

# Indoor radio location algorithm using empirical propagation models and probability distribution heuristics

Gábor Varga / Róbert Schulcz

Received 2011-10-28

## Abstract

Research is conducted at the Budapest University of Technology in the field of indoor positioning using radio waves, with practical application issues in focus. Our goal is to enhance and combine existing algorithms and create an implementation that is efficient enough to enable real-time operation in 3D space in multi-level office environments while retaining the accuracy of more complex systems. The proposed solution is based on proven empirical propagation models (the Motley–Keenan Model and the New Empirical Model of Cheung, Sau and Murch), with the ability to augment and refine with surveyed radio fingerprint data. The propagation model described in this paper can be used to generate estimated signal strength patterns for locations that were left out of the survey; and the values coming from the two different sources can be further handled in a common way. An algorithm is proposed to calculate a probability distribution over the floor plan using these expected values and the positioning-time measurements. The results from this location algorithm are combined with other probability distributions, generated by heuristics such as the distribution of walls, previous walking directions and typical indoor movement patterns, as well as data from G-sensors and digital compasses built into modern handheld devices.

Other real-world issues are investigated like reaction to changes in the placement of radio beacons, and the comparability of various devices with different signal characteristics. The proposed algorithm can achieve an average positioning error of 3 metres in a common office environment using fingerprinting data, and errors are only about 50% greater when relying solely on simulated signal strength values.

## Keywords

indoor radio location · indoor positioning · Wi-Fi positioning · location algorithm · empirical propagation model

Gábor Varga

Róbert Schulcz

Department of Telecomm., BME, H-1117 Budapest, Magyar tudósok krt 3., Hungary

## 1 Introduction

Development of context-sensitive applications has become a dynamically evolving area of information technology. Position as a context can on its own provide valuable information, but in conjunction with other knowledge bases, it makes services that grow to be integral parts of our daily lives possible. While outdoors the ability to determine geographic coordinates has given birth to a whole navigation industry, indoors this revolution is yet to be achieved. We, engineers, have to come up with solutions that integrate with our usual behaviour imperceptibly, but are able to provide such extra functions and ease of use that may change the way we look at our environment. Using indoor location, devices knowing the position of persons and objects surrounding them may very well realize such a vision.

Research projects and commercial products have been available for multiple years now that utilize the signals of Wi-Fi networks to calculate people's positions inside a building. However most of the currently used systems are based on the radio fingerprint principle that makes the thorough survey of the area necessary before deployment and after every change in the network infrastructure [5, 6, 15–17]. Realizing this burden, our goal was to develop an algorithm that lifts this constraint while retaining the accuracy and performance of the previous systems.

The motivation for developing a new location algorithm is threefold. Albeit being highly accurate and fast, methods utilizing radio fingerprints are strenuous to deploy and require full re-surveying every time an access point is replaced or moved [8]. For this reason they are not well-suited for an ever-changing office environment: a simple wave propagation model based algorithm would be preferable, where system parameters and access point locations can be changed independently. Secondly, the new algorithm has to be extremely resource-efficient in order to scale up to serving a multitude of people simultaneously in real time. It should run (possibly distributed) on central servers, so that regular equipment and clients with very low processing power could be used. Thirdly, its accuracy should not be much worse than the currently existing radio fingerprint technologies, with the option of finer calculation in specified locations. This is done by integrating the radio fingerprint based principle and

the propagation model, selectively relying on one or the other, having surveyed only the critical areas of the floor plan.

The selectively surveyed data set is overlaid on the map, and the missing areas are determined, where a mathematical model is used to calculate the expected signal strength values. Thus a distribution pattern is generated, much like a radio fingerprint map, so location-time operation is reduced to calculating probabilities of signal strength vectors matching. This way where accurate survey data exists, it will be used to determine the location, but we can still use another approach where it is not available. To increase accuracy, we statistically filter input; output is also filtered, its credibility is tested, and its oscillation is smoothed out. Different device models may be used simultaneously, since measured signal strength levels can be translated to a common scale [14].

The recalculation of the expected signal strengths is only necessary when the environment is changed. Calculated values are stored in a database as a materialized view, merged with the surveyed data set. In order to achieve fast computation of the propagation model based algorithm, we chose a rather simple equation based on the Motley–Keenan Model [11], augmented with factors from the New Empirical Model of Cheung, Sau and Murch [2]. On its own it still would not be sufficiently accurate, so we added some heuristics to the algorithm taking into account the floor plan, distribution of walls, previous walking directions and typical indoor movement patterns, as well as data from G-sensors and digital compasses built into modern handheld devices. Each of these information sources and the propagation model define a probability distribution over the map, these layers are then weighted and summarized to get the final coordinates. To avoid the need to run a lengthy simulated annealing in continuous space at every iteration, we decided to represent the map as a set of discrete points at the intersections of a grid over the map. This way calculations can be simplified, each probability layer can define the points to be evaluated on, and the introduced error is not greater than half of the grid diagonal, practically the space occupied by a person.

We used a novel approach when implementing the algorithm: we moved all the application logic into the database. Every calculation is represented as a PL/SQL language stored procedure in the Oracle Database, and both the algorithm's input and output are stored in database tables. We strongly rely on the Oracle Locator technology, which is a geographical extension of the database system: it allows us to do indexed 3D calculations and searches in a very convenient and elegant way.

This project started out as a university research at the Mobile Innovation Centre of the Budapest University of Technology. In cooperation with an industrial partner, it is the subject of a tender sponsored by the Economic Development Operational Programme of the Hungarian National Development Agency, partially financed by the European Regional Development Fund. The goal of the project is to develop a mobile based indoor location and navigation system usable in office environments, of-

fering a complete solution including the client, the server infrastructure, the operator interfaces and the location algorithm itself, with basic context-sensitive navigation services built on the top.

