

Hyperbolic Location Fingerprinting: A Calibration-Free Solution for Handling Differences in Signal Strength

Mikkel Baun Kjærgaard and Carsten Valdemar Munk
University of Aarhus, Department of Computer Science
Aabogade 34, DK-8200 Aarhus N
Email: [mikkelbk|cvm]@daimi.au.dk

Abstract

Differences in signal strength among wireless network cards, phones and tags are a fundamental problem for location fingerprinting. Current solutions require manual and error-prone calibration for each new client to address this problem. This paper proposes hyperbolic location fingerprinting, which records fingerprints as signal-strength ratios between pairs of base stations instead of absolute signal-strength values. The proposed solution has been evaluated by extending two well-known location fingerprinting techniques to hyperbolic location fingerprinting. The extended techniques have been tested on ten-hour-long signal-strength traces collected with five different IEEE 802.11 network cards. The evaluation shows that the proposed solution solves the signal-strength difference problem without requiring extra manual calibration and provides a performance equal to that of existing manual solutions.

1 Introduction

Location Fingerprinting (LF) based on signal strength is a promising location technique for many awareness applications in pervasive computing. LF has the advantage of exploiting already existing network infrastructures, like IEEE 802.11 or GSM, and therefore avoiding extra deployment costs and effort. LF is based on a database of pre-recorded measurements of signal strength, denoted as location fingerprints. A client's location can be estimated from the fingerprints by comparing these with the current measured signal strength. Clients can be in the form of, e.g., a tag, a phone, a PDA, or a laptop.

A fundamental problem for LF systems is the differences in signal strength between clients. Such signal-strength differences can be attributed to inequalities in hardware and software and lack of standardization. For IEEE 802.11 dif-

ferences above 25 dB have been measured for same-place measurements with different clients by Kaemarungsi [5]. Such differences have a severe impact on LF systems' accuracy. Our results show that signal-strength differences can make room-size accuracy for the Nearest Neighbor algorithm [1] drop to unusable 10%.

Current solutions for handling signal-strength differences are based on manually collecting measurements to find mappings between signal strength reported by different clients. Such manual solutions are: (i) time consuming because measurements have to be taken at several places for each client; (ii) error prone because the precise location of each place has to be known; (iii) unpractical considering the huge number of different IEEE 802.11 and GSM clients on the market. For instance, due to such issues the company Ekahau maintains lists of supported clients [3]. Solutions have been proposed by Haeberlen et al. [4] and Kjærgaard [6] that avoid manual measurement collection by learning from online-collected measurements. However, both of these solutions require a learning period and they perform considerably worse in terms of accuracy than the manual solutions.

This paper proposes *Hyperbolic Location Fingerprinting (HLF)* to solve the signal-strength difference problem. The key idea behind HLF is that fingerprints are recorded as signal-strength ratios between pairs of base stations instead of as absolute signal strength. A client's location can be estimated from the fingerprinted ratios by comparing these with ratios computed from currently measured signal-strength values. The advantage of HLF is that it can solve the signal-strength difference problem *without* requiring any extra calibration. The idea of HLF is inspired from hyperbolic positioning, used to find position estimates from time-difference measurements [2]. The method is named *hyperbolic* because the position estimates are found as the intersection of a number of hyperbolas each describing the ratio difference between unique pairs of base stations. We have evaluated HLF by extending two well-known LF techniques to use signal-strength ratios: *Nearest Neighbor* [1]

and *Bayesian Inference* [4]. The HLF-extended techniques have been evaluated on ten-hour-long signal-strength traces collected with five different IEEE 802.11 clients. The traces have been collected over a period of two months in a multi-floored building. In our evaluation the HLF-extended techniques are compared to LF versions and LF versions extended with a manual solution for signal-strength differences.

We make the following contributions: (i) we show that signal-strength ratios between pairs of base stations are more stable among IEEE 802.11 clients than absolute signal strength; (ii) we propose the novel idea of HLF and show that the HLF-extended LF techniques perform clearly better than their LF versions and equal to their manual-solution-extended LF versions; and (iii) we show that the HLF-extended techniques place the same requirements as LF techniques on common parameters.

