International Global Navigation Satellite Systems Society
IGNSS Symposium 2015

Outrigger Gold Coast, Qld Australia
14-16 July, 2015

A Novel Approach for Wi-Fi Fingerprinting Using Logical Sequences of Intelligent Checkpoints

Guenther Retscher
Department of Geodesy and Geoinformation, Vienna University of Technology, Austria
+43 1 58801 12847, +43 1 58801 12894, guenther.retscher@tuwien.ac.at
Hannes Hofer
Vienna University of Technology, Austria
+43 1 58801 12804, +43 1 58801 12894, h.hannes@gmx.at

ABSTRACT

Commonly location fingerprinting is employed for Wi-Fi positioning. In the training phase RSSI scans have to be measured at known reference points (RPs) which are usually distributed in a grid within the area of interest. The grid has to be rather dense and therefore RSSI measurements are very labour consuming. The newly developed approach aims at a significant reduction of the number of required RPs. For that purpose particular points are selected which we call intelligent checkpoints (iCPs). To navigate a user to a certain room in a building certain iCPs have to passed on the way. Doors, stairs and corridors can serve as iCPs. Hence, a logical sequence can be derived due to the interdependence of these points along the way. iCPs are twofold intelligent, i.e., because of the fact that they depend on the selection of the points for the RSSI scans and because of the logical sequence in their correct order along the path depending on the building. While navigating then always the following iCP is known due to the vector graph allocation in the fingerprinting database. Further constraints between the iCPs are definable, e.g. the heading of the moving user can be logically defined while walking along a corridor or even the step length while climbing stairs. Thus only a small limited number of iCPs needs to be tested when matching the current RSSI values. The required processing time is significantly reduced. Furthermore the scanned Wi-Fi Access Points can be weighted on the RSSI value level. From field tests in an office building it could be seen that the iCP approach achieves a higher success rate for correct matching of the RSSI fingerprints than conventional approaches. In average correct matching results of 90.0 % were achieved using a joint Wi-Fi database including training measurements of all employed smartphones. An even higher success rate is achieved if the same mobile device is used in both phases. Thus the positioning accuracy can be crucially increased.

KEYWORDS: Location fingerprinting, training phase, reference points, intelligent checkpoints (iCPs), logical sequence.

1. INTRODUCTION

For indoor positioning a variety of localization technologies have been developed. Li and Rizos (2014) classified indoor positioning technologies into three categories. Firstly, designated technologies based on pre-deployed signal transmission infrastructure such as systems using infrared or ultrasonic signals, magnetic fields, Ultra Wide Band (UWB) or other RF-based systems. The second category are technologies based on 'signals-of-opportunity', i.e., systems using RF signals not intended for positioning, for instance, wireless fidelity (Wi-Fi), digital television, FM radio, mobile telephony, and others. The third group are technologies not based on signals. Dead reckoning (DR) using inertial sensors (accelerometers and gyroscopes) as well as vision/camera systems belong to this category.

Especially the use of Wi-Fi has become very popular as it uses signals which are available in many indoor environments and public spaces. The first system called RADAR developed by Bahl and Padmanabhan (2000) at the Microsoft Research Labs contributed to demonstrate the feasibility and usefulness of positioning systems for indoor spaces by exploiting the ubiquity of wireless local area networks. For Wi-Fi positioning the signal strength (i.e., Received Signal Strength Indicator RSSI) to the surrounding Access Points (APs) is measured. In this context, location fingerprinting is most commonly used as it is more robust to environmental effects on the RSSI than other methods. This is because the fingerprinting algorithm constructs a search space according to either a simulated environmental model (e.g., a model of the building) or previously-measured RSSI distributions in the radio maps. In the second case a training phase is conducted where the RSSI measurements or scans are previously measured on known reference points (RPs) distributed throughout the area of interest. The RSSI measurements define a so-called fingerprint on that particular RP. Then from all RSSI scans a radio map of RSSI fingerprints is constructed within the area of interest. The RSSI values of the RPs measured in the training phase are stored in a fingerprinting database. For the establishment of such a database the RPs are usually distributed in a regular grid throughout the area of interest. To achieve acceptable results for localization determination of a mobile user with positioning accuracies on the few meter level or at least to locate the user in the correct room in a building the grid has to be rather dense (see e.g. Li *et al.*, 2005). Such a construction of a fine radio map of RSSI distributions, however, leads to high workload and therefore RSSI scans for the training database construction are very labour consuming. The newly developed approach aims not only at achieving similar levels or an increase of positioning accuracies than conventional approaches but also at a significant reduction of the number of required RPs. In this paper the novel approach is presented and field experiments conducted in an office building analysed. It could be seen that the novel approach achieves a much higher success rate for correct matching of the RSSI fingerprints to the respective RPs than conventional approaches. In average correct matching rates of over 90.0 % were achieved when also considering the heading of the user and a weighting of the RSSI scans to the APs in addition.

The remainder of the paper is organized as follows: Firstly, the initial situation is briefly elaborated and the field test site and test arrangements are presented. Then the principle of operation of the intelligent checkpoint approach is introduced followed by a description and discussion of the major results of the Wi-Fi location fingerprinting experiments conducted in a multi-storey office building. Several different processing modes and variants are analysed and assessed in this section. Finally, a brief summary and outlook on future work concludes the paper.

2. INITIAL SITUATION

To reach a specific room in a multi-storey building certain waypoints must be passed along the way to the destination. Coming from outdoors, first an entrance has to be chosen and used. Then one will enter a foyer or similar area. To reach the next floor, either the stairs or an elevator must be used. Before one can reach the designated room one has to walk along a corridor. Thereby the route can be divided into waypoints which are dependent on each other. They have to be passed following a logical sequence to reach the designated destination. Doors, stairways and corridors can be considered as points or nodes along the way which define the possible path. Hence, obviously these waypoints have to be recognised and the logical connections between them in the process of navigation to the desired destination. We call these waypoints intelligent checkpoints (iCPs). In our novel approach, they are the particular points were the RSSI scans are performed in the training phase for location fingerprinting.

From measured RSSI scans on RPs established in a rather high density throughout the area of interest – which is a multi-storey office building in this study at hand – iCPs are selected from the field tests which can be identified and revealed very well. By the intelligent choice of these iCP waypoints the peculiarities of the building are considered as always bottlenecks must be passed along the way. These bottlenecks are formed by building structures such as entrances, corners, doors, walls, etc., which influence the RSSI measurements to the surrounding APs significantly. In contrast to common fingerprinting approaches where the RPs are often distributed in a regular grid the iCPs are chosen in an intelligent manner on important and well distinguishable decision points such as crossings, entrances, and other important waypoints. Besides this, they are these locations which must be passed and lead to new distinctive areas in the building. Hence, iCPs are twofold intelligent because of the fact that they depend on their intelligent selection for the RSSI scans and because of the logical sequence in their correct order along the path. Then a logical sequence can be derived due to the interdependence of iCPs along the way. In the following the test sites are presented and the analyses of extensive field experiments are discussed in detail.