### 1.1 Positioning in radio networks

Before we analyse our indoor positioning system let's review the basic methods a positioning system can use. There are four main approaches associated with general location problems, even usable in radio networks such as GSM, UMTS or GPS, however these are not well suited for indoor environments and low-end consumer wireless devices.

The first is Time of Arrival (ToA). This method is based on the arriving time of the electromagnetic signals. The basic assumption of this is that we have positions with known coordinates. Points with equal distances from a known location form a sphere. In reality the reference positions are access points and the distance is measured by the time of the signal propagation. To use this approach a restriction is that the clocks of the mobile device and the reference station have to be accurately synchronized and have a resolution of at least 3.3 nanoseconds to achieve a positioning accuracy of 1 metre (this is calculated from the speed of light and electromagnetic waves).

Time Difference of Arrival (TDoA) is similar to ToA. In this case the system measures the difference of the running time of the signals received from already known transmitters. When given at least three known transmitters the unknown point can be determined as the intersection of the hyperbolas defined by the difference of the distances measured from two access points. This positioning principle abolishes the tight synchronization requirements of ToA, but it still requires clocks with extremely high resolution.

The requirement ensuring synchronization can not be satisfied easily, because the clock resolution of common wireless networking devices, as reported by their operating systems, is in the microsecond ranges, thus the maximum achievable accuracy is over 300 metres.

The other way to determine an unknown position is the Angle of Arrival (AoA) principle. Its precondition is only to be able to measure directions with respect to the reference coordinate system. The actual position can be calculated from the known coordinates of reference objects, using triangulation. We have to mention the Angle Difference of Arrival (ADoA), which is similar to AoA, but the orientation of the coordinate system is also unknown. What we are able to measure is the difference of the angles at known points' coordinates.

The main problem with measuring angles is that in wireless communication networks the antennas are non-directional, so we can not apply these theorems. To implement them, we would need complicated antenna systems, and the error would still increase with the distance. We have to mention the multipath problem too: electromagnetic signals propagating indoors are subject to reflections and refractions on the surfaces of objects in the environment, and these reflections cause the signals to ar-

rive at certain points from different angles.

Other physical parameters of electromagnetic waves need to be evaluated besides propagation time and direction. These parameters may include the phase, signal to noise ratio, spectrum, or most importantly the observed signal strength at the receiver end.

The propagation time and phase can not really be used reliably as the devices are not able to measure time as accurately as it is needed there. The SNR can be easily measured and calculated but we are not able to predict the amount of noise at a given point in time since we have no knowledge of the people's movement and the usage of interfering electronic appliances. The spectrum distribution of narrowband Wi-Fi channels does not seem to be measurably varying over distance, and its information may be distorted by the attenuation of walls. Signal strength can be measured and evaluated easily so it's promising to analyse it as features.

An increase of distance can be observed as a signal strength decrease (or the increase of the absolute value of the received power level expressed in dBm). By conducting measurements we can interpolate functions that assign distance values to various statistical properties of the received signal strengths. We have shown that because a great variance is introduced due to unpredictable elements in the environment, using the average (or rather the median) of multiple measurements yields better results than applying the function to every single value. Even the standard deviation shows a weak correlation with the distance: interestingly it decreases as the distance grows.

Signal attenuation based solutions measure the signal strengths at the receiver, and use this information to estimate the distance of the access points. There are three groups of models based on this approach:

- deterministic,
- semi-deterministic and
- heuristic models.

Deterministic ones are strictly based on mathematical, physical and electrical theorems of signal propagation and attenuation. The basis of these theorems are ray tracing, reflection and refraction. Adapting ray tracing is quite expensive with respect to calculations because of multiple reflections and refractions. Also the Maxwell-equations can be considered as a basic model, but it is also as expensive as the first mentioned solution. Neither of them can be considered in real-time systems because of the lack of resources.

Semi-deterministic models use simplified theorems modified with regards to a large number of empirical experiments and measurements. These models are based on scientific facts, augmented with empirical variables which are refined by tests comparing the predictions of the models with reality.

Heuristic models are solely formed by observation, they are not based on any valid engineering or physical models. Although these models are not the best at describing the causality

of facts and processes observed in reality, they estimate the positions quite well. The most common of these approaches are the ones using radio fingerprints, which are the unique signatures of signal strength vectors measured for the access points, peculiar to distinct locations on the map.

Fingerprinting approaches are based on experimental models that relate the measured signal strength values to the measurement position. These models are based on off-line collected data from several locations that cover the area to operate in. Fingerprinting algorithms are considered to be more robust against errors and extreme cases of signal propagation models as they make use of location dependent error characteristics. The way for the radio map's creation is called calibration, survey or training, the acquired data is called calibration or training data.

A great number of models have been developed so far to determine the positions of devices (and persons) in radio networks in all of these categories. The next section will introduce some of these ideas that are used in existing indoor positioning systems.

## 1.2 Existing solutions

After analysing basic positioning methods, we present some existing solutions which are in use nowadays. We show some familiar solutions and some interesting ones. There are three main directions in the techniques of indoor positioning.

The first, called fingerprinting has already been mentioned in the previous section. It uses signal strength patterns. The algorithm compares vectors of signal strength levels of the seen access points, the comparison is usually based on Euclidean distance.

There is an approach which uses probabilistic models (e.g. Bayesian model) to determine the location. It is based on prior probabilities which we already know, a posterior probability which we seek, and assumes a hypothesis. The Bayesian model for example is as follows:

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)}$$

where  $p(A)$ ,  $p(B)$  are the prior probabilities or marginal probabilities of events  $A$  and  $B$ ;  $p(A|B)$  is the conditional probability of  $A$  given  $B$  (also called posterior probability) and  $p(B|A)$  is the conditional probability of  $B$  given  $A$  (also called the likelihood).

This solution is not the most accurate one but in practice it gives quite accurate results.

The third, already mentioned common technique is based on wave propagation, with simplified equations describing the attenuation of the signal strength, possibly taking walls and other properties of the environment into account.

Next we introduce some of these solutions in real functioning products.