The paper is structured as follows: signal-strength ratios are quantified to be more stable than absolute signal strength among IEEE 802.11 clients in Section 2. The definition of HLF and the extension of two well-known LF-techniques are presented in Section 3. The results of evaluating the HLF-extended techniques for five different IEEE 802.11 clients are then given in Section 4. Afterwards, a discussion of the results are given in Section 5 and Section 6 discuss related work. A conclusion and a discussion of further work are given in Section 7.

2 Signal-Strength Differences

For IEEE 802.11 signal-strength differences can mainly be attributed to the standard's lack of specification of how clients should measure signal strength [5]. In the standard, signal strength is specified as the received signal-strength index with an integer value between $0, \dots, 255$ with no associated measurement unit. The standard also states that this quantity is only meant for internal use by clients and only in a relative manner. The internal use of the value is for detecting if a channel is clear or for detecting when to roam to another base station. Therefore, IEEE 802.11 client manufacturers are free to decide what their interpretation of signal-strength values is. Most manufacturers have chosen to base signal-strength values on dBm values. However, different mappings from dBm values to the integer scale from $0, \dots, 255$ have been used. The result of this is that most signal-strength values represent dBm values with different limits and granularity. However, inequalities in hardware also attribute to the problem.

This paper explores the use of signal-strength ratios between pairs of base stations. The following definitions are needed: $B = \{b_1, \dots, b_n\}$ is an ordered set of visible base stations and $O = \{o_1, \dots, o_m\}$ a finite observation space. Each observation o_i being a pair of a base

station $b \in B$ and a measured signal-strength value $v \in V = \{v_{min}, \dots, v_{max}\}$ according to a discrete value range. For the range of V the following restriction is necessary: $v_{max} < 0$. The signal-strength ratio r is defined for a unique base station pair $b_i \times b_j \in B \times B$ with the constraint $i < j$ for uniqueness. The signal-strength ratio r can be computed from two observations $o_i = (b_i, v) \in O$ and $o_j = (b_j, y) \in O$ as follows:

$$r(o_i, o_j) = \frac{v}{y} \quad (1)$$

However, because the signal-strength ratios are non-linear with respect to changes in either of the signal-strength measurements, normalized log signal-strength ratios are used. These are calculated from the signal-strength ratios as follows:

$$nlr(o_i, o_j) = \log(r(o_i, o_j)) - \log\left(\frac{1}{v_{max}}\right) \quad (2)$$

where the last term normalizes the ratios in order to keep them on a positive scale. When we refer to signal-strength ratios in the rest of the paper it will be in their log-normalized form.

2.1 Data Collection

For our analysis and evaluation data have been collected at a two-floored test site covering $2256 m^2$ and offering an 802.11 infrastructure with 26 reachable base stations. Signal-strength data have been collected as continuous traces with five different IEEE 802.11 clients, which are listed in Table 1. The five clients have been picked to cover different manufactures, options of antennas and operating systems. For each client three separate 40-minute traces have been collected, totaling about 10 hours of data. The traces were collected over two months and for each client the three separate traces were collected at different days and time of day to make sure the data was affected by temporal variations. Each entry in the traces consist of a time stamp, measured signal strength to surrounding base stations, and current ground truth. The ground truth was manually specified by the person collecting the trace by clicking on a map. The area of the test site were divided up into 126 clickable cells, with an average size of $16 m^2$, corresponding to rooms or parts of hallways, and spanning two floors. The cells approximately represent a coarse grained four meter fingerprinting grid. The people collecting the traces walked at moderate speeds, with several pauses through the test site on both floor levels, as illustrated for one trace in Figure 1. Signal strength were measured with a sampling rate of 0.5 Hz for the Fujitsu Siemens Pocket Loox 720 and 1 Hz for the four other clients.

