match it to the correct training trajectory, and the less frequent the system switches routes.

Table 2 Comparison of the normalised DTW measure for the test trajectories

	Max DTW measure	Min DTW measure	Average DTW measure
Fully observed test trajectories	12.6	3.2	6.8
Partly observed test trajectories	84.7	3.3	31.2
Completely new test trajectories	116.7	5.4	84.1

Figure 14 Destination prediction accuracy with the routine database. Each horizontal line corresponds to the test trajectory. Each cross demonstrates a position along the journey where the destination prediction is correct

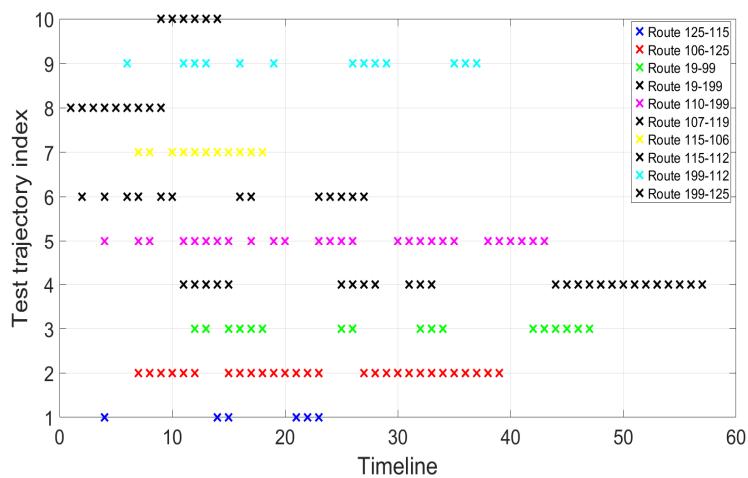
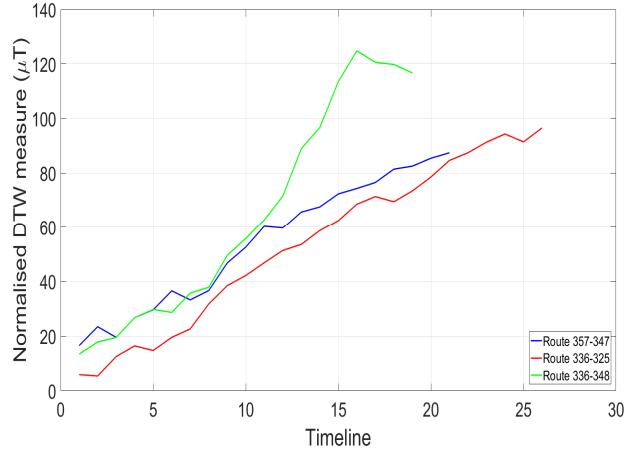


Figure 15 The three training trajectories that start from office 106. The blue circle denotes the starting point (see online version for colours)



Figure 15 The three training trajectories that start from office 106. The blue circle denotes the starting point (see online version for colours) (continued)



Figure 17 The minimum DTW measure at each position on the test trajectory throughout the journey (continued)

(c) New test trajectories. The DTW measure increases rapidly from the very beginning

