Indoor Navigation by WLAN Location Fingerprinting

Reducing Trainings-Efforts with Interpolated Radio Maps

Dutzler Roland & Ebner Martin
Institute for Information Systems and Computer Media
Graz University of Technology
Graz, Austria
r.dutzler@student.tugraz.at, martin.ebner@tugraz.at

Brandner Robert
Department of System Development
evolaris next level GmbH
Graz, Austria
robert.brandner@evolaris.net

***Abstract*—Due to the fact that smartphones are today already used by about one out of seven persons worldwide and their capabilities concerning hardware and sensors are growing, many different indoor navigation solutions for smartphones exist. The solution presented in this paper is based on Wireless Local Area Network Location Fingerprinting. Fingerprinting is a method where signals at a number of specific points are measured once and stored in a database that is needed to determine the position later on. Measuring each and every fingerprint makes the off-line phase a complex and very time-intensive process, especially for big buildings. The bigger the building, the higher is the effort to create the map needed for the on-line phase to determine the position of a device. In order to cope with this complexity, an approach for optimizing the off-line phase is realized. The system substantially lowers the number of positions at which fingerprint measurements have to be taken by identifying ideal positions. All other non-measured fingerprints are determined by using a form of the Log-Distance Path Loss Model.**

Keywords-Indoor Navigation; Location Fingerprinting; WLAN; Smart-phone; Log-Distance Path Loss Model

I. Introduction

Wireless Local Area Network (WLAN) is the most widely used and studied technology for realizing an Indoor Positioning System (IPS). There are also other ones, such as Bluetooth ([1], [2]), Infrared ([1], [3]), the earth's magnetic field [4] or dead reckoning approaches [5], which can also be combined to make positioning more accurate. WLAN based IPSs have two basic advantages compared to other systems: usually a number of WLAN Access Points (APs) are already deployed in buildings and no additional hardware has to be installed to realize an IPS based on WLAN. Also, most smartphones as well as other mobile devices like tablets support WLAN by itself.

WLAN IPS are usually based on Location Fingerprinting (LF). LF approaches are generally divided into two phases: an off-line or learning phase and an on-line or positioning phase.

In the off-line or learning phase, the target environment is split up into a grid of a certain width. At each and every grid point (GP) reference measurements of the surrounding APs are taken. These reference measurements are called Fingerprints (FPs) and consist of a location information as well as the Received Signal Strengths (RSS) of the different APs at this position. The collected FPs are stored in a so called Radio Map (RM).

In the on-line or location phase, a mobile device, which is aware of the stored RM of the target environment, can be used to determine the position. The device measures the signal strengths of the APs and compares it to the FPs stored in the RM in order to find its current position. There are many ways to calculate the position. A very simple is to compare the measured FP to the FPs stored in the RM and to find the three most similar ones by its mean square error. The position of the device then can be determined by calculating the balance point of the triangle that is made up by the location information of these three FPs.

Many research projects focus on improving the accuracy of the LF approach. The studies in [6] focus on improvements during the off-line or learning phase. The authors pointed out that additional direction information for every RSS measurement at one FP location might be helpful. The evaluation showed that a map of averaging RSS measurements from four directions to one FP showed the best values concerning accuracy and memory load, comparing to a map having a separate FP measurement for four directions at one location and another one considering a lower amount of all samples.

Kaemarungsi and Krishnamurthy [7] did some research on factors which influence the RSS values. They investigated that the user shadowing the signal from a device has an impact on the measured RSS values and so does the user's orientation. Measured RSS values at the same location tend to be different when examined over time, for example a day.

Zhao et al. [8] implemented a Differential Evolution (DE) algorithm to optimize the number and locations of APs for WLAN IPSs. It maximizes the variety of the FPs to improve location determination. Their tests show that symmetrical placement is worse than the calculated one by the DE algorithm. According to the authors it is better to place the APs in a 'zigzag' pattern. An increasing number of APs stops improvements when the system proposes to place them closely together.