EkaHau [4] was created to supplement GPS indoors in 2002, since GPS signals are not visible inside buildings. The system is a software-based solution, so it can be used with existing wireless systems, of course by optimising the deployment of

the access points. The system offers sub room-, room-, floor-, and building-level accuracy. An average error of 1.5 metres is achieved by having three Aps in a 640 square metres site, 4 metres distance between sample points for measuring fingerprint data and 1.5 seconds scan interval. It can reach 1 metre accuracy by using 7 APs.

The patented algorithm first needs a floor plan and legal paths (rail tracking) then some calibration data measured continuously sampled between 3-5 metres (model calibration). These data contain user-clicked map coordinates, received signal strength intensity (RRSI) samples and other network information. During the positioning phase the positioning engine queries signal strength information from clients and estimates locations using a mathematical algorithm. This algorithm is based on a Bayesian probabilistic model, and assumes that the world is stochastic, and signals are inherently noisy, so the system is able to cope with the impreciseness of the measured signal levels.

In Ekahau Positioning Engine 2.0 [3] a hidden Markov model was added, which is based on the calibration data and is used to calculate transition probabilities from one location to another; and to decrease random errors a Kalman filter is used at the output. This solution has great importance because it was the first enterprise level indoor positioning system used in real operation.

Locadio [7] also uses a probabilistic model based on Markov chains. The solution was developed by engineers working at Microsoft, published in 2004, and never turned into a commercial product. The authors studied the differences of signal properties of moving and still wireless devices. They found that these states can be distinguished by their characteristics, and created a two-state Markov model (a simple Bayesian probabilistic filter) to decide whether the client is moving or not: if the Wi-Fi receiver is moving then the variance of the signal strength value is significantly higher than in the still state, so a signal level changing with greater amplitude means that the user is moving, otherwise he is standing still.

$$p(\text{moving}|\sigma_{\max}^2) = \frac{p(\sigma_{\max}^2|\text{moving})p(\text{moving})}{p(\sigma_{\max}^2|\text{moving})p(\text{moving}) + p(\sigma_{\max}^2|\text{still})p(\text{still})}$$

where  $\sigma_{\max}^2$  is the windowed variance of the currently strongest AP;  $p(\text{moving}|\sigma_{\max}^2)$  is the probability of moving if  $\sigma_{\max}^2$  is given;  $p(\text{moving})$  is the probability of moving;  $p(\sigma_{\max}^2|\text{still})$  is the probability of still if  $\sigma_{\max}^2$  is given and  $p(\text{still})$  is the probability of being still.

This information, together with the data of a rough fingerprinting survey, is used as the input of another Markov chain, representing actual positions and transition probabilities to estimate the location. The algorithm uses the Viterbi algorithm to compute the a posteriori probabilities. The mean error of the system is 1.53 metres.

The Cisco Wireless Location Application is exclusively based on the wireless network called Cisco Unified Wireless Network. In this approach the network needs to be more intelligent than

the devices, using the information collected and ensuring a comprehensive control of the whole system. Basically the application is for assuring contextual services in real time.

The architecture of the system is as follows: the primary components are specialised access points, WLAN controllers, and the Cisco Wireless Control System. The access points measure and report wireless networking properties of the consumers' devices, the WLAN controllers aggregate this information from the APs. A special challenge is the Cisco Wireless Control System's role: it has to maintain an overall picture of the network state and use this to locate individual devices.

The system's main goal is to provide security, IT management and asset tracking. This solution is special and interesting because of the usage of intelligent, vendor specific hardware to move the required software to the network infrastructure from the client devices.

WLANPos [12] was developed at the Budapest University of Technology and Economics. The authors developed two separate algorithms: one which is the improvement of the nearest neighbour in signal space (NNSS) search implementation, and another model which is probabilistic.

The NNSS is less accurate; it was just a starting point to be able to refine the positioning method. The algorithm compares stored signal strength vectors to measured ones and estimates an approximate location based on the best matches. Similar solutions use the Euclidean distance as a measurement of similarity, but because of the differences of various network adapters and antenna configurations, signal strengths measured by different devices do not have the same characteristics. The authors prevent this problem by comparing the signal strength difference of two neighbouring APs, thus eliminating the relative fault of the value levels.

$$d_n = \left( \frac{P_n}{P_{n+1}} - \frac{\hat{P}_n}{\hat{P}_{n+1}} \right) \cdot \frac{\hat{P}_n}{\hat{P}_{n+1}}$$

$$d = \sqrt{\sum_{n=0}^{N_{AP}} d_n^2}$$

where  $d_n$  is the  $n$ -th component of the distance;  $d$  is the total distance;  $P_n$  is the measured signal strength,  $\hat{P}_n$  is the stored signal strength of the  $n$ -th AP at a given point; and  $N_{AP}$  is the number of APs with signal strengths greater than a limit value, but at least three.

The probabilistic approach uses presumptions based on the Bayesian thesis by calculating probabilities over the position, AP existence and the strength of the AP. Instead of defining a single value describing the signal characteristics at a given location, a histogram is recorded. These histograms are then compared to calculate the probabilities of the client device being at any of the surveyed locations.

$$p(X_i|I, S) = \frac{p(I, S | X_i) \cdot p(X_i)}{\sum_{i=1}^n [p(I, S | X_i) \cdot p(X_i)]}$$

where  $p(X_i)$  is the probability of being at position  $X_i$ ;  $I$  is the

name of a list consisting of true or false values, which gives every AP whether it was seen (measurable) or not;  $S$  is the signal strength list of a given AP;  $p(X_i|I, S)$  is the probability that  $I$  is true and we accurately measured  $S$  at  $X_i$ ;  $p(I, S |X_i)$  is the probability that  $I$  is true and  $S$  was measured at point  $X_i$ .

The authors claimed in [12] that the algorithms' average errors were 3.53 and 2.27 metres respectively, but unfortunately no further information is given in their article regarding either the environment in which these measurements were carried out, or the standard deviation of the errors.