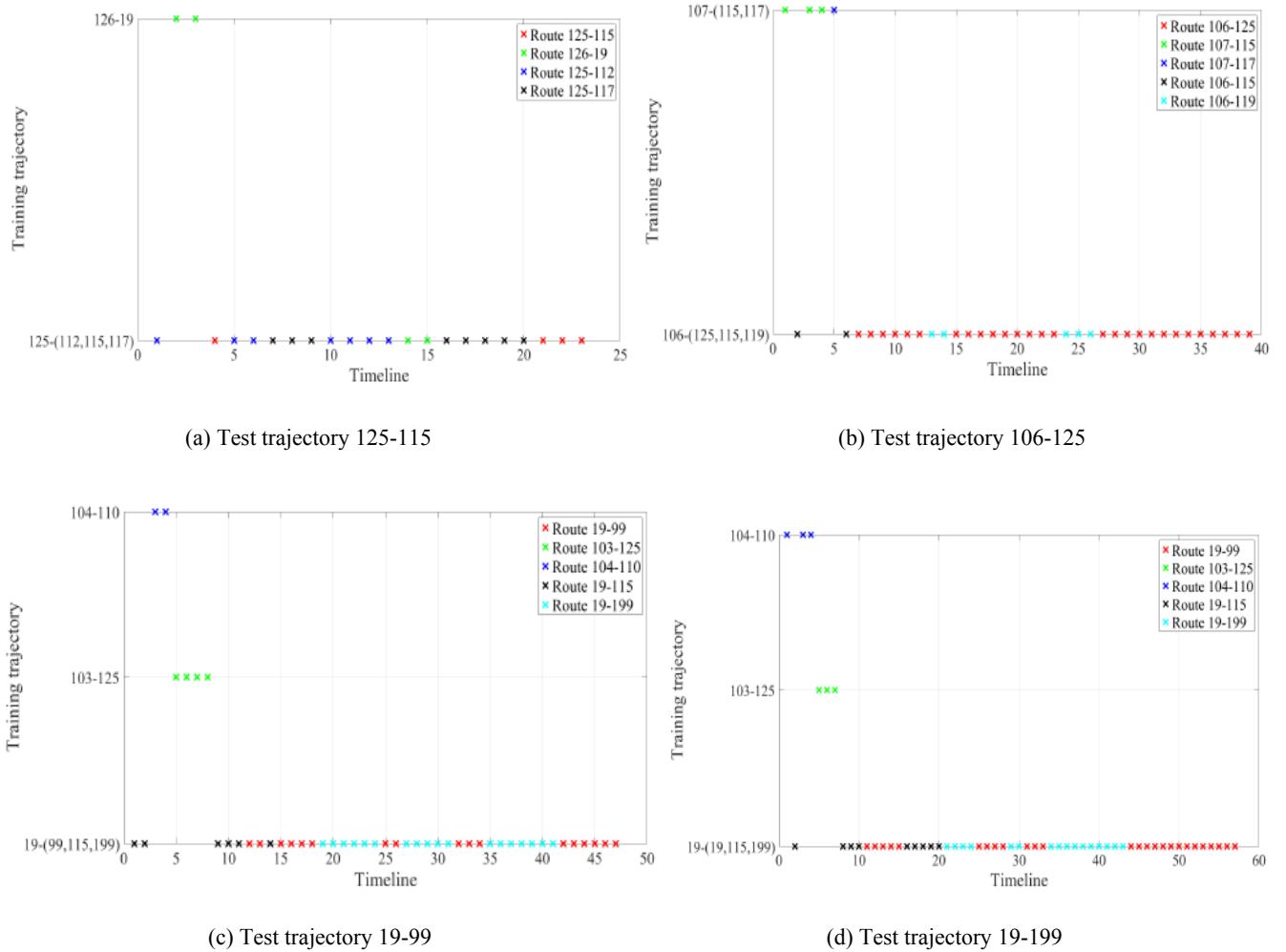
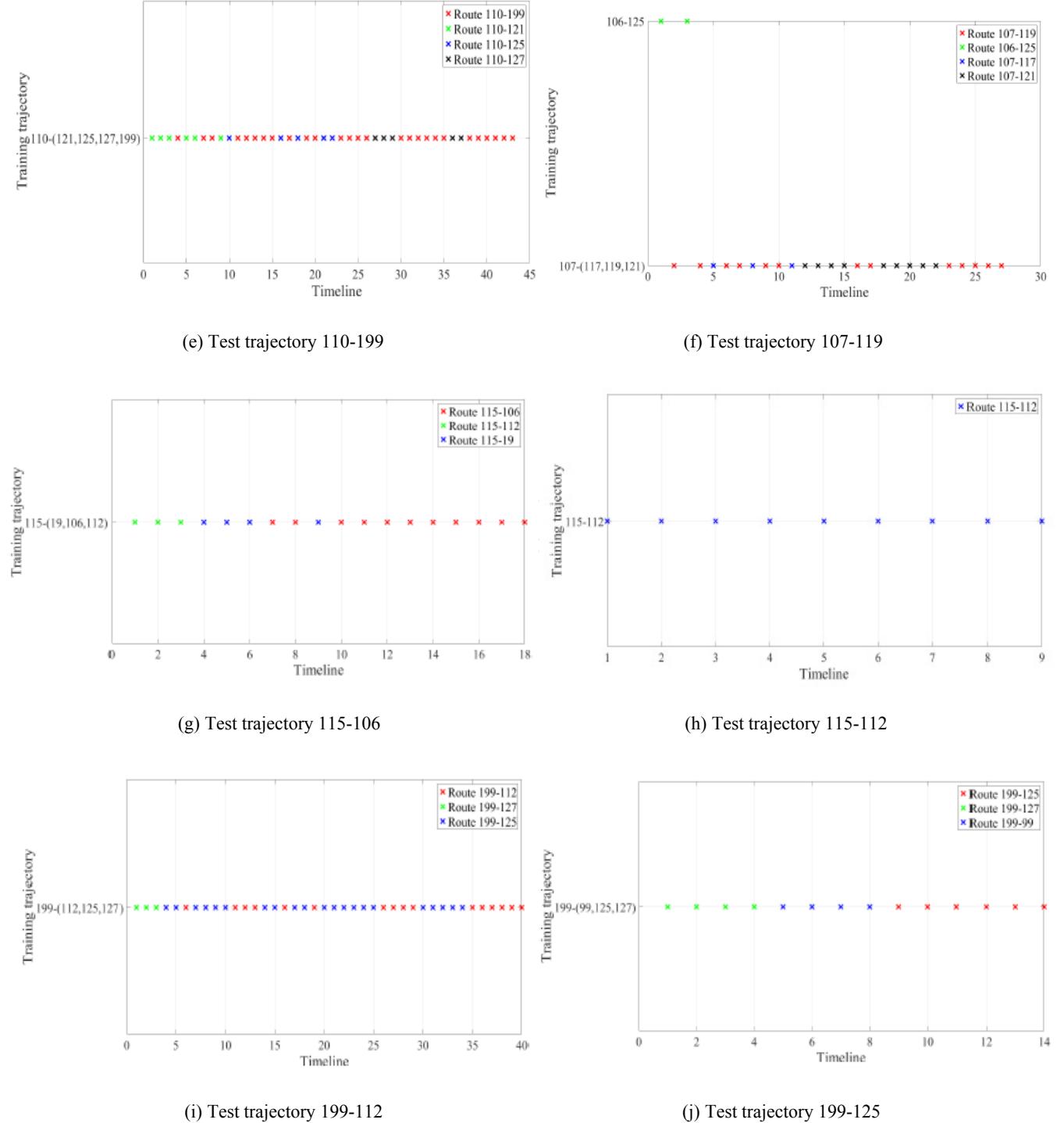
Figure 18 How often the system switches route, throughout the journey of the whole test trajectory