It can be summarized that a general advantage of the WLAN LF approach is that it deals very well with factors that have influence on the RSS at a certain point, e.g. attenuation caused by walls. However, there is one major drawback for this method. The off-line phase is a complex and very time-intensive process, especially for big buildings. The bigger the building, the higher is the effort to create the map needed for the on-line phase to determine the position of a device. The number highly depends on the buildings size and on the accuracy the final system shall achieve. Common grid widths can range from one to five meters. For a large shopping center and a grid width of five meters the number of reference FPs that have to be measured for the RM can easily grow up to 2500 or more. In order to cope with these complexities, a new, very promising approach for optimizing the off-line phase has been realized.

Section II describes the goals of the research work. Section III explains how the implemented algorithm works. The evaluation and evaluation results are described in section IV. The last section concludes the research and points out problems that have to be investigated in future studies.

II. Objectives

The research work of this publication aims to develop a prototype application for an Android smartphone, which can be used for indoor navigation in large public buildings like shopping centers, hospitals or museums. The application shall determine the position of a device as accurate as possible and work robustly. In order to overcome the complexity of measuring each and every FP for the RM in the off-line phase, an algorithm has been implemented that substantially lowers the number of FP measurements. The algorithm suggests ideal GPs at which reference measurements should be taken on the basis of the building's plan, the building's dimensions, the exact course of the walls and the exact places of APs.

The idea behind the algorithm is that the fingerprint for each and every GP which distance to a certain AP is crossed by the exact same walls can be interpolated easily and exactly by measuring only one of them. Hence, the signals for these GPs travel through the same walls, their attenuation and other signal influence factors are nearly the same. Their signal strengths are of course dependent on the distance the GP is away from the AP. The next chapter explain the algorithm in more detail.

III. Method

The algorithm basically consists out of two parts:

- Part 1 – Determine Ideal GPs
- Part 2 – Interpolate FPs

In order to determine the GPs at which measurements have to be taken and subsequently interpolate the non-measured fingerprints, the algorithm needs the following information:

- True scale building's plan
- Dimensions of the building
- Exact course of walls
- Exact positions of APs
- GPs

A. Part 1 – Determine Ideal GPs

Part 1 of the algorithm determines ideal GPs at which measurements have to be taken. The requirement of good GPs depends on the number of APs, the grid density and hence the number of GPs, and the number of unique walls an AP crosses on the path to the GPs.

The algorithm takes the total of all GPs and determines for every GP and every AP the walls, which are crossed by the connecting line between GP and AP. All GPs whose connecting lines to a certain AP are crossed by the same walls, are grouped up and one of them gets suggested as ideal GP, which has to be measured, and is used to interpolate all other GPs of the group in part 2. The ideal GP is the one that is closest to the central point out of the GPs of one group, to calculate all surrounding ones as good as possible. Fig. 2 shows an example for one suggested GP and one AP. The solid line points to the suggested GP. All other GPs, which are marked by the dotted arrows, will be interpolated by the measurement of the suggested one, because the distances between these and the suggested GPs are crossed by the same wall.

B. Part 2 – Interpolate Fingerprints

After the suggested ideal GPs have been measured, in part 2 of the algorithm the non-measured FPs are determined by interpolation, with the help of the Log-Distance Path Loss Model for Line-of-Sight (LOS) environments, as described in [9]. The signals of all GPs which are covered by the exact same walls from an AP have nearly the same attenuation factors. So if the signal strength of one of those GPs is measured, the signal strength of all other surrounding GPs which are covered by the same walls can easily and very precisely be calculated by using a signal attenuation model for LOS environments, the Log-Distance Path Loss Model (1).

$$\overline{PL}(d) = \overline{PL}(d_0) + 10n\log\left(\frac{d}{d_0}\right) \qquad (1)$$

The wanted value PL(d) is the path loss or signal strength at a distance d. PL(d_0) is a measured signal strength or path loss at a reference distance d_0, whereas n is the path loss exponent. The exponent varies for every environment and is usually evaluated using empirical data. For LOS environments the path loss exponent is 2. The distances shall be given in meters. Fig. 1 shows how the ideal grid points are selected using pseudo code.

```
function findGridPoints()
Input: accessPoints, gridPoints, walls ∈ List
Output: foundGPs ∈ List
    for all accessPoints do
        wallHashMap ← newHashMap < String, List < Point >> ()
        for all gridPoints do
            for all walls do
                if wall.intersects(gridPoint, accessPoint) then
                    wallHash ← wallHash + wall.hashCode()
                end if
            end for
            wallHashMap.add(wallHash, gridPoint)
        end for
        for all wallHashMapEntries do
            goodGridPoint ← selectGoodGridPoint(wallHashMapEntry.value)
            if !foundGPs.contains(wallHashMapEntry.key) then
                foundGPs.add(wallHashMapEntry.key, goodGridPoint)
            end if
        end for
    end for
    return foundGPs
end function
```