3. FIELD EXPERIMENT SITE AND TEST ARRANGEMENT

The field test site is a multi-storey office building of the Vienna University of Technology (TU Wien), Austria. For this paper the test results of three different areas have been selected, i.e., the entrances and foyer on the ground floor, the staircase up to level 3 and hallway on the 3rd floor as well as selected test rooms along a corridor on the 3rd floor of the building (see Figures 1 to 3). In total 49 RPs were selected as candidates for the iCPs in these areas. The iCPs are represented as red dots in the three Figures. For the construction of the RSSI fingerprinting database on these points RSSI scans were performed repeatedly in four different orientations with two different smartphones. This resulted in 5726 independent RSSI scans[1]. Additionally, for the positioning phase, i.e., the phase in where the mobile user has to be located, 796 RSSI scans were measured on the tested waypoints. For the data acquisition an App was developed which logs the RSSI scans along with the heading information of the moving user obtained from the magnetometer of the used smartphones. In addition, the accelerometers were used to count the steps of the user. These observations, however, are not employed for the presented analyses in this paper.

---

[1] For the interested reader these observations stored in Matlab files may be downloaded from https://drive.google.com/folderview?id=0B2sMc_nmy1A1UmtCa0MxaUZaSFE&usp=sharing

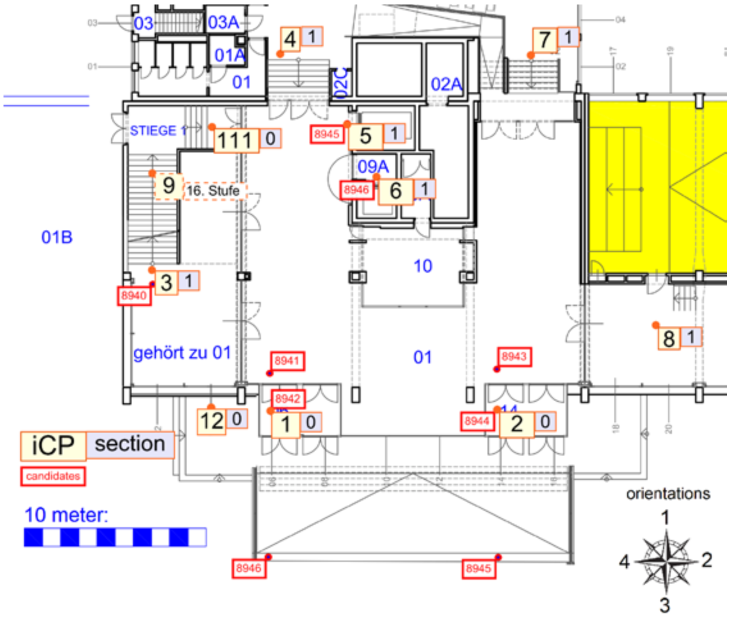

Figure 1. iCPs at the entrances, in the ground foyer and staircase of the multi-storey office building

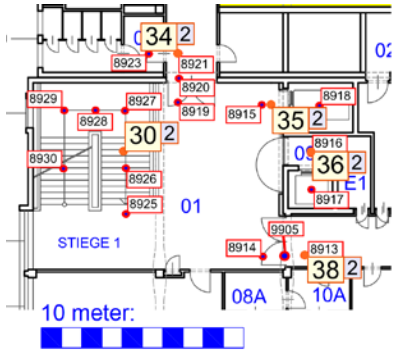

Figure 2. iCPs in the foyer and staircase in the 3rd floor

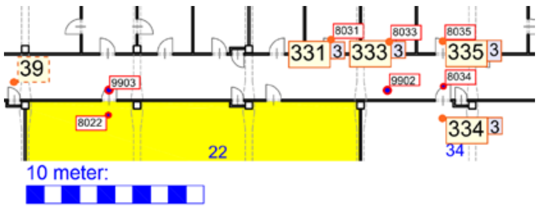

Figure 3. iCPs along the corridor in the 3rd floor and at entrances of four different test rooms

As a preliminary work RSSI scans were measured for the establishment of the fingerprinting database in the training phase throughout the test sites in the multi-storey office building to be able to choose representative iCPs.

4. OPERATIONAL PRINCIPLE OF THE iCP APPROACH

As mentioned above the iCPs are twofold intelligent. On the one hand, due to the choice of their location as they are selected on waypoints resulting from bottlenecks where a decision has to be made which way to go and, on the other hand, due to their logical sequence for two consecutive waypoints which are divided by structures such as, for example, doors, walls or glass walls and doors in between. Such a selection can guarantee that they are chosen in such a way that they can be distinguished clearly by Wi-Fi location fingerprinting. Therefore the number of possible candidates to be tested in the fingerprinting matching algorithm is significantly reduced compared to the common approach with a large number of RPs distributed in a regular grid. For such connections a vector graph allocation offers a suitable data structure. Then always the following iCP is known due to the vector graph allocation in the fingerprinting database. In this work, a simple logical sequence is applied in which only the iCPs are used which are located in a certain section of the building. Thus an optimization of the calculation method and weighting of the APs can be performed following a simple logic. As it is also possible to determine a user's heading with a smartphone with the in-built magnetometer this observation can be used additionally. Thus we tested whether it leads to better results if the heading is considered and then only the RSSI scan in this direction is used (note that the RSSI scans have been performed in four different orientations aligned to the axes of the building in the training phase). Besides comes along that the selected consecutive iCPs, for example, when walking through a door or along a corridor, can only be passed in certain directions according to the heading of the user.

5. MAJOR RESULTS FROM THE FIELD EXPERIMENTS

In this section selected results of the field experiments are presented and analysed. Firstly, results of different calculation variants from a Wi-Fi location fingerprinting approach using averaged RSSI values in the training database for each RP are presented in section 5.1 followed by results from the novel fingerprinting approach where either the direction of movement obtained from the magnetometer (section 5.2) or the use of all RSSI scans from the training database (section 5.3) or the logical sequence of iCPs (section 5.4) and an additional weighting of the AP scans according to their occurrence (section 5.5) is considered. A final comparison of the obtained results is given in section 5.6.

