# Comparability of RF-based indoor localisation solutions in heterogeneous environments: an experimental study


## Tom Van Haute*, Eli De Poorter and Ingrid Moerman

Department of Information Technology (INTEC),
Ghent University – iMinds,
Gaston Crommenlaan 8, 9050 Ghent, Belgium
Fax: +32 9 331 49 46
Email: tom.vanhaute@intec.ugent.be
Email: eli.depoorter@intec.ugent.be
Email: ingrid.moerman@intec.ugent.be
*Corresponding author

## Filip Lemic, Vlado Handziski and Adam Wolisz

Telecommunication Networks Group (TKN),
Technische Universität Berlin (TUB),
10587 Berlin, Germany
Fax: +49 30 314 23818
Email: lemic@tkn.tu-berlin.de
Email: handziski@tkn.tu-berlin.de
Email: wolisz@tkn.tu-berlin.de

## Niklas Wirström and Thiemo Voigt

SICS Swedish ICT,
Stockholm, SE-164 29 Kista, Sweden
Fax: +46 8 633 15 38
Email: niwi@sics.se
Email: thiemo@sics.se


**Abstract:** The growing popularity of indoor localisation research has resulted in a significant amount of research papers describing and evaluating innovative localisation solutions. Unfortunately, the results from most of these research papers cannot easily be compared since they are evaluated in different environments, use different evaluation criteria and typically tailor their solutions towards a single testbed environment. To evaluate how these different conditions influence the localisation performance, in this paper an exhaustive set of experiments has been performed, in which three different localisation solutions have been evaluated using multiple metrics in three different test environments: two types of office environments and an industry-like factory environment. None of the used localisation solutions was previously optimised for any of these test environments and they were all evaluated under similar conditions. The results reveal several weaknesses in the evaluation methods used in the majority of existing scientific literature of indoor localisation solutions.

**Keywords:** indoor localisation; experimental comparison; benchmarking methodology; performance metrics; fingerprinting; time of arrival; RSSI-based localisation; IEEE 802.11; IEEE 802.15.4.

**Reference** to this paper should be made as follows: Van Haute, T., De Poorter, E., Moerman, I., Lemic, F., Handziski, V., Wolisz, A., Wirström, N. and Voigt, T. (2016) 'Comparability of RF-based indoor localisation solutions in heterogeneous environments: an experimental study', *Int. J. Ad Hoc and Ubiquitous Computing*, Vol. 23, Nos. 1/2, pp.92–114.

**Biographical notes:** Tom Van Haute is a Doctoral Researcher at Ghent University. He received his MSc (cum laude) in Computer Science Engineering from Ghent University, Belgium, in 2012. In September 2012, he joined the Department of Information Technology (INTEC) at Ghent University.

Within this department, he is working at the Internet Based Communication Networks and Services research group (IBCN). His research is focused on wireless sensor networks combined with indoor localisation and indoor navigation in particular. Additionally, he is providing guidance to Computer Science Engineering Master thesis and 'Informatics' course.

Eli De Poorter is a Professor at the Ghent University, Belgium. He received his Master and PhD degrees in Computer Science Engineering at the same university in 2006 and 2011, respectively. He is a Member of the Program Committee of several conferences and is the author or co-author of more than 70 papers published in international journals or in the proceedings of international conferences. He is the Creator of the Patented IDRA Architecture (http://idraproject.net), a flexible communication framework for heterogeneous networked devices. His main research interests include wireless network protocols and architectures, wireless sensor networks, self-learning networks, and next-generation network architectures.

Ingrid Moerman received her degree in Electrical Engineering (1987) and the PhD (1992) from the Ghent University, where she became a part-time Professor in 2000. She is a Staff Member of the Research Group on Internet-Based Communication Networks and Services, IBCN (www.ibcn.intec.ugent.be), where she is leading the research on Mobile and Wireless Communication Networks. Since 2006, she joined iMinds, where she is coordinating several interdisciplinary research projects. Her main research interests include sensor networks, cooperative and cognitive networks, wireless access, self-organising distributed networks (internet of things), and experimentally supported research.

Filip Lemic is a Junior Researcher and a PhD candidate at the Telecommunication Networks Group at the Technical University of Berlin. He finished his Bachelor and Master studies at the Faculty of Electrical Engineering and Computing at the University of Zagreb. His main scientific interests are in context awareness, with an emphasis on the indoor localisation.

Vlado Handziski is a Postdoctoral Researcher in the Telecommunication Networks Group at the Technische Universität Berlin, where he coordinates the activities in the areas of wireless sensor networks, cyber-physical systems and the internet-of-things. He received his Doctoral degree in Electrical Engineering from TU Berlin (summa cum laude, 2011). His research interests focus on software architecture aspects of networked embedded systems and their large-scale experimental evaluation. He has led research activities in several large European projects and participated in international standardisation efforts on testing indoor localisation systems. He is Chief Architect of the TWIST testbed at TU Berlin. He is the Member of IEEE and ACM.