Figure 1. Pseudo Code – Find Ideal Grid Points

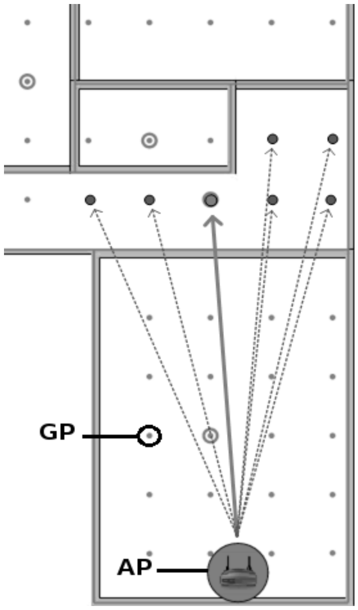

Figure 2. Part 1 – Determine Ideal GPs Example

The described method is another similar approach compared to the dominant path model described in [12]. The dominant path to any GP is the shortest path which crosses the lowest number of walls. The idea is that the signal is attenuated less when the dominant path crosses only one wall compared to the direct path that crosses three walls. The result achieved by the method described in this paper is at least the same or even more precisely. Difference is that the dominant path model suggests only one GP to measure for every AP in one room, but it also assumes that the attenuation along a long path is less compared to a direct short path.

Fig. 3 shows the differences between the calculated RSS attenuation in a LOS environment and real recorded attenuation.

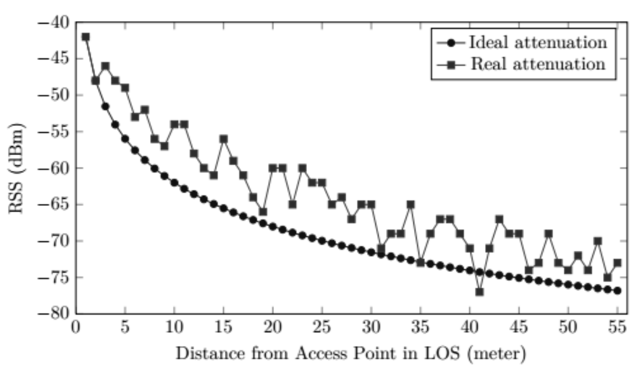

Figure 3. Differences between calculated RSS attenuation and real recorded attenuation

It is evident that the real recorded attenuation is smaller than the ideal, calculated signal attenuation.

IV. Experiments and Results

The evaluation of the prototype aimed to investigate the performance of the algorithm in complexity-savings as well as accuracy. The tests were done in three different surroundings: two Non-Line-of-Sight (NLOS) environments - one office building and a flat - and one LOS environment, a tennis court. For each test scenario two RMs were created to compare the accuracy. One RM where all GPs of the test environment were measured and one RM that used our optimization algorithm to create it. In each environment some Test Points (TPs) were randomly picked and afterwards the average error rate in meters is determined for all implemented positioning algorithms [10]:

- Simple Triangulation Algorithm (STA)
- Weighted Triangulation Algorithm (WTA)
- K-Nearest-Neighbor Algorithm (KNN)
- Weighted K-Nearest-Neighbor Algorithm (WKNN)

To make the evaluation as accurate as possible and hence ensure the integrity of all distances including distances to APs, GPs and TPs were exactly displayed by the application and measured in the real environment using a digital range measure. Fig. 4 shows a screenshot of the activity of the application that is used for creating the RM. The distances from the selected GP to the surrounding walls are displayed.

The creation of the RMs as well as the evaluation of the algorithm in the following test scenarios were done within one day. All FPs in the learning phase were measured in four directions and averaged with the user heading in the given direction and holding the smartphone in front of the chest. The TPs in the evaluation phase were measured in one random direction, also with the user holding the smartphone.