Figure 18 How often the system switches route, throughout the journey of the whole test trajectory (continued)

4.4 Single positioning estimation evaluation

Although the main purpose of the routine database is to predict the potential walking trajectory and the destination, it is interesting to investigate the positioning accuracy while the user is travelling on the predicted route. At any stage of the journey, the user's position will be regarded as the Cartesian label of the last position on the training trajectory returned by DTW. It is worth recalling that the OE-DTW variant of DTW employed in this paper tests different

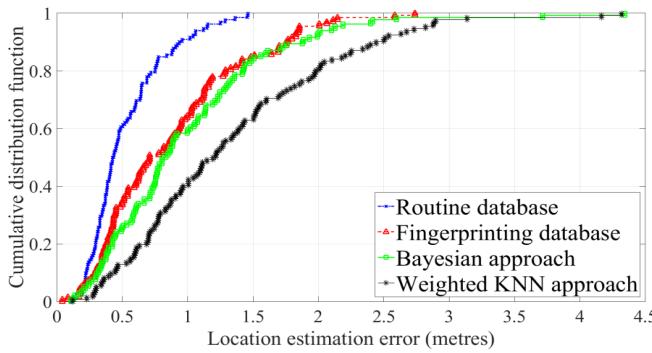
incomplete versions of the training trajectory by relaxing the endpoint constraint to find the best matched version. For this experiment, all 317 positions on the ten fully observed test trajectories will be used as test positions. To compare the positioning estimation with the fingerprinting database, the WiFi RSS vector at each test position was used to estimate the user position.