RedPin [1] is an open source indoor positioning system. The system is not only special because of this fact but because its users can collaborate by stating their positions and submitting the networking data recorded from the environment, therefore it shows some similarities with the collaborative efforts of Wikipedia. The goal of the authors was to reach at least room-level accuracy, therefore they did not really deal much with developing the most exact positioning method. The system does not provide geographic coordinates but rather symbolic identifiers such as names and numbers of rooms. It enables users to put labels on rooms instead of calibrating the whole territory.

The algorithm's first step is called sniffing during which the mobile phone gets signal strength information from active GSM cells, Wi-Fi APs and the IDs of seen non-portable Bluetooth devices. This system, using a multitude of information sources, can adapt better to environmental changes, for example AP replacements.

## 2 Developing a new algorithm

Having seen the advantages and disadvantages of the basic methods described above we decided to integrate a signal attenuation based solution with a fingerprint based method, using probabilistic calculations. Most of the currently used systems are based on the radio fingerprinting principle that makes the thorough survey of the area necessary before deployment and after every change in the network infrastructure. Realizing this burden, our goal was to develop an algorithm that lifts this constraint while retaining the accuracy and performance of the previous systems.

Our goal with the new algorithm is to develop a positioning system which accuracy is not worse than the earlier indoor positioning systems – which are mainly based on the fingerprinting method –, easy to install and able to be used real time.

What is the main problem with fingerprint based solutions? We have to calibrate the whole area after deployment and after each and every change in the network infrastructure, either when the topology changes or devices are simply replaced due to ordinary maintenance. This causes unnecessary and partly avoidable extra work for the maintenance staff of the building.

Of course we have to consider the specialities of the building. It would be reasonable if we could do this only by having a map. The algorithm has to be able to cope with multi-level buildings: it needs to decide on which floor we are and further-

more to determine the position on that level. The mean error – the difference of the estimated position and the real position – must not be more than 5 metres: we demand at least sub-room accuracy in a Wi-Fi covered building. It has to be able to track floor transitions made either by stairs or elevators.

Because our main goal is to develop a universal software-based solution for indoor positioning we can only use the information that we can obtain from the operating system (and from its driver programs). The only common reachable network parameter is the list of the visible APs and the signal strength values associated with them.

The system's deployment time and the amount of work to be done at network infrastructure changes can be decreased if we choose a wave propagation based model instead of fingerprinting ones [13]. Since at extremes (the noisy parts of the building) these can be quite imprecise we have to enable the measurement of signal characteristics at critical points. Thus we have to integrate the two approaches.

The algorithm based on wave propagation has to work in real-time with low delay. So instead of a deterministic model with high computing demand we use a narrowband empirical model which gives the correspondence between the distance and the measured signal strength as a simple mathematical equation.

The topology of the building has to be usable by the algorithm, so that the floor maps have to be in an interpretable format.

A design decision was made to assume a common office environment, thus restricting and building upon the movement patterns of people using the system, using this information to predict possible future locations and filter erroneous output.

We carried out our own experiments to examine the correspondence between signal strength levels and distances. The results showed that fluctuation and variance of the values is significant, however the trend of the function is clearly visible. That is why we decided to deal mainly with this approach.

We also use fingerprints to determine the location at surveyed regions of the map, and a probabilistic approach to resolve the multi-device problem in order to avoid the need to measure calibration data for each and every access point and mobile device pair.

Indoor radio positioning systems may use two different approaches: radio fingerprint based methods or wave propagation models.

Radio fingerprints are the unique distributions of measured signal strengths identifying distinct locations on the floor plan. Basic fingerprint positioning only requires the survey of the area: the signal strengths of the available access points are recorded in advance at different locations, possibly determined by a grid laid over the map. When positioning of devices is requested, the current distribution of the signal strengths, as measured by the device to be located is compared to the database of previous measurements, and the most probable location is calculated from these informations.

The main disadvantage of this approach is that it necessitates re-surveying of the entire area when the network topology changes due to simple maintenance like the movement or replacement of an access point. However this method gives the most accurate results, as it operates on actual data on both sides of the comparison.

Since signal levels have a significant variance due to multiple path propagation, interference and the effect of unpredictable factors in the environment (movement of people, other signal sources), both the beforehand survey and the real-time positioning should include the aggregation of the measured values by some kind of statistical function to filter out these errors. We have found that calculating the median value of at least 25 (preferably more) measurements in the surveying phase, and 5 measurements (or all measurements in the last few seconds) is the best approach to eliminate the random deviations.

In the signal strength vectors, where no data is available (because an access point cannot be seen from the location in question), we substituted the sensitivity limit of the client devices,  $-99$  dBm, and used this value in the aggregation.

It is assumed that the distribution of the measured signal strengths is the normal distribution, when the values are measured in Watts. Then the probability of the device being at a specific location is calculated from the cumulative distribution function of the assumed normal distribution of the signal strengths. As the signal strength is a continuous quantity, and we can measure it with a resolution of 1 dBm, we have to calculate the difference of two continuous cumulative distribution functions, measured in Watts at  $\bar{P}_i \pm 0.5$  dBm.

$$p(\vec{x}|\bar{P}_i) = \frac{\text{normcdf}\left(\frac{\bar{P}_i \cdot 10^{0.05} - \hat{P}(\vec{x})_i}{\sigma_i}\right) - \text{normcdf}\left(\frac{\bar{P}_i \cdot 10^{-0.05} - \hat{P}(\vec{x})_i}{\sigma_i}\right)}{\text{normcdf}\left(\frac{\bar{P}_i \cdot 10^{0.05} - \hat{P}(\vec{x})_i}{\sigma_i}\right) - \text{normcdf}\left(\frac{\bar{P}_i \cdot 10^{-0.05} - \hat{P}(\vec{x})_i}{\sigma_i}\right)}$$

where  $\bar{P}_i$  is the aggregated signal strength value for the  $i$ -th access point;  $\hat{P}(\vec{x})_i$  is the calibration data for the same access point at location  $\vec{x}$ ; and  $\sigma_i$  is the standard deviation of the survey data. The cumulative distribution function of the standard normal distribution cannot be calculated analytically, as it references the Gauss error function whose defining integral cannot be evaluated in closed form in terms of elementary functions, so we use Marsaglia's estimation algorithm [10] to compute the values using Taylor series.