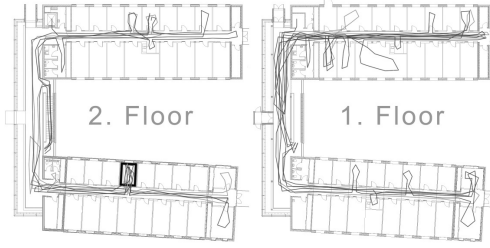


Figure 1. Path for one 40-minute client trace.

Table 1. Evaluated IEEE 802.11 clients

| Client name | Antenna | OS / Driver |
|---------------------------------|-----------|--------------------------------------|
| Apple AirPort Extreme | In laptop | Mac OS X (10.4) / OS provided |
| D-Link Air DWL-660 | In card | Windows XP / D-Link 7.44.46.450 |
| Fujitsu Siemens Pocket Loox 720 | In PDA | Windows Mobile 2003 SE / OS provided |
| Intel Centrino 2200BG | In laptop | Windows XP / Intel 10.5.0.174 |
| Orinoco Silver | In card | Windows XP / OS provided (7.43.0.9) |

2.2 Stability of Signal-Strength Ratios

If normalized log signal-strength ratios should be able to solve the signal-strength difference problem they have to be more stable than absolute signal-strength values among IEEE 802.11 clients. To quantify if this is the case the variations in absolute signal strength and signal-strength ratios have been analysed among different IEEE 802.11 clients. The analysis is based on statistics calculated from the collected traces. To make the statistics directly comparable the presented values have been converted to percentages of mean values.

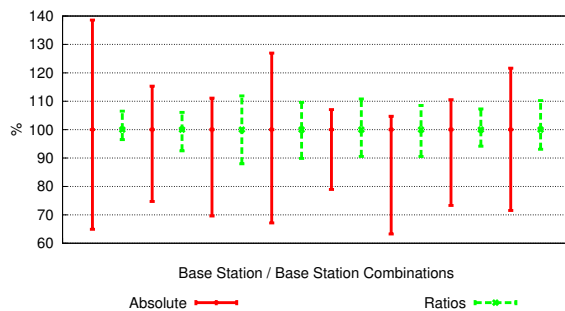


Figure 2. Absolute versus Ratios

The analysis uses trace data for all five clients from the black-rectangle-highlighted cell on Figure 1. The calculated

statistics from this trace data are shown in Figure 2. The figure shows the minimum and maximum values of absolute signal strength and signal-strength ratios for base stations and combinations, respectively. For the first base station the clients' absolute signal-strength values are at anytime at most 35.1% below and 38.6% above the mean absolute signal strength for this base station. For the first base station combination the signal-strength ratios are at any time only 4.5% below and 6.5% above the mean signal-strength ratio for this combination. Looking at all base stations and combinations the results show that the variations are only $\pm 10\%$ for signal-strength ratios but $\pm 20\%$ for absolute signal strength. Similar results were obtained in an analysis using data from all cells contained in the traces. The results confirm that signal-strength ratios vary less between IEEE 802.11 clients than absolute signal strength. Furthermore, because the used signal-strength traces were collected spread out over two months the signal-strength ratios are also shown to be stable over time.

3 Hyperbolic Location Fingerprinting

This section presents the extension of two well-known LF-techniques to HLF. The main change is the replacement of absolute signal-strength with signal-strength ratios. This change affects both the representation of location fingerprints and the calculation of location estimates. The extended techniques are the techniques of Nearest Neighbor [1] and Bayesian inference [4]. Both techniques are in this paper applied for cell-based localization, i.e. locations are represented as cells. A cell may correspond to a room or a part of it, or a section of a hallway. The following definitions are needed: $C = \{c_1, \dots, c_n\}$ is a finite set of *cells* covered by the location system, a *sample* s is a set of same-time same-place observations, one for each visible base station and a *fingerprint* f is a set of samples collected within the same cell.