Figure 19 demonstrates a remarkable positioning accuracy using the routine database with up to 2.2 metre positioning error, 95% probability with the routine database, compared to

3.5 metres, 95% probability with the fingerprinting database. This result emphasises that when the system is certain of the user's walking trajectory, it can estimate the user's whereabouts using the individual positions recorded on the training trajectories much quicker with higher accuracy. In contrast, the fingerprinting approach must go through the whole training database to make prediction for every position.

There are two reasons why the estimated position using the routine database may not perfectly match the true position. Firstly, the training trajectory does not line up completely with the real-time trajectory. The version of the training trajectory returned by DTW may be a few measures before or after the real-time trajectory. Secondly, there are many training trajectories starting in the same position, the current predicted one is not necessarily the correct one.

Figure 19 Performance accuracy of single position estimation with the routine database



5 Related work

The earliest use of the magnetic field for indoor positioning was first seen in robotics navigation. It helps the robot orientate itself by maintaining a correlation to the magnetic North (Jung et al., 2013; Suksakulchai et al., 2000). However, the structure of most buildings greatly distorts the direction of the magnetic field (i.e. compass does not work reliably indoors).

Some early work suggested the use of the magnetic field with fingerprinting. However, they shared the same idea that the position's fingerprint is directly represented by the magnetic data (Chung et al., 2011; Li et al., 2012; Kim et al., 2012; Subbu et al., 2013; Haverinen and Kemppainen, 2009). This approach is challenging, as explained earlier, since the magnetic field only contributes a maximum of three measures per location. The work described in this paper makes use of the robust magnetic data via a sequence of measures.

Most earlier works in fingerprinting were considered in an active tracking context, to provide the positioning service to the user when requested (Bahl and Padmanabhan, 2000; Youssef and Agrawala, 2005; Weber et al., 2010; Wang et al., 2015). This paper provides an alternative approach to monitor the user's positions continuously in the background. The proposed system predicts not only the user's position, but also the walking route the user may take in advance. This opens up new possibilities for smart applications, which were not achievable with only a single position.

While the WiFi signal and the magnetic field data are popular for fingerprinting, they have also been previously applied in other fields, such as swarm intelligence, and hardware security (Abraham et al., 2006; Bi et al., 2014; Bi et al., 2015).

6 Summary and further work

Traditional fingerprinting has normally been considered as a tracking service to guide the user to his wanted destination. This paper considered fingerprinting as a sentient service for the regular users to allow the system to predict their intended walking route and the potential destination in advance. The paper showed that such objective was possible by using the magnetic field data to generate a routine database for the user. The major difference of the proposed approach and other fingerprint-based ones is that the real-time readings are accumulated as a sequence to distinctively represent the user's walking trajectory. Using this routine database, the system can predict the user position with higher accuracy than using just the fingerprinting database. The paper also demonstrated that the longer the system observes the magnetic field data, the more accurate the predicted routes are, and the less frequently the system changes route during the journey.

The limitation of the proposed approach is that it is more feasible for the buildings with narrow corridors (e.g. those used in the test bed). It was not designed to identify the user's position in a large area (e.g. in the middle of the lecture hall), for which the fingerprinting database will be more suitable.

Lastly, some practical discussions are outlined here. These discussions are not presented in the main body to maintain the focus of the paper.

- i *How does the system cope when the user temporarily turns off his phone, or when he takes a break during his journey?* When the walking route is broken, the system simply makes a single position-by-position prediction using the fingerprinting data set.
- ii *What happens if the user turns on the tracking app in the middle of his journey?* With the current design, the system predicts the trajectory at the beginning of the user's journey. It may be possible to make prediction midway, although this concept has not yet been examined in this paper.
- iii *Why not use a universal routine database for every user in a similar manner as the fingerprinting database?* Firstly, the paper assumes that most users often follow their own preferred path to navigate the building. Secondly, a universal database may contain too many trajectories, for which certain paths may not be used at all by some users.