Adam Wolisz received his degrees (Diploma 1972, PhD 1976, Habil. 1983) from Silesian University of Technology, Gliwice, Poland. He joined TU-Berlin in 1993, where he is a Chaired Professor in Telecommunication Networks and Executive Director of the Institute for Telecommunication Systems. He is also an Adjunct Professor at the Department of Electrical Engineering and Computer Science, University of California, Berkeley. His research interests are in architectures and protocols of communication networks. Recently he has been focusing mainly on wireless/mobile networking and sensor networks.

Niklas Wirström is a Researcher in the Networked Embedded Systems (NES) group at SICS and a PhD student at Uppsala Universitet, Sweden. His research focus is on Machine Learning techniques for localisation in WSN and other resource constrained systems.

Thiemo Voigt is a Professor at the Uppsala University. He also leads the Networked Embedded Systems group at SICS Swedish ICT. His main interests are networking and system issues in wireless sensor networks, and the internet of things. He has published papers at flagship sensor networking conferences such as ACM SenSys and IEEE/ACM IPSN, and received awards for several of these publications. He has also been TPC Co-Chair for IEEE/ACM IPSN and EWSN, the European Conference on Wireless Sensor Networks as well as Track Chair for IEEE DCOSS, IEEE MASS and IEEE ICDCS. He is on the Editorial Board for the *ACM Transaction on Sensor Networks* and the *IEEE IoT Newsletter*.

## 1 Introduction

### 1.1 Why indoor localisation

The emergence of satellite navigation systems – mainly GPS (Bulusu et al., 2000) – has resulted in a significant increase of personalised location-based services suitable for guidance, navigation, tracking, recreation, security, etc. However, the use of GPS is limited to outdoor environments, whereas many commercial applications are envisioned in indoor environments. Location-based services are envisioned in many different indoor environments: hospitals, airports, underground mines, detention houses, etc.

A significant amount of work is available in scientific literature describing and evaluating innovative techniques or solutions for localisation inside buildings. As a result a wide range of indoor localisation solutions has been proposed using a variety of different RF technologies (such as WiFi, RF, Bluetooth, 60 GHz, etc.) and non-RF technologies (such as infrared and ultrasonic). However, a major problem is the lack of comparability between indoor localisation solutions.

- The majority of evaluations of indoor localisation solutions (Hightower and Borriello, 2001; Pahlavan et al., 2002) focus mainly on the accuracy of the results whilst ignoring crucial application-level metrics such as scalability, delay, energy consumption, cost, simplicity, etc. Moreover, even the reported accuracy is typically calculated using different calculation statistics (average, median, percentiles, etc.), thereby making comparison of solutions is almost impossible.
- In addition, even though each of the targeted application domains has different environmental characteristics, most of the existing solutions were evaluated in one specific test environment. As a result, it is impossible to gain insight in the overall performance of these solutions under different conditions.

Based on these observations, we argue that the current state of the art is lacking comprehensive comparative analysis of different localisation approaches in multiple deployment enthronements. The main reason for this lack of comparability studies is the significant effort that is currently required to perform localisation experiments in multiple experimentation facilities. The main goal of this paper is to identify to what extent these shortcomings influence the comparability of results in existing scientific literature and to provide suggestions for improvement. Therefore, we implemented three typically used localisation approaches and evaluated their performance in multiple test environments using the same evaluation methodology. Amongst the evaluated solutions, we include:

- two popular RF technologies (IEEE 802.11 and IEEE 802.15.4)
- three localisation approaches (ToA, fingerprinting, weighted RSSI)
- four evaluation metrics (point accuracy, room accuracy, energy consumption, response time)
- three different test environments: two office environments and an industrial-like open environment.

To the best of our knowledge, we are the first to evaluate multiple localisation solutions (and not only different parameterisations of the same localisation solution class) in multiple environments using the same evaluation procedures.

The remainder of this paper is organised as follows. Section 2 discusses related work, including ongoing efforts to standardise the evaluation of indoor localisation solutions. Next, Section 3 describes the evaluated localisation solutions:

- a ToA-based IEEE 802.15.4 solution
- a fingerprinting-based IEEE 802.11 solution
- an IEEE 802.15.4 RSSI-based solution.

Section 4 discusses the used evaluation methodology and evaluation metrics. Section 5 gives an overview of characteristics of the used experimentation testbeds. Afterwards, the localisation solutions are evaluated and the performance results of the solutions are compared and discussed in Section 6 for office environment with brick walls (TWIST), in Section 7 for office environment with plywood walls (w-iLab.t I) and in Section 8 for open industrial like environment (w-iLab.t II). This is followed by a general overview in Section 9 where several lessons learned are discussed. Finally, Section 10 concludes the paper.