5.1 Wi-Fi location fingerprinting approach with averaged RSSI values

First of all it is investigated if whether the arithmetic mean or median is more suitable for averaging the RSSI scans measured in the training phase on a certain RP. Furthermore it is examined how the RSSI measurements on a particular RP to a certain AP should be considered in the DB if not in every scan epoch a RSSI value is obtained during the whole duration of the measurements in the training phase on that point. This effect is mainly caused due to high fluctuation and variation of the RSSI values. Hence, either a minimum value of -101 dBm is assigned for a certain RSSI scan or it is ignored and no value is stored in the DB. The second case is referred to with the term ‘Not a Number’ (NaN) following the respective Matlab function. Equations (1) and (2) show the relationship for the first case where an RSSI value of -101 dBm is assigned for APs where no RSSI is obtained in a certain epoch:

$$\text{Scan1} = \left[S1_{AP1}, S1_{AP2}, \ldots, S1_{APn}\right] \quad S1_{APx} \in \left\{\mathbb{Z}_{<0} \cap \{-100, -99, \ldots, -1\}\right\} \tag{1}$$

with

$$S1_{APx} = \begin{cases} \text{RSSI APx in dBm} & \text{if RSSI value to APx is obtained} \\ \text{-101 dBm} & \text{if RSSI value to APx not obtained} \end{cases} \tag{2}$$

where n is the number of APs given in the vector $AP_x = [AP_1, AP_2, \ldots, AP_n]$ in the test area.

The second case is described by Equations (3) and (4):

$$\text{Scan2} = \left[S2_{AP1}, S2_{AP2}, \ldots, S2_{APn}\right] \quad S2_{APx} \in \left\{\mathbb{Z}_{<0} \cap \{-100, -99, \ldots, -1\} \cup \text{NaN}\right\} \tag{3}$$

with

$$S2_{APx} = \begin{cases} \text{RSSI APx in dBm} & \text{if RSSI value to APx is obtained} \\ \text{NaN} & \text{if RSSI value to APx is not obtained} \end{cases} . \tag{4}$$

Due to this definition the vectors Scan1 and Scan2 have the same dimension. The APs are distinct-able by their BSSID (Basic Service Set Identification) number which either corresponds to the MAC address of the Wi-Fi network or which is randomly generated as a substitute. If for a certain AP no RSSI value is obtained in a certain epoch only the available measurements in the vector Scan2 (with NaN) are used for the calculations (see equation (3)). While using the vector with the minimum value of -101 dBm all RSSI values are given this value for not available scans in a certain epoch. This procedure can also be seen as a weighting of the observations as in the case where RSSI values to an AP cannot be measured in most epochs, the rather the average value will reach the minimum value of -101 dBm. For instance, if for a certain AP only once a RSSI value of -91 dBm out of ten scans is measured then the arithmetic mean would result in -100 dBm when using the Scan1 vector, whereas when ignoring the other nine scan values and using the Scan2 vector the arithmetic mean would result in -91 dBm. With such a differentiation of the calculation methods the effects of fluctuations and temporal variations of the RSSI scans can be efficiently considered. As the arithmetic mean and the median can be calculated following these two methods four different fingerprinting databases are formed, i.e., the mean DB (-101 dBm), mean DB (NaN), median DB (-101 dBm) and median DB (NaN). The success of the fingerprinting matching approach using the nearest neighbour (NN) algorithm in the positioning phase is termed as correct matching rate in the paper. It is as follows:

$$\text{matching rate} = \frac{\text{number of correctly assigned RSSI scans to RPs}}{\text{total number of all RSSI scans in positioning phase}}. \tag{5}$$

Rather than the positioning accuracies defined in metric units in the paper the matching rate defined in equation (5) is used to indicate the performance of the different calculation approaches. In this study only static observations on the RPs and iCPs have been considered and no kinematic measurements. Thus it is justified to indicate the performance of the fingerprinting approach using the matching rate on the static observed points instead of giving the positioning accuracy. If in the positioning phase an RSSI scan on a certain location is correctly matched to the respective iCP then the matching algorithm could find the correct location of the user.

Table 1 summarizes the results for the average matching rates of all tested calculation variants using different DBs and all directions of the measured four orientations. Note that RSSI scans in four orientations are commonly performed to encounter the influence of the human body of the mobile user in the training phase (compare e.g. Li *et al*., 2007). In total eight combinations are formed and included in the results summarized in Table 1. Two test DBs are formed in the same manner like the two vectors Scan1 and Scan2, i.e., DB1 with a minimum value of -101 dBm and DB2 with NaN. They are used to indicate test cases where one RSSI scan is assigned differently to a certain value in the DB. In the most commonly employed nearest neighbour (NN) matching algorithm the Euclidean distance for all test RSSI scans is calculated in the positioning phase from the DB values. As mentioned before two calculation methods are examined, i.e., either the calculation of the arithmetic mean or median. This resulted in eight calculation variants as shown in the columns in Table 1 whereby test scans with a DB with a minimum value of -101 dBm are highlighted in blue and the ones with the ignored measurements (NaN) in green. Furthermore the difference in the results is analysed if either a joint DB including RSSI scans of all mobile devices or two DBs containing only the scans of a particular smartphone are used. In the first two rows the results for smartphone SM1 (SM1 DB) and SM2 (SM2 DB) respectively are given whereas in the third row the results for the joint DB. The best results for the matching rate are highlighted in bold in these rows. Finally, the fourth row summarizes the mean matching rates.

In the multi-storey office building used in this study several Wi-Fi networks exist whereby one physical AP can have parallel several different names and MAC addresses (e.g. TUNET, and eduroam). For the first group of results all RSSI values for all readable APs are used. This is referred to as all AP RSSI values in Table 1. In the second group called unique AP RSSI values the matching rates are shown for the case where only RSSI measurements for one respective AP at the same location are used. In other words, for the calculation only one unique RSSI value for the same AP's physical location in the building is used.