We think this paper has laid the foundation for using the routine database to predict the walking trajectory. Some potential further works to improve this concept are outlined below.

- i The personal routine database can be further personalised by adding the approximate starting time for each journey. Depending on the user's timetable, he may favour one particular route at a certain time, and this extra information may improve the accuracy of the decision-making process. More importantly, the timetable also tells how often the event takes place, which can provide the initial value for the frequency of the trajectory. For example, the timetable indicates the user has four lectures per week, thus the $F(\text{office, lecture}) = 4$.
- ii It may also be useful to set a threshold for the user's real-time trajectory, so that the prediction engine only engages when the trajectory reaches a certain length. This approach has two benefits. Firstly, it guarantees that the accumulated user trajectory is long enough to avoid many initial similar matches with the training trajectories. Secondly, it avoids the scenario where the recording process is falsely triggered by the user who temporarily picks up the phone to make a call.
- iii Since the routine database is constantly updated with new WiFi RSS and magnetic field data, it is also a great resource for crowdsourcing. The training examples from the fingerprinting database may be updated with the latest data extracted from the individual routes in the routine database.
- iv It may be beneficial to reduce the number of times the system has to search through the routine database by introducing a DTW measure threshold. There are two scenarios for this threshold.
 - a) If the DTW measure between the current training trajectory and the real-time trajectory is below the threshold, the system believes that the user still follows the predicted route. Hence, it does not bother checking other training trajectories.
 - b) If the measure rises above the threshold, the system believes that the predicted trajectory is no longer correct. It recalculates the measure to other training trajectories to find a better match.

There are two parameters to consider here. Firstly, a DTW measure threshold needs to be decided in advance, so that the system knows when to drop the current predicted trajectory to find a better one. The challenges here are that if this DTW measure threshold is too low, the system keeps following the wrong training trajectory, although the user has turned to a new route. In contrast, if this DTW threshold is set too high, the system becomes too sensitive, and may drop the correct training trajectory because of the noises recorded in the real-time trajectory. Secondly, a time window parameter also needs to be considered to tackle the challenge with the magnetic noises. This time window ascertains that the system does not accidentally drop the correct training trajectory. Only when several low DTW measures are recorded within this window, the current training trajectory is dropped.

References

Abraham, A., Guo, H. and Liu, H. (2006) 'Swarm intelligence: foundations, perspectives and applications', in Nedjah, N. and Mourelle, L.D.M. (Eds): *Swarm Intelligent Systems*, Springer, New York, pp.3–25.

Bahl, P. and Padmanabhan, V.N. (2000) 'RADAR: an in-building RF-based user location and tracking system', *Proceedings of IEEE Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies*, Vol. 2, pp.775–784.

Bi, Y., Gaillardon, P-E., Hu, X.S., Niemier, M., Yuan, J-S. and Jin, Y. (2014) 'Leveraging emerging technology for hardware security – case study on silicon nanowire FETs and graphene SymFETs', *2014 IEEE 23rd Asian Test Symposium*, IEEE, Hangzhou, China, pp.342–347.

Bi, Y., Yuan, J.S. and Jin, Y. (2015) 'Beyond the interconnections: split manufacturing in RF designs', *Electronics*, Vol. 4, No. 3, pp.541–564.

Chung, J., Donahoe, M., Schmandt, C., Kim, I-J., Razavai, P. and Wiseman, M. (2011) 'Indoor location sensing using geomagnetism', *Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services*, ACM, Bethesda, MD, pp.141–154.

Dey, A.K., Abowd, G.D. and Salber, D. (2000) 'A context-based infrastructure for smart environments', *Managing Interactions in Smart Environments*, Springer, Dublin, pp.114–128.

Faragher, R. and Harle, R. (2013) 'SmartSLAM – an efficient smartphone indoor positioning system exploiting machine learning and opportunistic sensing', *ION GNSS*, Vol. 13, pp.1–14.

Giorgino, T. (2009) 'Computing and visualizing dynamic time warping alignments in R: the DTW package', *Journal of Statistical Software*, Vol. 31, No. 7, pp.1–24.

Haverinen, J. and Kemppainen, A. (2009) 'Global indoor self-localization based on the ambient magnetic field', *Robotics and Autonomous Systems*, Vol. 57, No. 10, pp.1028–1035.