The following hardware was used for doing the evaluation:

- 4 APs – Cisco Linksys Wireless-G Broadband Router (WRT54GL)
- Smartphone – HTC Evo 3D (X515m)
- Silverline Ultrasonic Digital Range Measure

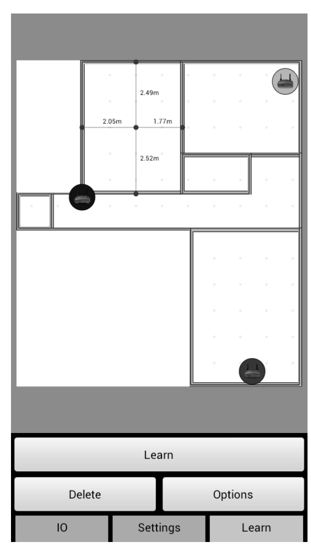

Figure 4. Screenshot of the Application

 ISBN: 978-1-61208-289-9

C. *Test Scenario I (Flat)*

Test scenario I is a typical small flat with two rooms, a kitchen, a bathroom, a toilet and a corridor of the size of 10.5m x 12.5m and a total area of 77.0m². In that scenario a grid width of 1.0m and 3 APs were used, which results in a total of 66 GPs. The number of ideal GPs that were suggested by the optimization algorithm was 22, which results in savings of 66%.

The evaluations were done at ten equally distributed TPs. The overall accuracy achieved by the Measured-RM (M-RM) was slightly better with used all algorithms. At two TPs the positioning accuracy reached with the Interpolated-RM (I-RM) was even better. This could be due to measuring inaccuracies caused by the test person covering signals from surrounding APs. The mean accuracy averaged over all TPs and algorithms by the M-RM was 1.66m and 1.89m with the I-RM. If we keep in mind that only one-third of all GPs were measured for the I-RM the error increase of 23cm is a very encouraging result. Fig. 5 shows the mean positioning error averaged over all TPs for each algorithm and every RM. The performances reached by the different algorithms are nearly the same for the M-RM. For the interpolated one, the WTA achieved the best overall accuracy. Fig. 6 shows the absolute RSS differences between the M-RM and the I-RM. The highest difference at that scenario is at the marked circled position in Fig. 6. The absolute RSS difference at this point is -13 dBm. Hence, the overall differences between both RMs are low.

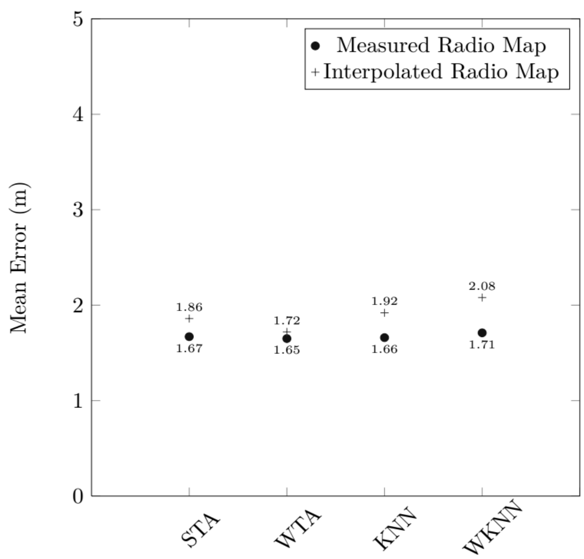

Figure 5. Mean Positioning Error of all Algorithms of Test Scenario I

D. *Test Scenario II (Office Building)*

Test scenario II is a typical office environment with some small offices and meeting rooms. The environment is 21.0m x 15.0m in size and covers an area of about 200m². A grid width of 1.3m and 3 APs were used in this scenario, which results in a total of 85 GPs. The optimization algorithm suggested 47 GPs, which are savings of 45%.

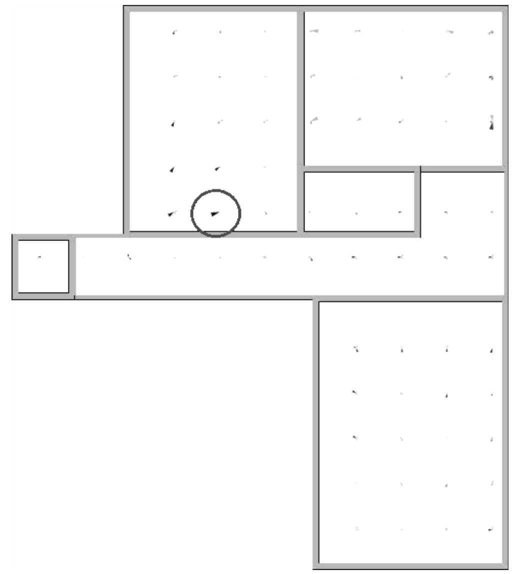

Figure 6. Absolute Fingerprint Differences between M-RM and I-RM of Test Scenario II