<table>
<tr><th colspan="2"></th><th colspan="8">Results of test using all four orientations</th></tr>
<tr><th colspan="2"></th><th colspan="4">DB1 (-101dBm)</th><th colspan="4">DB2 (NaN)</th></tr>
<tr><th colspan="2"></th><th colspan="2">mean DB</th><th colspan="2">median DB</th><th colspan="2">mean DB</th><th colspan="2">median DB</th></tr>
<tr><th colspan="2"></th><th>-101dbm</th><th>NaN</th><th>-101dbm</th><th>NaN</th><th>-101dbm</th><th>NaN</th><th>-101dbm</th><th>NaN</th></tr>
<tr><td rowspan="4">all AP RSSI values</td><td>SM1 DB</td><td>66.2 %</td><td>46.9 %</td><td>38.3 %</td><td>35.3 %</td><td>20.4 %</td><td>20.7 %</td><td>72.5 %</td><td>72.0 %</td></tr>
<tr><td>SM2 SB</td><td>56.9 %</td><td>49.4 %</td><td>50.6 %</td><td>44.4 %</td><td>34.8 %</td><td>33.6 %</td><td>67.9 %</td><td>69.4 %</td></tr>
<tr><td>joint DB</td><td>42.6 %</td><td>30.2 %</td><td>41.8 %</td><td>34.4 %</td><td>17.0 %</td><td>16.1 %</td><td>55.5 %</td><td>53.8 %</td></tr>
<tr><td>mean matching rate</td><td>55.2 %</td><td>42.1 %</td><td>43.6 %</td><td>38.0 %</td><td>24.1 %</td><td>23.4 %</td><td>65.3 %</td><td>65.1 %</td></tr>
<tr><td rowspan="4">unique AP RSSI values</td><td>SM1 DB</td><td>61.0 %</td><td>48.9 %</td><td>52.1 %</td><td>39.3 %</td><td>24.2 %</td><td>23.7 %</td><td>62.7 %</td><td>61.0 %</td></tr>
<tr><td>SM2 SB</td><td>68.7 %</td><td>66.2 %</td><td>56.4 %</td><td>55.4 %</td><td>53.6 %</td><td>53.9 %</td><td>64.9 %</td><td>68.9 %</td></tr>
<tr><td>joint DB</td><td>63.8 %</td><td>55.4 %</td><td>55.8 %</td><td>45.0 %</td><td>33.0 %</td><td>33.2 %</td><td>51.4 %</td><td>51.6 %</td></tr>
<tr><td>mean matching rate</td><td>64.5 %</td><td>56.8 %</td><td>54.8 %</td><td>46.6 %</td><td>37.0 %</td><td>36.9 %</td><td>59.7 %</td><td>60.5 %</td></tr>
</table>

Table 1. Matching rates of Wi-Fi location fingerprinting with consideration of all four orientations

The best results are achieved if the minimum value of -101 dBm is assigned to not available RSSI scans during one measurement epoch when calculating the arithmetic mean. When using the median the best results are achieved if the not available RSSI measurements are either set to the minimum value (see results for smartphone SM1) or are ignored, i.e., if no values (NaN) are assigned (see smartphone SM2). If one unique representative RSSI value is chosen and stored as one RSSI scan for the particular AP location in the fingerprinting DB in average a 12 % higher matching rate is achieved (compare rows of the first group with all AP RSSI values and of the second group with unique RSSI AP values). As can be seen from column mean DB (-101 dBm) of DB1, a matching rate of 63.8 % for the joint DB and 68.7 % for a single DB for smartphone SM2 is achieved if a minimum value of -101 dBm is assigned in the fingerprint database for the RSSI scans which cannot be measured in a respective epoch of the training phase and which have therefore no measured RSSI value when using the arithmetic mean, on the one hand. On the other hand, the best results for the matching rate when calculating the median are achieved if the scan is ignored and therefore the database does not contain a RSSI value. In this case the matching rate resulted in 68.9 % (median DB (NaN) of DB2 for smartphone SM2). In general, the results indicate that the average matching rate is around 6.0 % higher if the same device is used in the training and positioning phase.

In general, the following conclusions can be drawn from the results in Table 1. The matching rate is better in most cases where

- own databases are considered for the two used smartphones SM1 and SM2,
- the database mean (-101 dBm) and test DB1 (-101 dBm) are combined if the arithmetic mean is used,
- the database median (-101 dBm) and median (NaN) with test DB2 (NaN) are combined for the median calculation, and
- the APs of multiple networks at one physical location are reduced to one unique AP if a joint DB for the two smartphones SM1 and SM2 is used.

The achieved results with this fingerprinting approach, however, are not very satisfying. Thus a new approach is developed using intelligent CPs (see section 5.4. and 5.5). But first it is analysed how a consideration of the heading of the moving user can improve the matching rate in the following section.

5.2 Fingerprinting in consideration of the heading of the mobile user

The heading of the moving user can be obtained from the measurements of the digital compass or magnetometer embedded in the smartphone and shall therefore be considered additionally. Then it is possible to determine the correct orientation from the four measured orientations in the training phase and only the orientation in direction of movement has to be tested in the positioning phase due to the known heading. Thus the processing speed is decreased. Furthermore an advantage is that not only averaged RSSI values from all four orientations are used but also the relevant RSSI scans for the certain user's orientation. The resulting matching rates are summarized in Table 2. The Table has a similar structure and includes the same calculation variants as Table 1.

		Results of test with known orientation							
		DB1 (-101dBm)				**DB2 (NaN)**			
		mean DB		**median DB**		**mean DB**		**median DB**	
		-101dbm	*NaN*	*-101dbm*	*NaN*	*-101dbm*	*NaN*	*-101dbm*	*NaN*
all AP RSSI values	**SM1 DB**	62.5 %	46.3 %	39.5 %	36.0 %	40.8 %	46.3 %	**70.5 %**	69.8 %
	SM2 SB	52.6 %	46.4 %	49.6 %	49.1 %	43.4 %	46.4 %	**66.4 %**	66.2 %
	joint DB	46.2 %	34.9 %	44.8 %	41.5 %	27.8 %	34.9 %	**56.9 %**	55.8 %
	mean matching rate	**53.8 %**	**42.5 %**	**44.7 %**	**42.2 %**	**37.3 %**	**42.5 %**	***64.6 %***	**63.9 %**
unique AP RSSI values	**SM1 DB**	**65.2 %**	55.9 %	50.6 %	46.6 %	44.6 %	55.9 %	58.7 %	57.4 %
	SM2 SB	**75.2 %**	65.7 %	63.4 %	48.4 %	58.9 %	65.7 %	64.7 %	63.7 %
	joint DB	**66.8 %**	62.1 %	56.4 %	53.1 %	45.4 %	62.1 %	55.2 %	50.8 %
	mean matching rate	***69.1 %***	**61.2 %**	**56.8 %**	**49.4 %**	**49.6 %**	**61.2 %**	**59.5 %**	**57.3 %**

Table 2. Matching rates for fingerprinting with consideration of the heading of the user

As can be seen from the results for the matching rates in Table 2 an improvement of in average 3.0 % for the arithmetic mean is obtained compared to the location fingerprinting approach where all orientations are used (compare with Table 1). If looking at the median an overall improvement of around 5.3 % is achieved. Thereby the highest improvement rates are achieved in the calculation variants which have the lowest matching rates (see the arithmetic mean calculation method for test DB2 in both Tables 1 and 2). Although only a slight improvement is achieved this method is better due to the fact that only one orientation has to be tested out of four in the positioning phase. Hence, only one quarter of the RSSI scans of the whole training measurements have to be tested which results in a significant reduction of the required processing time. Additionally, a further advantage is the fact that only iCPs shall be accepted if they are passed in a certain heading when walking to the destination. Such constraints improve the matching rate, for instance, if entrances to the building or corridors as well as doors to rooms are passed along the way.