Since the access points are independent of each other, the resulting probability is defined as the product of the probabilities assigned to the access points.

$$p(\vec{x}|\bar{P}) = \prod_{i=1}^{N_{AP}} p(\vec{x}|\bar{P}_i)$$

It is not yet clear if multiplication of the probabilities of all the access points, or only the ones with the highest values provide greater accuracy. It is similar to the problem of the nearest

neighbour search in vector space, where accuracy and performance both depend on the size of the signal strength vectors, and surprisingly the best option is to select only around 5 signal sources.

Because of the limited number of measurement data, population standard deviation cannot be calculated reliably. In extreme cases, where all the measured values are the same (which may very well happen in case of short duration measurements at survey points), the variance would be zero, yielding problems in the probability calculations. Instead of this we found that the standard deviation can be approximated by a polynomial, dependent on the distance from the access point. The constants are measured in Watts, and are specific to the environment.

$$\sigma(d) = 4.1270 \cdot 10^{-7} \cdot d^{-3.5119}$$

One could also compare entire histograms instead of only median values, however it would require at least an order of magnitude higher number of measurements for each location in advance, and preferably at positioning time too.

The result of the positioning can be obtained by selecting the location with the highest probability. A simple level of filtering may be applied by selecting more than just the first location from the ordered list, and calculating their centre of gravity (weighted average of the coordinates).

Another issue arises from the difference of the sensitivities and antenna characteristics of various device models. We found that the signal strengths reported by two different models for the same access point are linearly dependent. These values are now measured in dBm, in contrast to the previous equations using Watts.

$$\bar{P}_0 = a \cdot \bar{P}_1 + b$$

When a device of unknown type is added to the system, an iterative calibration of the linear parameters is started, and continued until the variance of the result of the linear equation stays below a predefined threshold of 3 dBm. A calibration step consists of two phases:

- The current location is estimated from the measured signal strength values already compensated by the previous value of the linear parameters, calculating the Pearson product-moment correlation coefficient of the signal strength vectors for each location and the measured vector (in effect this is very similar to the calculation of the probability distribution)
- And then using the signal strengths recorded for that location as reference points in a linear regression.

$$\begin{aligned} \vec{x} &= \min_{\vec{x}} \frac{\text{cov}(\bar{P}, \hat{P}(\vec{x}))}{\sigma(\bar{P}) \cdot \sigma(\hat{P}(\vec{x}))} \\ a &= \frac{\sum_{i=1}^{N_{AP}} (\hat{P}(\vec{x})_i - \bar{P}_i) - n \cdot \sum_{i=1}^{N_{AP}} \hat{P}(\vec{x})_i \cdot \sum_{i=1}^{N_{AP}} \bar{P}_i}{\sum_{i=1}^{N_{AP}} (\hat{P}(\vec{x})_i)^2 - n \cdot \sum_{i=1}^{N_{AP}} (\bar{P}_i)^2} \\ b &= \sum_{i=1}^{N_{AP}} \hat{P}(\vec{x})_i - a \cdot \sum_{i=1}^{N_{AP}} \bar{P}_i \end{aligned}$$

Electromagnetic wave propagation models may relieve us of the burden of constant re-surveying, enabling us to calculate the positions from the known behaviour of access points and client devices, as well as the building topology and material properties [9]. The two main classes of wave propagation models are deterministic and empirical models. Deterministic models need a detailed representation of the 3D space and the objects in it, and require very high computational power for tracing the paths of electromagnetic waves either by ray-tracing or dominant path tracing, so they are impractical for real-time and scalable location purposes. However they are very accurate, as they take multi-path propagation, reflection, wave-guiding, diffraction and shadowing into account.

Empirical models on the other hand provide approximate but acceptable results using only a simplified subset of the electromagnetic phenomena, chosen and tuned differently in each model. We chose to implement features from the Motley-Keenan Model, the COST 231 Multi-Wall Model and the Cheung-Sau-Murch Model.

To combine this method with the radio fingerprint technique, thus allowing higher accuracy where it is needed either by user demand or the failure of algorithmic computation at extreme locations, we set up a grid over the map (allowing multiple floors), assign survey data to the nearest grid points, interpolate values where possible (at small enough gaps), and generate fingerprint data by an algorithm implementing the wave propagation model. This way where accurate survey data are available, it will be used, and where it is missing, positioning will still be possible. After the radio fingerprint database is complete for all the grid points, the positioning algorithm is yet again reduced to the previous probability calculation, thus resulting in a fast solution integrating the best parts from both the fingerprint based method and the empirical wave propagation model.

The following data items are required for the propagation model:

- Access point characteristics: signal levels observable by a reference device at a distance of about 1 metre, thus eliminating near-field antenna effects and differences between transmitters
- Floor plans and floor heights, allowing the application to construct a rough 3D model of the building
- Placement of access points
- Wall and slab attenuations, as functions of the material and thickness
- An empirical constant as an attenuation factor for free space path loss, with a value of  $-2.5$  (we considered two distinct constants for the first Fresnel zone and the rest of the propagation path, but in an average indoor environment the first Fresnel zone covers almost all of the visibility area of the access point)

The predicted signal level of an access point at a given distance is defined as the sum of multiple attenuation terms, measured in dBm and dB:

- The access point's signal strength perceived at 1 metre
- Free space path loss for the given distance, if it is over the distance where previous value was measured
- Cumulated attenuation of intersected walls and slabs, inversely proportional to the cosine of the angle between the wave's path and the surface normal vector

$$P = P_0 + 10 \cdot n \cdot \log_{10} \left( \frac{|\vec{d}|}{d_0} \right) + \sum_{w \in W} \frac{L_w}{\cos \theta_w}$$

The resultant signal levels are overly pessimistic, mainly in areas where the density of walls is high. We are currently investigating the possibility of decreasing the attenuation factors of intersected walls and slabs as their number rises, similarly to the COST 231 model.