To evaluate the accuracy between the fully M-RM and the I-RM created with the optimization algorithm, measurements at 12 equally distributed TPs were taken. The results are similar to the results of test scenario I. The overall mean error for the M-RM was 2.11m and 2.54m for the I-RM. This means that the error that occurred due to only measuring about the half of all GPs and interpolating the other ones is 43cm or about 20% higher. Fig. 7 shows the mean errors of all implemented algorithms for both Radio Maps. The absolute FP differences are shown in Fig. 8.

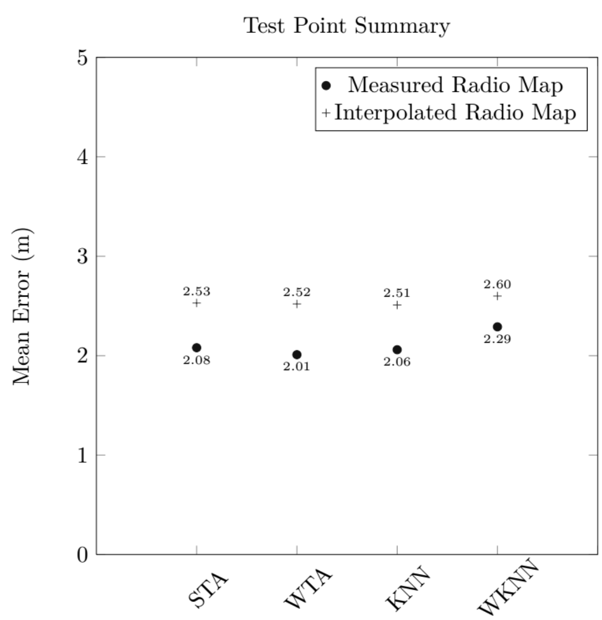

Figure 7. Mean Positioning Error of all Algorithms of Test Scenario II

 ISBN: 978-1-61208-289-9

Figure 8. Absolute Fingerprint Differences between M-RM and I-RM of Test Scenario II

E. Test Scenario III (Tennis Court)

The last test aimed to test the performance of the algorithm in a fully LOS environment and was done outdoors on two tennis courts. The area was 36.0m x 36.0m in size which result in a total area of 1296m². For the evaluation 4 APs and a grid width of 5.0m was used, which leads to a total number of 49 GPs. Because they are no walls in that scenario the optimization algorithm suggested only one GP in the mid of the environment for the I-RM, which are savings of 98%.

Measurements from 13 equally distributed TPs were taken. The overall accuracy for both RMs were substantially lower compared to the results from Test Scenario I and II. The mean error for the M-RM was 4.44m and 1.14m higher for the I-RM, which is 5.58m. Fig. 9 shows the mean errors for all algorithms and both RMs. Fig. 10 again shows the FP differences between the M-RM and the I-RM.