3.1 Nearest Neighbor

A common *deterministic* LF technique calculates the nearest neighbor in Euclidian space between a client's measured samples and the fingerprints in the database [1]. The cell with the lowest Euclidian distance is picked to be the current one of the client. In the nearest-neighbor calculations each fingerprint is represented as a vector with entries for each visible base station. Each entry contains the average signal-strength for a base station computed from the samples of the fingerprint.

To extend this technique to HLF, both the fingerprint representation and the nearest-neighbor calculation have to be changed. The HLF fingerprint representation has entries for each unique pair of visible base stations in the fingerprint.

The entries of the vector are computed as the average signal-strength ratio from the fingerprint's sample set. Let f_{c_x, b_i} denote the set of observations from the fingerprint taken in cell c_x that refers to base station b_i . Each entry of a fingerprint representation vector v for a cell c_x and unique base station pair $b_i \times b_j$ can be computed as follows:

$$v_{c_x, b_i \times b_j} = \frac{1}{n} \sum_{o_i \in f_{c_x, b_i}} \sum_{o_j \in f_{c_x, b_j}} (nlr(o_i, o_j)) \quad (3)$$

where n is the number of observation combinations. An example with three base stations is given in Table 2. The table includes both the LF average absolute signal-strength and the HLF average signal-strength ratios.

Table 2. Example of representation

| | Entry | Average |
|-----|------------------|---------|
| LF | b_1 | 81.8 |
| | b_2 | 62.1 |
| | b_3 | 85.1 |
| HLF | $b_1 \times b_2$ | 2.12 |
| | $b_1 \times b_3$ | 1.98 |
| | $b_2 \times b_3$ | 1.86 |

The HLF location estimation step computes the nearest-neighbor with Euclidian distances in signal-strength ratio space. Euclidian distances are computed using the set of signal-strength ratios R calculated from the currently measured sample. The following formula is used with B_o as the set of base stations currently observed by the client:

$$E(c_x) = \sqrt{\sum_{b_i \times b_j \in B_o \times B_o, i < j} (R_{b_i \times b_j} - v_{c_x, b_i \times b_j})^2} \quad (4)$$

3.2 Bayesian Inference

Several LF systems use Bayesian inference [4, 10], which represents a *probabilistic* method. In simple terms, for each cell in the system a probability is calculated based on the currently measured sample. The probabilities are computed using Bayesian inference. The cell associated with the highest probability is picked to be the current location of the client. In Bayesian inference each fingerprint for each base station $b \in B$ is represented as a probability distribution over the range of absolute signal-strength values V .

To extend this technique to HLF both the fingerprint representation and the Bayesian inference calculation have to be changed. The HLF fingerprint representation is for each unique pair $b_i \times b_j \in B \times B$ a probability distribution over the range of signal-strength ratios $V' = [0 : nlr(v_{max})]$.

The probability distributions over V' are computed using the histogram method [10] from the fingerprints' samples. An example of a distribution is shown in Figure 3 for a specific fingerprint and a unique base station pair. A parameter that can be used to tune the histogram method is the size of the discrete steps; a size of 0.02 was used for the histogram on Figure 3 and for the evaluation in Section 4. This value was chosen by the authors based on evaluations that showed that larger values would deteriorate accuracy and smaller values would not improve it.

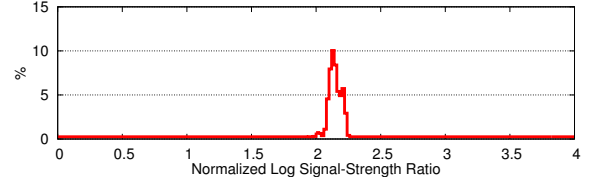


Figure 3. HLF Histogram

The HLF location estimation step performs Bayesian inference from signal-strength ratios computed from currently measured samples. The HLF fingerprint representation is used to describe the conditional probability of measuring a specific signal-strength ratio in a specific cell. The conditional probabilities over all cells are defined for a finite observation space $O' = \{o'_1, \dots, o'_m\}$ with each observation o'_i being a tuple with a unique pair of base stations $b_i \times b_j$ and a normalized log signal-strength ratio $v' \in V'$. The probabilities are calculated for an observation $o'_j \in O'$ within a cell $c_x \in C$ with fingerprint f_{c_x} as:

$$P(o'_j | c_x) = \text{Histogram}(o'_j, f_{c_x}) \quad (5)$$

where the function *Histogram* is the probability of the observation computed from the HLF-histogram fingerprint representation. The HLF location estimation step follows the LF procedure and returns the cell with the highest probability as the current cell of the client.