The location selected by this algorithm is independent of previous measurements and does not use any extra information available about the device's (or person's) movement. However previous results may be used to test the credibility of the calculated position. This filtering can be used to drop erroneous results altogether, or to limit the distance a new point can be from the previous ones, thus providing us with a new set of coordinates weighted towards the resultant position.

The new state may also be predicted by the current movement of the device. Modern handheld devices are equipped with accelerometers and magnetometers (G-sensors and digital compasses), which can be used to deduce some extra pieces of information, despite not being accurate enough for positioning on their own. The accelerometer can be used to sense if the user is walking or standing still, as there is a measurable difference between the variance of the sensor readouts. When the variance of the absolute value of the acceleration vector exceeds an empirically defined limit, the user is deemed to be moving. The magnetometer gives the direction of the movement, we only need to store the orientation of the map to use it. The speed of movement cannot be obtained this way, but assuming that the user is walking, it may be given as a random variable with normal probability density. In case these information sources are not available, the direction of movement may be guessed from the previous locations.

Considering the direction to be normally distributed too, another probability distribution  $p_m$  may be laid over the grid points. The probabilistic vector originates at the last known position, and points towards the direction of the movement. Its standard deviation is defined as an empirical (linear) function of its length.

The final probability distribution is calculated by multiplication of the probability distributions obtained from the fingerprint database and the movement informations.

$$p_r = p \cdot p_m$$

The resulting position may be selected as the location with highest value from this ordered list, or an aggregate as described before.

### 3 Evaluation

To measure the accuracy of our algorithm we made experiments. The test environment was our lab at the Budapest University of Technology and Economics in the Mobile Innovation Centre (building Z, level 3). We set up an environment to simulate an average office building. It has eleven small rooms, one bigger open space (the lab) and corridors connecting these rooms. We placed 8 access points uniformly in this 300 square metre area. We used TP-Link TL-WR941ND access points conforming to the 802.11g standard, operating at the 2.4 GHz frequency range on channel 6.

We decided to test the algorithm on one floor only, because this method would yield more comprehensible results. Anyway, the proper floor can be guessed quite easily by the algorithm, as slabs separating floors have typical attenuation values of  $-20$  dBm, thus resulting in very distinguishable signal strength values.

After placing the APs uniformly we first calibrated an HTC Touch HD device by measuring the whole site at every two metres and stored the observed signal strengths in the database. In the calibration phase 25 measurements were taken and sent to the server at each point. The client device measures the signal strength levels of the surrounding access points. Operating systems usually report signal levels with a resolution of 1 dBm. It takes about 1.5 seconds to measure a single set of values.

This step may be repeated more times to gather more survey data, as the algorithm is prepared to use any number of input vectors. When creating the fingerprint map on the server side, the median of these measurements is calculated for each access point, with the client's lower threshold (usually  $-99$  dBm) substituted for any unseen access point.

In the positioning phase the client continuously reports measured signal strength vectors to the server, these values are then stored in a database. The algorithm uses the median values of the input data aggregated over the last few seconds for each access point instead of only the latest measurement to filter out random errors and the effects of signal level fluctuations. This time interval is 10 seconds when the device is considered standing still (from the data of the accelerometer), and less when moving to avoid lags.

When evaluating the accuracy we conducted over 100 measurements at every chosen test location. We marked seven points uniformly on the map (three in small rooms, three in the lab and one in the corridors) and saved the positions estimated by the algorithm. We let the mobile device to take 100 measurements and observed the outcome of the algorithm on the server. We experienced that extreme results occurred but most of the time the algorithm was able to decide to the position claimed. These extreme cases should be able to detect and filtered out in later

steps of the program operation.

### 3.1 Results

We exported the result into a file and made statistics. First we calculated the average, the median and the standard deviation of the errors. Most of the errors were below 4 metres so the accuracy of the algorithm is in fact comparable to other solutions. The average error of the system was 3.09 metres. In the graph of Fig. 1 we can see the result of these individual measurements.

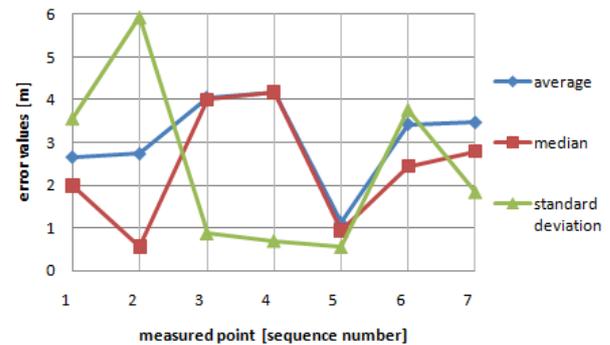


Fig. 1. Statistical properties of the error

The seven items at the horizontal axis of the graph are the locations where the measurements were made: (0.3; 8.3), (1.6; 23.5), (3.7; 4.2), (3.5; 12.0), (3.9; 15.8), (5.0; 24.5) and (6.2; 1.4).

The average error was between 1.12 and 4.09 metres for different locations. It can be evaluated as a good estimation but most of the time the mean value is above three metres. This error is acceptable for the whole system, if we expect further filtering of the results (and maybe fitting on navigation paths). The median value follows the trend of the average, but is lower due to its property of eliminating extreme errors. The overall median of the positioning error was 2.73 metres. The most unstable among the three statistical values is the standard deviation: it was 3.25 metres. As mentioned before some random errors were recorded which has a big effect this property. These extremes should be filtered from the output, as they are the result of unpredictable noise that we have no influence on. These can occur because of sudden and great electromagnetic noise or a human body held up the way of the signal. A possible solution to lower the effects of this phenomenon is to use a filtering method, for example the Kalman filter.