Heiskanen, W.A. and Meinesz, F.A.V. (1958) *The Earth and Its Gravity Field*, McGraw-Hill Book Company, Inc., New York.

Holm, S. (2009) 'Hybrid ultrasound-rfid indoor positioning: combining the best of both worlds', *2009 IEEE International Conference on RFID*, IEEE, New York, pp.155–162.

Jung, J., Lee, S., Kim, H., Park, B. and Myung, H. (2013) 'Mobile robot relocation using ambient magnetic fields and radio sources', *2013 13th International Conference on Control, Automation and Systems (ICCAS)*, IEEE, Gwangju, Korea, pp.1766–1768.

Kim, S-I. (2011) 'Agent system using multimodal interfaces for a smart office environment', *International Journal of Control, Automation and Systems*, Vol. 9, No. 2, pp.358–365.

Kim, S-E., Kim, Y., Yoon, J. and Kim, E.S. (2012) 'Indoor positioning system using geomagnetic anomalies for smartphones', *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, IEEE, Sydney, Australia, pp.1–5.

Li, B., Gallagher, T., Dempster, A.G. and Rizos, C. (2012) 'How feasible is the use of magnetic field alone for indoor positioning?', *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, IEEE, Sydney, pp.1–9.

Martin, E., Vinyals, O., Friedland, G. and Bajcsy, R. (2010) 'Precise indoor localization using smart phones', *Proceedings of the International Conference on Multimedia*, ACM, Firenze, Italy, pp.787–790.

Müller, M. (2007) *Information Retrieval for Music and Motion*, Springer, Berlin.

Priyantha, N.B. (2005) *The Cricket Indoor Location System*, PhD Thesis, Massachusetts Institute of Technology.

Subbu, K.P., Gozick, B. and Dantu, R. (2011) 'Indoor localization through dynamic time warping', *2011 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, IEEE, Anchorage, AK, pp.1639–1644.

Subbu, K.P., Gozick, B. and Dantu, R. (2013) 'LocateMe: magnetic-fields-based indoor localization using smartphones', *ACM Transactions on Intelligent Systems and Technology (TIST)*, Vol. 4, No. 4, 27pp.

Subrt, L. and Pechac, P. (2012) 'Intelligent walls as autonomous parts of smart indoor environments', *IET Communications*, Vol. 6, No. 8, pp.1004–1010.

Suksakulchai, S., Thongchai, S., Wilkes, D. and Kawamura, K. (2000) 'Mobile robot localization using an electronic compass for corridor environment', *2000 IEEE International Conference on Systems, Man, and Cybernetics*, Vol. 5, pp.3354–3359.

Wang, H., Sen, S., Elgohary, A., Farid, M., Youssef, M. and Choudhury, R.R. (2012) 'No need to war-drive: unsupervised indoor localization', *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services*, ACM, New York, pp.197–210.

Wang, B., Zhou, S., Yang, L.T. and Mo, Y. (2015) 'Indoor positioning via subarea fingerprinting and surface fitting with received signal strength', *Pervasive and Mobile Computing*, Vol. 23, pp.43–58.

Want, R., Hopper, A., Falcao, V. and Gibbons, J. (1992) 'The active badge location system', *ACM Transactions on Information Systems (TOIS)*, Vol. 10, No. 1, pp.91–102.

Weber, M., Birkel, U., Collmann, R. and Engelbrecht, J. (2010) 'Comparison of various methods for indoor RF fingerprinting using leaky feeder cable', *2010 7th Workshop on Positioning Navigation and Communication (WPNC)*, IEEE, Dresden, pp.291–298.

Xie, H., Gu, T., Tao, X., Ye, H. and Lv, J. (2014) 'MaLoc: a practical magnetic fingerprinting approach to indoor localization using smartphones', *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ACM, Seattle, Washington, DC, pp.243–253.

Youssef, M. and Agrawala, A. (2005) 'The horus WLAN location determination system', *Proceedings of the 3rd International Conference on Mobile Systems, Applications, and Services*, ACM, Seattle, Washington, DC, pp.205–218.