4 Evaluation

Our evaluation uses the traces collected as described in Section 2.1. In addition to traces a set of fingerprints have been collected for the test site's 126 cells one month before the traces. Each cell was fingerprinted by a person walking around in the cell for 60 seconds using a laptop with an Orinoco client. The evaluation uses this set of fingerprints for each technique's database of fingerprints. The evaluation is performed as emulated localization. This means that trace samples are given as input to a technique and the returned cell estimates are compared with trace ground truth. The evaluation results are given in terms of accuracy: the

percentage of samples where the ground truth and the estimated cell matched. Both the algorithms and the emulation environment were implemented by the authors in Java.

Our evaluation covers the techniques of *Nearest Neighbor (NN)* [1] and *Bayesian Inference (BI)* [4] implemented in three setups: a HLF version (implemented as presented in Section 3), a LF version, and a LF version extended with a manual solution for signal-strength differences. The manual solution handles signal-strength differences using linear mapping, as described in Kjærgaard [6]. The linear mapping transforms one client's samples to match another client's samples. The parameters for the linear mapping are found by comparing fingerprints collected with both clients using least squares estimation. The linear mapping is then applied to all samples before they are forwarded to a LF technique. The linear mapping parameters used in the evaluation were calculated from separate data collected with each of the clients.

Results of emulated localization with traces are given in Table 3 for each client and as an average over all clients. Accuracy for LF (first column) was highest for Orinoco (65% for BI) which can be attributed to the absence of signal-strength differences. However, for Intel and Apple BI accuracy is only 2% and 12%, respectively. The Fujitsu and D-link clients have higher accuracy and the NN accuracies are generally also a bit higher across all clients but for Intel only 10%. The results demonstrate that signal-strength differences have a large impact on LF accuracy for both NN and BI. Accuracy for LF extended with a manual solution (second column) is again highest for Orinoco. However, accuracy improves on average compared to LF for Apple, Fujitsu and Intel with 27% for BI and 22% for NN. For D-Link and Orinoco no improvement can be observed. One thing that can be noticed is that the BI accuracy for Apple and Intel do not improve as much as one could expect. This issue will be further analysed below. Accuracy with HLF (third column) improves on average compared to LF for Apple, Fujitsu and Intel with 22% for BI and 14% for NN. For D-Link there is a small improvement and no improvement for Orinoco. However, again it can be noticed that the BI accuracy for Apple and Intel do not improve as much as one could expect.

To give a more detailed analysis error distributions are shown in Figure 4 to 6. The error distributions for Apple and D-Link have been omitted because they are nearly similar to Intel and Orinoco, respectively. For Intel the distributions reveal a high percentage of large errors for LF, in comparison, both LF + Manual and HLF have much less large errors. The distributions also show that HLF for Intel recovers from the low accuracy in terms of percentage of large errors. For Fujitsu the better performance of LF is also apparent in lower errors which converge towards the distributions for LF + Manual and HLF. The lower accu-

racy of NN compared to BI is also visible as larger errors for NN than for BI. For Orinoco the distributions form a narrow band again with BI having the smallest percentage of large errors.