The actual distribution of the errors is given in Fig. 2:

The most common rounded error values are from 1 to 5. For example at point (3.9; 15.8) the error distribution is: 88% 1 metre, 7% 2 metres, 4% 3 metres and 1% 4 metres.

Fig. 3 plots the relative errors. The size of the circle depends on the frequency of the error pair, the origin stands for the actual position for every measurement. We can state that almost every error is in the five metre range of the origin and just a negligible proportion has greater value.

The two positioning results in the upper left corner of Fig. 3 are caused by unexpected factors. Due to the fact that our al-

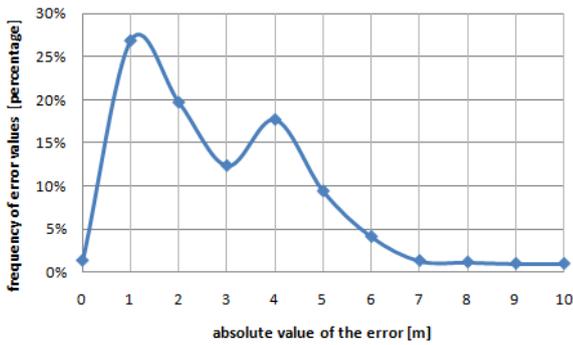


Fig. 2. Frequency of error absolute values

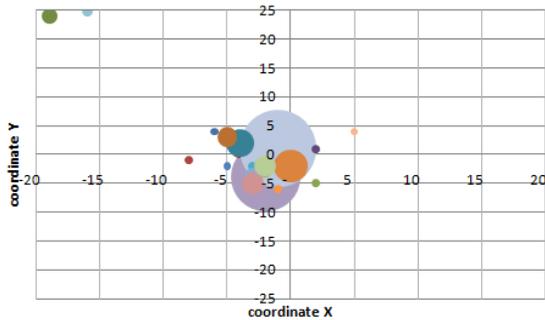


Fig. 3. Relative errors weighted by frequency

algorithm always chooses the best result, even if all possible positions have equally low (or zero) probabilities, such deviations may occur: in case of equal probabilities, the upper left corner of the floor plan was returned. These erroneous probability distributions may be caused by contradicting input vectors, possibly due to radio interferences, measurement errors or missing items from the signal strength vectors. This effect could be mitigated by more sensibly ordering locations with equal probabilities, or by testing the credibility of the positioning results.

We also tested the algorithm with fingerprinting completely turned off. We conducted the same measurements, and reached the conclusion that although it yields less accurate results, it is still usable: it had a mean error of 4.61 metres, with a median of 3.73 and a standard deviation of 4.47 metres. The distribution of the errors is given in Fig. 4. This means that it can be successfully used to abolish the need to survey each and every point of the grid laid over the floor plan.

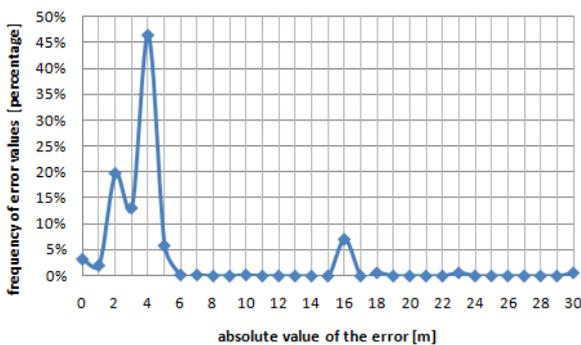


Fig. 4. Frequency of error absolute values with fingerprinting turned off

### 3.2 Implementation Details of the Test System

We implemented the algorithm inside the Oracle Database 10g Express Edition relational database management system. The rationale of this decision is threefold: it guarantees platform (operating system) independence, decreases the number of components to install when deploying the positioning system, and allows us to use a proven geometry storage and management module, integral part of the Oracle software suite. Oracle Locator, subset of the Oracle Spatial software option, is a software package supporting geographic information systems and location based services. Its primary focus is the representation and querying of street-level geographic data on the Earth's surface, but it is also suited for 3D Cartesian coordinate systems, thus it can be used to model office buildings and other indoor locations. The Express Edition may be used free of charge, however it imposes constraints on the computer resources: it limits the amount of memory in 1 GB, the number of processor cores in one 32-bit core, and the maximum storage space in 4 GB; these are enough for small-scale location purposes, but may require the purchase of a full-featured version for enterprise-level indoor location services.

The keystone of the Oracle Locator technology is the SDO\_GEOMETRY data type. It is an object capable of holding a number of geometric shapes: points, line segments, circular arcs, polygons, polygons with holes, and collections of the above. We chose to use a simple three dimensional Cartesian coordinate systems with no references to the Earth's surface (Spatial Reference System Identifier: 262152), with metres as measurement units; and to represent our objects either as points (geometry type: 3001), sections (geometry type: 3002) or polygons (geometry type: 3003). Locations of access points, measurements, and previous results for devices are represented as points, so are the grid points on which the location algorithm operates. Walls on the floor plan are represented as line segments, while the floors themselves are stored as polygons above each other. After maps are drawn on the user interface, a 3D model is created by converting walls to upright polygons between floors and ceilings, and these objects are stored together with the polygons for slabs (commonly referenced as obstacles) in a materialized view for further use by the location algorithm.

Locator supports the indexing of the above geometry objects. To accomplish this, the database administrator has to insert dimension and coordinate system information into the USER\_SDO\_GEOM\_METADATA view for each geometry column, and create indexes of type MDSYS.SPATIAL\_INDEX. These indexes are then used to speed up spatial queries. Quite complex topological relations and filters are supported by Oracle Locator, however only a subset of these is capable to operate on three dimensional data. The functions we found valuable for our purposes include SDO\_FILTER for finding objects with intersecting bounding boxes, SDO\_GEOM.RELATE to find objects with exact in-

tersections, SDO\_GEOM.SDO\_DISTANCE to get the distance between two objects, and SDO\_GEOM.WITHIN\_DISTANCE to query all the objects within a certain distance. We created custom functions to calculate the lengths of sections and the angles between them, as well as custom constructors, accessors and wrappers to hide the overly verbose syntax of the SDO package.