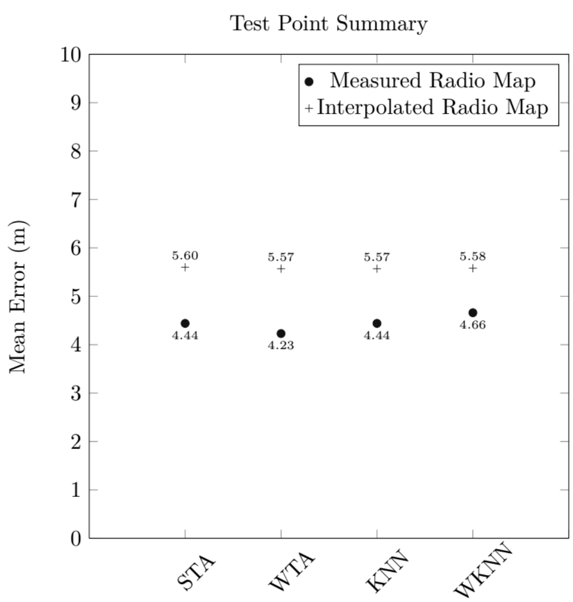

Figure 9. Mean Positioning Error of all Algorithms of Test Scenario III

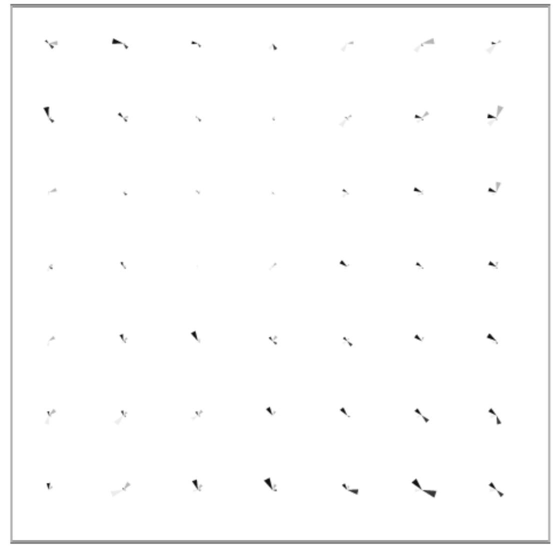

Figure 10. Absolute Fingerprint Differences between M-RM and I-RM of Test Scenario III

F. Summary and Overview

Table I gives an overview over the test scenarios, the environment they represent, their dimensions, the total area they cover, the number of APs that were used for the evaluation and the selected grid width. Table II then summarizes the performance of the implemented optimization algorithm. It states the complexity savings as well as the accuracy differences between the M-RM and the I-RM.

TABLE I. OVERVIEW TEST SCENARIOS DESCRIPTION

Scenario	Environ ment	Dimensions	Area	# APs	Grid Width
I	NLOS	10.5m x 12.5m	77.0m²	3	1.0m
II	NLOS	21.0m x 15.0m	200m²	3	1.3m
III	LOS	36.0m x 36.0m	1296m²	4	5.0m

TABLE II. OVERVIEW TEST SCENARIOS EVALUATION RESULTS

Scenario	# GP	# Ideal GP	Savi ngs	∅ Acc. M-RM	∅ Acc. I- RM	∅ Acc. Deviatio n
I	66	22	67%	1.66m	1.89m	+23cm
II	85	47	45%	2.11m	2.54m	+43cm
III	49	1	98%	4.44m	5.58m	+1.14m

The evaluation in NLOS environments showed a very satisfying overall performance. The mean error in test scenario I and a grid width of 1.0m for the M-RM of all positioning algorithms is 1.66m. The mean error for the I-RM is 1.89m and hence only 23cm or 13.98% higher. At test scenario II, another NLOS environment, and a grid width of

1.3m the mean error for the M-RM is 2.11m and 2.54m for the I-RM, which is a difference of 43cm or 20.37%, although there were high deviations at some TPs of nearly up to 5m (in some cases only for the I-RM or even both RMs). At test scenario III, a LOS environment, and a grid width of 5m the mean error for the M-RM is 4.44m and 1.14m or 25.60% higher, which is 5.58m, for the I-RM. In that scenario there were also some TPs at which the mean error grew up to over 8m. A further test aimed to determine the influence of the grid width on the positioning accuracy pointed out that a grid width of 2.5m is best. Lower as well as higher grid width has a negative impact on the resulting positioning accuracy.

V. CONCLUSION AND FUTURE WORK

The overall results are promising although there are some points that have to be further investigated. The WLAN LF approach has some basic restraints compared to other solutions. The RM of one environment is usually created once but the actual signal strengths may change over time because they are influenced by many factors. Furthermore every change in the environment can have an impact on the signal distribution of the WLAN routers. Walls, doors or even new or differently placed furnitures will change the received signal strengths at some points in the environment.

The tests were carried out using the same routers as well as the same devices for every test. One test case shows that different devices receive slightly different signal strengths from the same AP at the same place, depending on the built-in hardware. The evaluations do not care about what happens when different or other routers are used. Different routers have different antennas and hence do not cover the same area or have the same signal attenuation over the same distance. If different routers are used for creating one RM, a constant value might have to be added for every single AP to make the optimization algorithm accurate.