Table 3. % of correct estimations

| | LF | | LF + Manual | | HLF | |
|----------------|----|----|-------------|----|-----|----|
| | BI | NN | BI | NN | BI | NN |
| <i>Apple</i> | 12 | 31 | 28 | 42 | 32 | 30 |
| <i>D-Link</i> | 55 | 55 | 56 | 55 | 59 | 56 |
| <i>Fujitsu</i> | 23 | 32 | 51 | 45 | 48 | 41 |
| <i>Intel</i> | 2 | 10 | 39 | 53 | 25 | 45 |
| <i>Orinoco</i> | 65 | 58 | 65 | 59 | 65 | 57 |
| <i>All</i> | 31 | 37 | 48 | 51 | 46 | 46 |

Further analysis has shown that the smaller improvement for Apple and Intel can be attributed to a difference in the number of measured base stations at similar locations. Statistics calculated from the traces and fingerprints reveal that each D-Link and Fujitsu sample contains on average one extra observation than the Orinoco's samples. Apple and Intel samples contain on average approximately three extra base station observations. To address this problem we propose to use a K-strongest filter. The rationale behind this filter is that if a client makes more observations because of higher sensitivity we can filter out these by only keeping the K strongest measurements in each sample. K should here be set to match the sensitivity of the fingerprint client, from statistics calculated from the Orinoco fingerprints K was set to seven in our case. To evaluate this idea two emulations have been run for which results are given in Table 4 for BI. The first emulation applies a K-strongest filter to each sample before it is passed on to one of the techniques. The second emulation applies a ground-truth filter. This filter removes from each sample any extra observations that the Orinoco client did not observe at this location. For Apple and Intel the K-strongest filter has a large impact by improving BI accuracy with 15% and 20%, respectively, and reducing the percentage of large errors. The BI accuracy of the other clients is not improved by the K-strongest filter, which is consistent with the above calculations. The ground-truth filter improved BI accuracy for all clients except the Orinoco client. However, the ground-truth filter cannot be implemented in practice and are included to indicate an upper limit of performance for any filter. An interesting line of future work would be to develop a filter that using a prediction step could predict the base stations to sort out instead of only selecting the K strongest observations. Emulations were also run for LF where BI accuracy did not improve and LF + Manual where the filter made a small improvement in BI accuracy. For NN neither of the filters had a noticeable impact on accuracy.

For the preceding results a history of five samples were used. This means that, in addition to the current sample, the

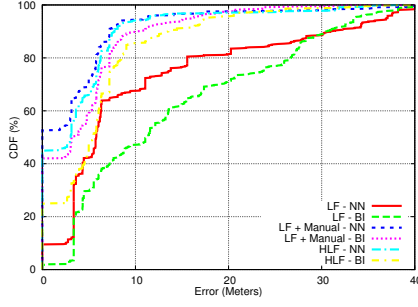


Figure 4. Error for Intel

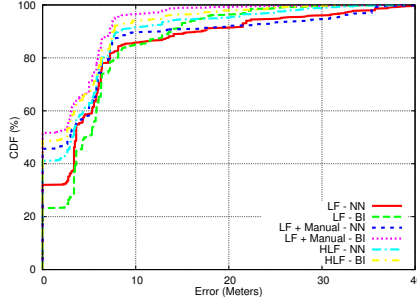


Figure 5. Error for Fujitsu

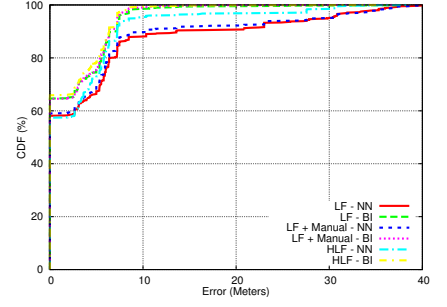


Figure 6. Error for Orinoco

Table 4. % of correct estimations for BI

| | HLF | HLF + K-Strongest | HLF + GT |
|----------------|-----|-------------------|----------|
| <i>Apple</i> | 32 | 47 | 72 |
| <i>D-Link</i> | 59 | 59 | 65 |
| <i>Fujitsu</i> | 48 | 48 | 53 |
| <i>Intel</i> | 25 | 45 | 73 |
| <i>Orinoco</i> | 65 | 64 | 65 |
| <i>All</i> | 46 | 52 | 66 |