5.3 Fingerprinting with use of all RSSI scans from the training phase

Besides averaging of the RSSI values on a particular RP using either the arithmetic mean or median the use of all scans from the training phase in the fingerprinting database DB is analysed. For the estimation of the most probable match the selection of the most best hits in k RSSI scans is performed whereby the Matlab function 'mode' is applied that determines the RSSI value which is contained most often in the scan vector. If two or more locations which are described by RSSI vectors equally often are present in the vector then the first one is chosen, i.e., the one which has the least differences of the RSSI values. The processing procedure is performed in the following four steps:

1. Calculation of all Eucledian distances to all RSSI scans in the DB,
2. Ascending sorting of the Eucledian distances and corresponding locations,
3. Selection of a set of the first *k* locations,
4. Allocation of the location which occurs most often in the selected set.

Using the determination of the most best hits in *k* RSSI scans similar results for the matching rates are obtained as for the methods where the arithmetic mean or median is calculated. The calculation expenditure, however, is higher than for the other method. Hence, a detailed discussion of the results is not given. But it can be mentioned that the best results are achieved for known heading of the user if unique AP RSSI values are considered and for DB1 the minimum value of -101 dBm is employed. Then a matching rate of 66.8 % is obtained for a joint DB of all two smartphones and 72.7 % for a single DB for one smartphone (i.e., SM2). But if this method is applied to the iCP approach and the concept of division of the area of interest into different building sections better matching results are achieved which goes along with a reduction of the RSSI scans to be tested. The section concept is introduced in the following section 5.4 and the achievable results are summarized in Table 4 and Figure 4.

5.4 Fingerprinting with consideration of the logical sequence of iCPs

From the total number of Wi-Fi RPs in the field test areas in the office building 17 iCPs are necessary which are located at decision points along the way to the destination and are therefore well distinguishable. The numbers of these points are highlighted in the large square boxes in Figures 1 to 3. For examination of the iCP concept the already measured RSSI scans on the RPs were used and additional test scans were carried out. Hence, for the analyses in total 1514 RSSI scans are used to build up the iCP database. For testing in the positioning phase 878 RSSI scans are available. The best matching rate for all investigated iCPs reached 66.8 % if a joint database of smartphone SM1 and SM2 is employed as can seen from Table 2. Table 3 and 4 show the matching rates of the already introduced methods and calculation variants. For all results the DBs of the two smartphones are combined. It was examined first, which matching rates are achieved if the total number of selected iCPs is observed. All these iCPs are matched correctly with a rate of 87.1 % in the case were the orientation is known as well as multiple AP readings are reduced to an unique AP at one physical location and the whole iCP database is used (see second group row whole iCP DB and column mean DB (-101 dBm) in Table 3). The mean matching rate can be raised in average to 90.4 % if the datasets are divided into sections for consideration of the logical sequence of iCPs to be passed (see mean matching rate for group 2 with unique AP RSSI values in Table 5). For that purpose sections following a logical sequence are defined regarding their occurrence. If navigating from outside the building to a certain room these sections have to be passed. Section 0 consists of the entrances to the building and section 1 aggregates the staircase or the elevator to the next floors. The iCPs of these two sections can be seen in Figure 1. Section 2 is in the staircase to the different floors which function as a junction. At one point this section is entered and at another left. The final section 3 consists of the entrances to the four different destination rooms. The location of these iCPS can be seen in Figure 3. The results for the matching rates in these different sections are summarized in Table 3 to 5 respectively.

Figure 4 shows graphically the most best hits in *k* RSSI scans in dependence on the parameter *k* in all four building sections. Thereby the values for *k* range from 1 to 25 and all variants with combined RSSI values for multiple APs are shown. All four sections are displayed in two different curves where one represents the DB1 (-101 dBm) and the second DB2 (NaN). In total 25 calculation runs were carried out whereby for each run the parameter *k* is increased by a factor 1 to find the suitable value for *k*. The orientation, however, is not considered but the differences can be viewed in detail in Table 4. In addition, the best matching rates (MR) are marked by red crosses in Figure 4 and their exact values can also be found in Table 4.