The Locator technology simplifies geometrical calculations, takes care of optimisation by indexing the data, and allows us to express the algorithm in a declarative language, SQL (a portion of the algorithm however is implemented in the procedural language PL/SQL). The same thing allows us to easily interface the geometry data with another part of our positioning application, which this algorithm is a part of: annotated locations (POIs, points of interest) and context-sensitive services. Two good examples for the benefits of this technology might be the cases of assigning measurements to the nearest grid points, or testing if the signal path between an access point and a grid point intersects any walls, and if so, returning the angles.

### 3.3 Conclusions

Our goal was to make a software-based algorithm with sub-room accuracy and real-time operation. By integrating fingerprinting and a wave propagation model based process we managed to handle special locations of the building where the signal can not be predicted properly by the model. The average error of the system is 3.09 metres with fingerprinting, 4.61 metres without it. These values are raw results, still missing any post-processing like validation and filtering, so this mixed algorithm will likely provide the required sub-room accuracy. Our results are promising, but they leave room for further refinery.

We intentionally omitted some heuristics that could have masked the real errors of the proposed algorithms. These heuristics include the addition of a probability layer with all ones near APs with significantly strong signals and all zeroes elsewhere (detecting that we are indeed very close to the AP), and the cut-off of areas that are unreachable from the previously known position in the elapsed time frame. However, in real-world test scenarios we experienced that these heuristics can greatly improve the positioning results. Another possibility that increases accuracy is the multiplication of access points: either by actually deploying more APs in a given area; or if an AP is set to broadcast multiple BSSIDs, more independent beacons will be received from it, allowing more accurate calculation of the average signal strength.

However, these results also indicate that mixing the two approaches of indoor radio-based positioning is feasible. We can survey the fingerprinting data where accuracy is important, and augment this data with simulated values generated by an empirical model where smaller accuracy is acceptable. This way the operational overhead of constant re-surveying can be lowered. In this paper we proposed a narrow-band empirical model that is easy to implement, can be executed efficiently, and, according to our measurements, is only 49% less accurate than the brute-force method based on radio fingerprinting.

## References

- 1 **Bollinger P.**, *Redpin – Adaptive, Zero-Configuration Indoor Localization through User Collaboration*, ETH Zürich, 2008.
- 2 **Cheung K W, Sau J H M, Murch R D.** *A New Empirical Model for Indoor Propagation Prediction*, IEEE Transactions on Vehicular Technology **47** (1998), no. 3, 996–1001, DOI 10.1109/ICCS.1994.474116.
- 3 *Ekahau Positioning Engine 2.0: 802.11-based Wireless LAN Positioning System*, Ekahau Corporation, 2002.
- 4 *Comparison of Wireless Indoor Positioning Technologies*, Ekahau Corporation, 2004.
- 5 **Hossain M, Nguyen Van H, Soh W.-S.**, *Fingerprint-based Location Estimation with Virtual Access Points*, IEEE International Conference on Computer Communications and Networks, posted on 2008, DOI 10.1109/ICCCN.2008.ECP.99, (to appear in print).
- 6 **Kaemarungsi K, Krishnamurthy P.** *Modeling of Indoor Positioning Systems Based on Location Fingerprinting*, INFOCOM Annual Joint Conference of the IEEE Computer and Communications Societies **2** (2004), 1012–1022, DOI 10.1109/INFCOM.2004.1356988.
- 7 **Krumm J, Horovitz E.** *Locadio: Inferring Motion and Location from Wi-Fi Signal Strengths*, Microsoft Research, 2004.
- 8 **Li B., Salter J, Dempster A G, Rizos C.**, *Indoor positioning techniques based on wireless LAN*, First IEEE International Conference on Wireless Broadband and Ultra Wideband Communications (2007), 3–16.
- 9 **Jaakko L.** *Indoor Propagation Models*, 2005.
- 10 **Marsaglia G.** *Evaluating the normal distribution*, Journal of Statistical Software **11** (2004), no. 4.
- 11 **Motley A J, Keenan J M.** *Radio coverage in buildings*, British Telecom Technology Journal **8** (1990), 19–24.
- 12 **Németh L H, Kis Z L, Szabó R.** *WLANpos: Wi-Fi Based Indoor Positioning (WLANpos: Wi-Fi alapú beltéri helymeghatározás)*, Híradástechnika **62** (2007), no. 8 (Hungarian).
- 13 **Sayrafian-Pour K, Perez J.** *Robust Indoor Positioning Based on Received Signal Strength*, IEEE Journal of Selected Topics in Signal Processing **3** (2009), no. 5, 821–831.
- 14 **Tsui A W, Chuang Y H, Chu H H.** *Unsupervised Learning for Solving RSS Hardware Variance Problem in WiFi Localization*, Mobile Network Applications **14** (2009), no. 5, 677–691, DOI 10.1007/s11036-008-0139-0.
- 15 **Varga G.** *Design and Analysis of an Indoor Radio Location Algorithm (Beltéri rádiós helymeghatározó algoritmus tervezése és vizsgálata)*, Budapest University of Technology and Economics, Department of Telecommunications, 2009 (Hungarian).
- 16 **Varga Zs.** *Wi-Fi Based Positioning in Mobile Environments (Wi-Fi alapú pozicionálás mobil környezetben)*, Budapest University of Technology and Economics, Department of Telecommunications and Media Informatics, 2006 (Hungarian).
- 17 **Webb N, Saragas V L D, Hung M.** *Indoor Navigation System for Handheld Devices*, Worcester Polytechnic Institute, 2009.