four preceding samples are supplied with each trace sample to the techniques. The preceding samples are treated by the Bayesian inference techniques in the same manner as the current sample. For the nearest neighbor method, samples are aggregated to the mean value for each base station. Additional emulations have shown that consistently for both LF, LF + Manual and HLF a history of samples smaller than five make accuracy slowly drop and larger histories does not improve accuracy. For the preceding results the size of fingerprints have been 60 samples. Additional emulations have shown that consistently for both LF, LF + Manual and HLF a size of fingerprints below 20 samples make accuracy drop. The number of deployed base stations needed for techniques to work is an important number in practice. The preceding results were based on using data for all 26 base stations reachable in some parts of the two-floored 2256 m^2 test site. Additional emulations have shown that consistently for both LF, LF + Manual and HLF if we randomly remove base stations accuracy drops.

5 Discussion

The results of the evaluation were that the average accuracy for BI (with K-strongest filter) was 51% for LF + Manual and 52% for HLF and for NN it was 51% for LF + Manual and 47% for HLF. These results show that the accuracy of HLF and LF + Manual are nearly similar and improvements compared to LF. Distributions of errors also revealed that HLF and LF + Manual lower the percentage of large errors compared to LF. In this paper two HLF techniques were proposed and evaluated but the use of signal-strength

ratios are possible with other LF techniques. The results in this paper are based on data from five IEEE 802.11 clients, which are representative in terms of hardware and antenna options for many other clients. However, clients also exist that cannot be used for LF and also for HLF because of faulty or poor signal-strength measuring capabilities, for lists of such clients see Ekahau [3] and Kjærsgaard [6].

The evaluation also revealed that accuracy depends on clients making same-place measurements to the same set of base stations. Because the client used for fingerprinting collection in our data measured least base stations we cannot evaluate if this also is a problem if fingerprints are collected with a client that measure the most base stations. But it is an interesting line of future work to collect such data to see if a recommendation could be to always use a client that collect measurements to a maximum number of base stations for fingerprinting. From our analysis we can conclude that if the client is not maximal you have to filter the samples of other clients to maximize accuracy.

The evaluation of the common parameters showed that the HLF-extended techniques have the same sensitivity as LF techniques to the history of samples, the size of the fingerprints and the number of deployed base stations.

6 Related Work

One of the first IEEE 802.11 LF systems was RADAR [1], which applied different deterministic mathematical models to calculate a client's position (in coordinates). Similar methods have also been applied to GSM by Otsason et al. [9]. In comparison to RADAR, later systems have used probabilistic models instead of deterministic models, following the definitions in Kjærsgaard [7]. An example of a probabilistic system, which determine the coordinates of a client, is published by Youssef et al. [11]. A probabilistic system determining the logical position or cell of a client is published by Haeberlen et al. [4]. The basic LF systems do not address the issue of signal-strength differences.

Haeberlen et al. [4] propose using a linear mapping for transforming a client's samples to match another client's

samples. They propose three different methods for finding the two parameters in the linear mapping. The first method is a manual one, where a client has to be taken to a couple of known locations to collect fingerprints and parameters are found using least squares estimation. The second method is a quasi-automatic one, for which a client has to be taken to a couple of unknown locations to collect fingerprints. For finding the parameters, they propose using confidence values from Markov localization and find parameters that maximize this value. The third is an automatic one requiring no user intervention. Here they propose using an expectation-maximization algorithm combined with a window of recent measurements. For the manual method, they have published results which show a gain in accuracy for three clients; for the quasi-automatic method it is stated that the performance is comparable to that of the manual method, and for the automatic one it is stated that it does not work as well as the two other methods. In comparison, HLF has a performance comparable or better than the manual method and does not involve any extra steps of collecting additional fingerprints.