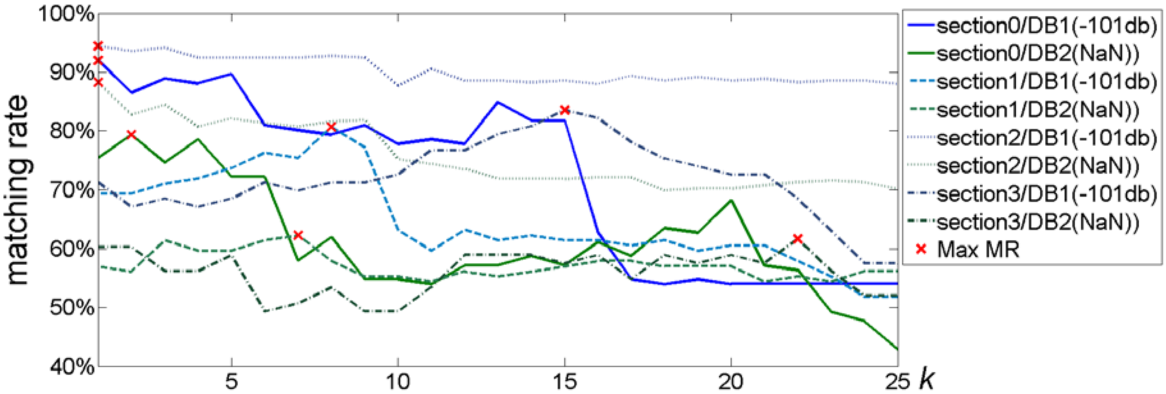

Figure 4. Most best hits in *k* RSSI scans in dependence of *k*

As can be recognized from Table 4 the consideration of the direction of movement does not lead to an improvement of the matching rate in most cases, except in section 2 and 3 with marginal differences of only up to 2.0 %. This is caused by the fact that the RSSI scans have been performed in the direction of movement only on the iCPs and thus the orientation faced either the corridor or a certain room. Thus the signals of other APs may be blocked due to the body of the test person. In the work of Hu (2013) it could be seen that typically Wi-Fi signals are blocked when people are located physically between the APs and the mobile devices. The main reason for this is that 2.4 GHz Wi-Fi signals can be greatly attenuated by water and the human body consist of about 70% water. Nevertheless, the orientation is indirectly considered in the calculation of the matching rate because all RSSI scans at all four orientations are included in the DB and no averaging has been performed. Thus also RSSI scans in the correct heading are compared in the case where all four orientations are used. This leads to smaller Eucledian distances than in the other three orientations. But one sees that only one quarter of RSSI scans in the DB have to be searched to achieve the same result. The division in sections reduces the number of RSSI scans to be searched even further as only the scans of a certain smaller number of iCPs need to be considered. Hence, the choice of the parameter *k* is also dependent on the number of RSSI scans which must be searched. If RSSI scans, for instance, are performed three times in all four orientations at a certain iCP then 12 scans belong to that location. Then *k* should be chosen only at most slightly greater. So had to go, actually, scans should be measured equally often at every location or equal RSSI values should be integrated into the DB for a particular location. It would be interesting to see from a detailed analysis of the effects resulting from a certain parameter *k* if there is a connection between *k* and the matching rate as well as the number of representative RSSI values for one iCP location. A question to be answered in further analyses is which value *k* must have if there is a certain number of RSSI scans for that particular location available and how many RSSI scans are meaningful. So it would be interesting to find the best average matching rate in relationship of the representative scan and *k*. One disadvantage, however, of this method is that all RSSI scans have to be considered in contrast to a DB where the scans are averaged for each iCP location and then only the number of iCPs have to be searched. For example one must calculate 5726 times the Eucledian distance from the whole DB values used in the study if one wants to search all scans on the 49 iCPs in the selected area of interest. If one calculates an average RSSI vector for all 49 locations then only 49 calculations are necessary. Even if only the required 17 iCPs are used for the most best hits algorithm still 1514 RSSI scans need to be used for the calculations of the Eucledian distances and 365 when considering only orientation 1 (Note that the used orientations are shown Figure 1). When looking at section 0 which has four iCPs still 314 RSSI scans need to be tested and in orientation 1 only 79 RSSI scans. Then the disadvantage described before can be annulled as the number of RSSI scans which need to be tested is substantially reduced. For future investigation it is presumably a good start point if one makes a selection of RSSI scans for a particular location which are representing the RSSI values well and in a variety of ways. Different times of the day with varying number of people in the building and situations such as how many of them currently use Wi-Fi may be considered to select representative RSSI scans which show large Eucledian distance differences.

		Results of test with known orientation							
		DB1(-101dBm)				DB2(NaN)			
		mean DB		median DB		mean DB		median DB	
		-101dbm	*NaN*	*-101dbm*	*NaN*	*-101dbm*	*NaN*	*-101dbm*	*NaN*
all AP RSSI values	**whole iCP DB**	**74.3 %**	60.1 %	62.2 %	54.9 %	25.1 %	60.1 %	70.8 %	70.7 %
	section 0 joint DB	**92.9 %**	67.5 %	76.2 %	57.1 %	69.8 %	67.5 %	89.7 %	86.5 %
	section 1 joint DB	58.8 %	48.2 %	56.1 %	53.5 %	32.5 %	48.2 %	**69.3 %**	65.8 %
	section 2joint DB	**82.7 %**	73.3 %	64.3 %	60.4 %	57.9 %	73.3 %	74.9 %	75.2 %
	section 3 joint DB	46.6 %	34.2 %	46.6 %	41.1 %	42.5 %	34.2 %	**64.4 %**	**64.4 %**
	section mean matching rate	**70.2 %**	**55.8 %**	**60.8 %**	**53.0 %**	**50.7 %**	**55.8 %**	***74.6 %***	**73.0 %**
unique AP RSSI values	**whole iCP DB**	**87.1 %**	83.6 %	75.7 %	69.8 %	58.5 %	83.6 %	68.5 %	66.1 %
	section 0 joint DB	**97.6 %**	96.0 %	83.3 %	61.9 %	85.7 %	96.0 %	83.3 %	81.7 %
	section 1 joint DB	**84.2 %**	78.9 %	70.2 %	58.8 %	63.2 %	78.9 %	64.0 %	52.6 %
	section 2 joint DB	88.0 %	**89.4 %**	78.8 %	79.7 %	74.1 %	**89.4 %**	71.6 %	69.6 %
	section 3 joint DB	72.6 %	74.0 %	69.9 %	60.3 %	**75.3 %**	74.0 %	69.9 %	63.0 %
	section mean matching rate	***85.6 %***	**84.6 %**	**75.6 %**	**65.2 %**	**74.6 %**	**84.6 %**	**72.2 %**	**66.8 %**

Table 3. Matching rates for iCP fingerprinting with consideration of the heading of the user [2]

		Results of most best hits in *k* RSSI scans			
		over all four orientations		with known orientation	
		DB1(-101dBm)	DB1(NaN)	DB1(-101dBm)	DB1(NaN)
all AP RSSI values	**whole iCP DB**	74.9 %	76.2 %	74.0 %	**76.9 %**
	section 0 joint DB	74.6 %	**90.5 %**	74.6 %	88.1 %
	section 1 joint DB	63.2 %	57.9 %	63.2 %	**62.3 %**
	section 2 joint DB	89.4 %	**90.3 %**	86.4 %	89.7 %
	section 3 joint DB	**64.4 %**	57.5 %	63.0 %	53.4 %
	section mean recognition rate	**72.9 %**	***74.0 %***	**71.8 %**	**73.4 %**
unique AP RSSI values	**whole iCP DB**	**86.2 %**	74.6 %	85.7 %	76.5 %
	section 0 joint DB	**94.4 %**	81.7 %	92.1 %	79.4 %
	section 1 joint DB	**82.5 %**	74.6 %	80.7 %	62.3 %
	section 2 joint DB	94.2 %	85.8 %	**94.4 %**	88.3 %
	section 3 joint DB	80.8 %	58.9 %	**83.6 %**	61.6 %
	section mean recognition rate	***88.0 %***	**75.3 %**	**87.7 %**	**72.9 %**

Table 4. Matching rates for iCP fingerprinting for most best hits in *k* RSSI scans without and with consideration of the user's heading [2]

[2] The term whole iCP DB means that all RSSI scans of the four sections are included in the DB and joint DB for the different sections stands for a combined database of smartphone SM1 and SM2 in each section

5.5 Fingerprinting with weighting of the RSSI measurements to APs

For the iCP concept it is always relevant to recognize the correct iCP when a certain way decision has to be made. Because the decision points are found mostly near to each other mainly the same APs are scanned at these locations. Then in most cases they differ not at all or only in a few scans. Hence, certain APs are more relevant for the differentiation than others. Therefore a weighting of the AP measurements is investigated and shall be performed. To test whether a certain AP is relevant the calculation of the Eucledian distance in the NN algorithm is complemented with a weighting of these value differences for each AP. This can be described by the following mathematical relationship for the Euclidean distance *d*:

$$d = \sqrt{(Sm_{AP1} - Si_{AP1})^2 \cdot g_{AP1} + (Sm_{AP2} - Si_{AP2})^2 \cdot g_{AP2}, ..., (Sm_{APn} - Si_{APn})^2 \cdot g_{APn}} \tag{6}$$

where $[g_{AP1}, g_{AP2}, ..., g_{APn}]$ is the weight vector, $[Sm_{AP1}, Sm_{AP2}, ..., Sm_{APn}]$ describes the measured RSSI vector for the positioning and $[Si_{AP1}, Si_{AP2}, ..., Si_{APn}]$ the reference for location *i* in the used fingerprinting DB. This calculation has to be done for all possible locations.