The approach was only designed and tested in 2-dimensional areas. It can be used for multi-story buildings by identifying the storey by the strongest signals, but the currently developed algorithm does not consider APs located on different storeys for optimization.

The evaluation revealed some incidents that might have an influence on the accuracy and have to be further investigated and improved. One very basic but important thing is the distribution and number of APs, which was topic of the studies [8] and [11].

At test scenario 3 there is only one GP out of a total number of 49 GPs used for the I-RM. The area is 36m x 36m in size. At some TPs the mean error for the I-RM is dramatically higher compared to the measured one. A final test pointed out that measuring more GPs (for example at least every 10 meters) and lowering the grid width improve the results arbitrarily.

Another issue that has to be reconsidered in detail is the Log-Distance Path Loss Model and the path loss exponent. An evaluation showed that the exponent could be slightly lower, probably this also depends on the APs.

The research work shows that the implemented optimization algorithm for the off-line learning phase of the WLAN LF approach is a legitimate way to reduce the complex, time-consuming way of establishing a RM for big buildings. The GPs at which measurements have to be taken can be substantially lowered and in addition an acceptable positioning accuracy reached compared to a fully M-RM. As already mentioned the evaluations revealed some factors that have to be considered and explored in future work to increase the positioning accuracy: The number as well as the distribution of APs, the maximum distance between two measured grid points to make the interpolation more accurate, the influence of different hardware – mobile devices as well as routers - and the calculation using the Log-Distance Path Loss Model. Especially the influences on the Path Loss exponent should be a starting point for further work.

ACKNOWLEDGMENT

I would like to express my gratitude to evolaris next level GmbH for offering and allowing me the work on this exciting topic of research.

REFERENCES

[1] Y. Gu, A. Lo, and I. Niemegeers, “A Survey of Indoor Positioning Systems for Wireless Personal Networks”, IEEE Communications Surveys and Tutorials, Volume 11, Number 1, 2009, pp. 13 – 32.

[2] G. Fischer, B. Dietrich, and F. Winkler, “Bluetooth Indoor Localization System”, Workshop on Positioning, Naviagtion and Communication, 2004, pp. 147 – 156.

[3] R. Want, A. Hopper, V. Falcão, and J. Gibbons., “The Active Badge Location System”, ACM Transactions on Information Systems (TOIS), Volume 10, Issue 1, 1992, pp. 91 – 102.

[4] IndoorAtlas, “Ambient magnetic field-based indoor location technology”, Technical Report, 2012.

[5] S. Beauregard and H. Haas, “Pedestrian dead reckoning: a basis for personal Positioning”, 3rd Workshop on Positioning, Navigation and Communication (WPNC), 2006, pp. 27- 35.

[6] H. Leppäkoski, S. Tikkinen, and J. Takala, “Optimizing Radio Map for WLAN Fingerprinting”, Ubiquitous Positioning Indoor Navigation and Location Based Service (UPINLBS), 2010, pp. 1-8.

[7] K. Kaemarungsi and P. Krishnamurthy, “Properties of Indoor Received SignalStrength for WLAN Location Fingerprinting”, Conference on Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous), 2004, pp. 14-23.

[8] Y. Zhao, H. Zhou, and M. Li, “Indoor Access Points Location Optimization Using Differential Evolution”, International Conference on Computer Science and Software Engineering, 2008, pp. 382-385.

[9] J. Mulligan, “A Performance Analysis of a CSMA Multihop Packet Radio Network”, Master Thesis, Polytechnic Institute of New York, 1997.

[10] H. Liu, H. Darab, P. Banerjee, and J. Liu, “Survey of Wireless Indoor Positioning Techniques and Systems”, IEEE Transactions on Systems Man and -Part C: Applications and Reviews, Volume 37, Number 6, 2007, pp. 1067 – 1079.

[11] O. Baala, Y. Zheng, and A. Caminada, “Toward Environment Indicators to Evaluate WLAN-Based Indoor Positioning System”, International Conference on Computer Systems and Applications, 2009, pp. 243 – 250.

[12] H. Wang, H. Lenz, A. Szabo, J. Bamberger, and U. Hanebeck, “Enhancing the Map Usage for Indoor Location-Aware Systems”, International Conference on Human-Computer Interaction (HCI 2007), 2007, pp. 151 – 160.