The method proposed by Kjærgaard [6] is also based on a linear mapping. This method is automatic, but it requires a learning period to find the parameters for the linear mapping. The solution is based on movement detection which is used to group same-place measurement into fingerprints. The parameters are then estimated from the grouped fingerprints using least squares estimation. The method, however, does only achieve lower or comparable performance to the manual approach, and it requires a learning period.

In addition to the above systems, which estimate the location of clients, a number of systems, such as NearMe [8], have been studied, for which the calibration step is only carried out by users for tagging relevant places. The systems propose simple metrics based on signal strength to quantify when clients are in proximity of calibrated places. One of the strengths of these simple metrics is that they overcome the problem of signal-strength differences. To summarize, HLF address signal-strength differences without requiring any extra steps.

7 Conclusion and Further Work

We showed that the proposed solution of HLF was able to address signal-strength differences. HLF records fingerprints as signal-strength ratios between pairs of base stations instead of as absolute signal-strength values. Signal-strength ratios factor out scaling differences in signal strength between clients. HLF is an improvement over existing solutions that require either error-prone manual steps or a learning period to work. Two LF techniques were extended to HLF and evaluated for five different IEEE 802.11 clients. The evaluation showed that the accuracy of HLF techniques is similar to that of existing manual solutions.

Two further issues subject to future work are proposed in the following. First, it would be interesting to evaluate other LF techniques with HLF and other technologies such as GSM where signal-strength differences are also present. Second, a further analysis is also interesting of how sensitivity affects the same-place measured base stations across clients. Here more data has to be collected to evaluate if a recommendation such as *always use a client which maximizes the number of measured base stations* can address the problem.

Acknowledgements

The research reported in this paper was partially funded by the ISIS Katrinebjerg competency centre.

References

- [1] P. Bahl and V. N. Padmanabhan. RADAR: An In-Building RF-based User Location and Tracking System. In *Proc. of the 19th Annual Joint Conf. of the IEEE Computer and Communications Societies*, 2000.
- [2] Y. Chan and K. Ho. A Simple and Efficient Estimator for Hyperbolic Location. *IEEE Trans. on Signal Processing*, 42(8):1905–1915, August 1994.
- [3] Ekahau. <http://www.ekahau.com/products/client/files/current/devices.html> (September 2007).
- [4] A. Haeberlen, E. Flannery, A. M. Ladd, A. Rudys, D. S. Wallach, and L. E. Kavraki. Practical Robust Localization over Large-Scale 802.11 Wireless Networks. In *Proc. of the Tenth ACM Int. Conf. on Mobile Computing and Networking*, 2004.
- [5] K. Kaemarungsi. Distribution of WLAN Received Signal Strength Indication for Indoor Location Determination. In *Proc. of the 1st Int. Symp. on Wireless Pervasive Computing*, 2006.
- [6] M. B. Kjærgaard. Automatic Mitigation of Sensor Variations for Signal Strength Based Location Systems. In *Proc. of the Second Int. Workshop on Location and Context Awareness*, 2006.
- [7] M. B. Kjærgaard. A Taxonomy for Radio Location Fingerprinting. In *Proc. of the Third Int. Symp. on Location and Context Awareness*, 2007.
- [8] J. Krumm and K. Hinckley. The NearMe Wireless Proximity Server. In *Proc. of the Sixth Int. Conf. on Ubiquitous Computing*, 2004.
- [9] V. Otsason, A. Varshavsky, A. L. Marca, and E. de Lara. Accurate GSM Indoor Localization. In *Proc. of the Seventh Int. Conf. on Ubiquitous Computing*, September 2005.
- [10] T. Roos, P. Myllymäki, H. Tirri, P. Misikangas, and J. Sievänen. A Probabilistic Approach to WLAN User Location Estimation. *Int. Journal of Wireless Information Networks*, 9(3):155–164, 2002.
- [11] M. Youssef and A. Agrawala. The Horus WLAN Location Determination System. In *Proc. of the Third Int. Conf. on Mobile Systems, Applications, and Services*, 2005.