In this work a single weight vector is applied for all locations in one section. For further investigations, however, own weight vectors could be employed for every possible location. For the optimization of the weighting a simple heuristic optimization procedure is employed. The algorithm works following the nearest neighbour search (Weber *et al.*, 1998) whereby local optima can be found in an optimization problem. It is used for the whole dataset in each section. Therefore optimum weight vectors could be determined for certain sections. Besides that, the optimization uses always the best calculation variant. At the beginning of the optimization procedure a value of 1 is set for all weights. Note that no weighting is made if all weights remain 1. Afterwards the first AP is selected and its weighting for the first value in the weight vector is increased by a factor 1. If an improvement is achieved the current weight value is maintained and the weight is again increased by 1. This is made as long as no further improvement is achieved. Then the process continues with the next weight value. If the algorithm comes to the last weight vector the algorithm starts again with the first one. The termination criterion in the optimization procedure is the time. In this work a search of up to five minutes is performed to achieve better results. So far only this simple optimization procedure was used but the results show already a slightly improvement of a few percentage points for the matching rate. In future investigations more advanced and complex optimization procedures shall be examined which might result in a better improvement rate. An overview over heuristic optimization procedures may be found in Boussaïd *et al.* (2013).

By using the above mentioned optimization procedure an improvement of the matching result is achieved if the APs are weighted (see column with best matching rates with AP weighting in Table 5). Thus the matching rate increased to around 93.0 %. If a joint DB for section 0 is used the matching rate reaches 97.6 % for DB1 (-101 dBm) with the mean DB (-101 dBm) for the second group with unique AP RSSI values (see Table 3). The overall best matching rate of 99.2 % is also achieved in section 0 (compare Table 5). The highest improvement rate is achieved in section 3, i.e., from 83.6 % to 93.2 %. In this building section the most difficult situation along the path to the destination has to be treated as it consists of the entrances to the three neighbouring rooms 31, 33 and 35 on one side of the corridor which are only 2.5 m apart (compare Figure 3). From Figure 5 the matching rates for each of these iCPs and their improvement if a weighting is applied can be seen. In the worst case on iCP 335, only a matching rate of 46.7 % is achieved without and 73.3 % with weighting. The allocation of this point is interchanged several times with its neighbour iCP 333. Obvious is from the results that the consideration of the user's heading is very beneficial as the RSSI scans were performed in the direction of movement, i.e., along the corridor or orthogonal to it pointing inward into the room. Then the situation can occur that for certain APs lower RSSI values are obtained or signal blocking occurs due to the human body of the test person. In the other three sections in the building a higher average matching rate of around 93.0 % without weighting is achieved as the iCPs lie further apart. Here the improvement rate when weighting the APs is lower than in section 3 as already a much higher matching rate is achieved before weighting. Especially in section 0 which includes of the entrances to the building the iCPs are correctly allocated with a high matching rate of 99.2 % (see Table 5). In combination with step detection using the smartphone's accelerometers and heading determination with the magnetometer a continuous navigation to the next iCPs can be achieved. Then it is even easier to recognize the correct iCP along the way to the destination due to a continuous determination of the user's trajectory with inertial navigation. Future work and developments will concentrate on this integration.

		best matching rate	*best matching rate with AP weighting*
all AP RSSI values	**whole iCP DB**	76.9 %	76.9 %
	section 0 joint DB	92.9 %	95.2 %
	section 1 joint DB	69.3 %	72.8 %
	section 2 joint DB	90.3 %	90.3 %
	section 3 joint DB	64.4 %	71.2 %
	mean matching rate	**79.2 %**	**82.4 %**
unique AP RSSI values	**whole iCP DB**	87.1 %	88.1 %
	section 0 joint DB	99.2 %	99.2 %
	section 1 joint DB	84.2 %	84.2 %
	section 2 joint DB	94.4 %	95.0 %
	section 3 joint DB	83.6 %	93.2 %
	mean matching rate	**90.4 %**	**92.9 %**

Table 5. Matching rates for iCP fingerprinting without and with weighting of APs [2]

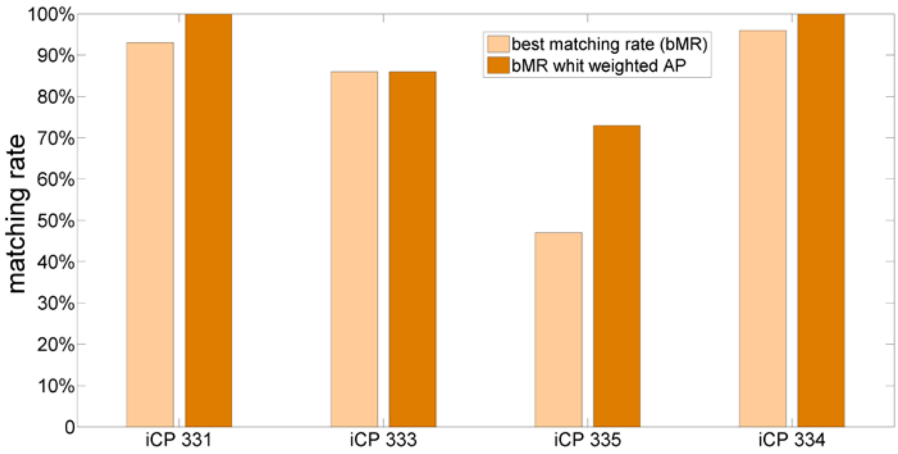

Figure 5. iCP matching rates in section 3 for the neighbour rooms 31, 33, 34 and 35 (Figure 3 shows their location along the corridor)

5.6 Comparison of the different approaches

The achieved results for the most commonly employed location fingerprinting approach are not very satisfying (see Table 1). Therefore a new approach is developed using intelligent CPs instead of training phase measurements on RPs usually distributed in a regular grid within the area of interest. In general, better results are achievable by consideration of the heading of the user and therefore the matching algorithm has to test the RSSI sans in only one orientation out of the four from the training phase. The improvements lie on average at a few percent points and they are achieved rather with methods which showed low matching rates in the beginning if the correct orientation is not considered. In the method where the most best hits for *k* RSSI scans are selected not relevant orientations can be excluded and the number of the RSSI scans to be tested is reduced by three quarter. In addition, iCPs are only accepted if they are passed in a certain orientation. As can be seen from the results summarized in Tables 3 to 5 significant better successful matching rates are achieved by using the iCP approach. The improvement results in around 18.0 % in the case of iCP fingerprinting with consideration of the heading of the user and around 24.0 % in the case of subdivision of the building into different sections. Then the best matching rates reach over 90.0 % in average and up to around 99.0 % for the determination of the correct building entrance, i.e., in section 0. This high success rates go along with a significant reduction of the required processing time. This is a mayor advantage compared to conventional fingerprinting approaches as not a high number of RSSI scans to RPs needs to be stored in the training database and tested in the positioning phase. As a matter of fact only a small number of carefully selected iCPs in their logical sequence have to be tested in the matching algorithm.

6. CONCLUSIONS AND OUTLOOK

Several development steps on the way to the iCP concept were shown and discussed in the paper. One of them is the matching of only one of the four orientations measured in the training phase due to the additional integration of the user's heading in the direction of movement obtained from the smartphone's magnetometer or digital compass. An App is developed in this study which logs the smartphone's inertial sensor readings together with RSSI scans. The consideration of the possible movement directions when building up the fingerprinting DB in the training phase is a major advantage as at entrances, for instance, only two possible movement directions have to considered, i.e., in and out, or when entering a room from a corridor only the orthogonal directions in and out are meaningful. Hence, iCPs are only accepted in the direction of possible movement. Thus the required acquisition time for collecting RSSI scans in the training phase is reduced by a factor of two as only two orientations instead of four have to be measured on each iCP. In addition, the large number of required RPs in the area of interest is significantly reduced. In the tests only 17 iCPs instead of 49 candidates are needed. In a previous study reported in Retscher (2012) training phase measurements on 75 RPs covering the same area on the 3rd floor in the office building were performed. Overall only 23.4 % of the test points could be located in the same room and another 25.0% in the neighbour room in the positioning phase using a conventional Wi-Fi fingerprinting approach. Using the novel iCP approach a much better successful matching rate of around 90.0 % is achieved as it could be shown for the tested neighbour rooms in building section 3. Especially, in the case where the Wi-Fi signals of a certain AP are blocked when the user is located physically between the AP and the mobile device better matching rates are obtained. The main reason for this is that it is then easier to distinguish between different orientations and locations. If one looks at the detailed results the use of the correct orientation leads to better matching rates in the case of weighting of the RSSI scans to the APs. Additionally, the processing time is decreased if the number of used APs is reduced in the DB. In our case the scan vectors had about approximately two-thirds smaller dimensions if the several appearing APs with different MAC addresses at a physical location are reduced to one.

The applied logical sequence of sections in the building is a simple attempt to reduce the number of possible user locations which have to be tested in the matching algorithm. For more complex buildings an advanced vector graph allocation can be applied and implemented. Further investigation and developments regarding such an approach are on the way. In addition, we work on the usage of the smartphone's accelerometers to provide continuous navigation to the next iCP. Then it is even easier to recognize the correct iCP along the way to the destination due to a continuous determination of the user's trajectory. As a matter of fact then the passing of a certain iCP has to be recognized if the iCP approach is combined with inertial navigation via dead reckoning using the smartphone's magnetometer for heading determination and accelerometers for step detection. Hence, only a certain number of iCPs come into consideration following their logical sequence along the way to the destination. Then only representative RSSI scans along the way are needed from the continuous scans. The representative RSSI scans of the possible iCPs have to be compared with the current scans to find the moment in time where the Eucledian distance shows a minimum. Exemplarily, the absolute differences of the RSSI vector in dependence of the time along a walked trajectory for a certain iCP are shown in Figure 6. The minimum value of the difference in the curve corresponds to the reference time obtained from manual recording when the iCP is passed along the way. In first experiments it could be seen that it is possible to find the iCPs with a divergence of less than four steps compared to the calculated steps using the accelerometer. Thus the iCP detection can be employed for absolute positioning to update the inertial navigation system and to reduce its drift rates. Further data acquisition and their analyses are currently on the way and will be published when more representative results are available.

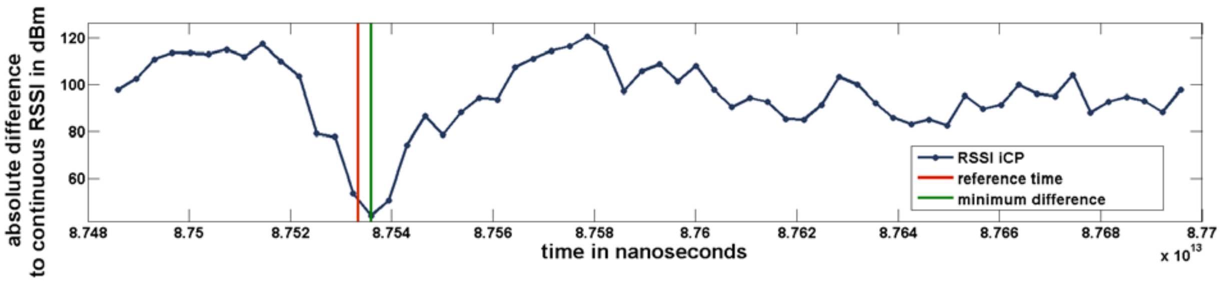

Figure 6. Eucledian distances of a certain iCP calculated from continuous RSSI scans while walking along a trajectory

REFERENCES

Bahl P, Padmanabhan V N (2000) RADAR: An In-building RF-based User Location and Tracking System. in: *Proceedings of the 19th Annual Joint Conference of the IEEE Computer and Communications Societies*. Tel-Aviv, Mar 26-30, Vol. 2, 775-784

Boussaïd I, Lepagnot J, Siarry P (2013) A Survey on Optimization Metaheuristics. *Information Sciences.* 237(0): 82-117

Hu B (2013) Wi-Fi Based Indoor Positioning System Using Smartphones. Master thesis, RMIT University, Melbourne, Australia, 77 pgs

Li B, Wang Y, Lee HK, Dempster AG, Rizos C (2005) Method for Yielding a Database of Location Fingerprints in WLAN, *Communications*, IEE Proceedings, Vol. 152, Issue 5, 580-586.

Li B, Kam J, Lui J, Dempster AG (2007) Use of Directional Information in Wireless LAN Based Indoor Positioning. in: *Proceedings of the International Global Navigation Satellite Systems Society IGNSS Symposium*, Sydney, Australia, December 4-6, 11 pgs

Li B, Rizos C (2014) Editorial: Special Issue International Conference on Indoor Positioning and Navigation 2012. Part 2. *Journal of Location Based Services*. 8:1: 1-2

Retscher G (2012) Wi-Fi Positioning with Smartphones. in: *Proceedings of the 9th International Symposium on Location-Based Services*, October 16-18, Munich, Germany, 9 pgs

Weber R, Schek H, Blott S (1998) A Quantitative Analysis and Performance Study for Similarity-Search Methods in High-Dimensional Spaces. in: *Proceedings of the 24rd International Conference on Very Large Data Bases*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA. 194-205