## 2 Related work

Recently, there has been a growing awareness that a more thorough way of comparing and evaluating localisation solutions is needed. This section gives an overview of efforts related to evaluation procedures for indoor localisation solutions.

### 2.1 Evaluation procedures for indoor localisation

The need for a systematic and objective evaluation methodology has been recognised by several authors (Liu et al., 2007). Although no standardised methodologies are currently available, several efforts are being made towards this goal.

- The FP7 EVARILOS project:[1] 'The EVARILOS project' (http://www.evarilos.eu) focuses on the *EVA*luation of *R*F-based *I*ndoor *LO*calisation *S*olutions. The project published a first draft of a benchmarking handbook (Van Haute et al., 2013a, 2013b), describing methods to calculate metrics, descriptive methods to describe evaluation environments and methods for deciding which evaluation points to use. The project is also the first to point out that current scientific literature lacks studies on the effect of interference on indoor localisation solutions, although interference is expected to be present at most sites where these systems are installed.[2]

- In parallel, the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) have established a joint technical committee, ISO/IEC JTC 1, to work on a ISO/IEC 18305 standard on "Test and evaluation of localisation and tracking systems" (http://www.iso.org/iso/home/store/catalogue_htc/catalogue_detail.htm?csnumber=62090).[3] The draft of the standard is not yet publicly available at the time of writing of this paper, but it currently includes a taxonomy of localisation solutions and describes a wide range of evaluation scenarios and performance metrics. In contrast to the EVARILOS project, which mainly focuses on RF-based localisation solutions, the ISO standard draft also considers other indoor localisation solutions that use a wide range of input sensors such as inertial sensors, ultra-sound sensors, etc. In terms of the evaluation approach, the ISO standard focuses on evaluation of fixed set of metrics and specifies a concrete enumerated set of evaluation scenarios under which the solutions should be evaluated. The EVARILOS benchmarking methodology, in comparison, is more broad and defines a basic 'vocabulary' for expressing different evaluation scenarios, instead of constraining to a set of few particular 'instances'. It also goes beyond evaluation of simple performance metrics, and defines a subsequent phase, in which they can be translated into use-case specific scores.

## 2.2 *Evaluation metrics for indoor localisation*

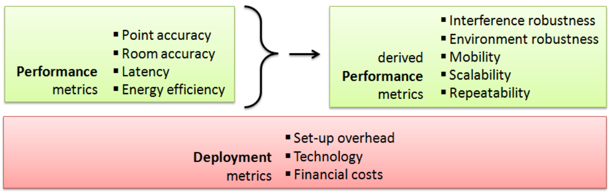
In more recent surveys, the importance of multiple metrics becomes visible.

- Liu et al. (2007) states that comprehensive performance comparison requires not only accuracy, but also needs to include precision, complexity, scalability, robustness and cost.
- In the EvAAL project ("Evaluating AAL Systems through Competitive Benchmarking") (http://evaal.aaloa.org/index.php?option=com_content&view=article&id=187:technical-annexes-localization2013&catid=15&Itemid=261), a competition is held that aims at establishing benchmarks and evaluation metrics for comparing Ambient Assisted Living solutions. For this competition, besides accuracy, also usability metrics are defined such as installation complexity, user acceptance, availability and interoperability with AAL systems (http://evaal.aaloa.org/index.php?option=com_content&view=article&id=187:technical-annexes-localization2013&catid=15&Itemid=261).
- A significant number of additional metrics can be found in the aforementioned EVARILOS handbook (Van Haute et al., 2013b) and ISO/IEC 18305 draft (http://www.iso.org/iso/home/store/catalogue_htc/catalogue_detail.htm?csnumber=62090), both including additional functional metrics, such as response delays, and non-functional (deployment) metrics such as setup time and required infrastructure.

The full list of potential metrics from these sources is very large, especially since many of these metrics can be calculated using multiple statistics (percentiles, averages, median, distributions, etc.). Some metrics are important mainly from a theoretical point of view and as such are well-suited for analysing and improving algorithms of researchers (Dezhong et al., 2014), whereas other focus on the performance of end-systems and as such are more important for the industry. Unfortunately, although the above sources strongly emphasise the need for utilising multiple criteria for evaluating indoor localisation solutions, none of these sources mention, which of the metrics are considered most important for different application domains, nor do they offer insight on the relation between different metrics (e.g., inherent trade-offs).

Therefore, in the evaluation section of this paper, we have included four functional metrics: point accuracy, room accuracy, response delay and energy consumption. These are the performance metrics from the EVARILOS handbook (Van Haute et al., 2013b), which can be found in Figure 1.

**Figure 1** The EVARILOS metrics: a graphical overview (see online version for colours)




## 2.3 *Evaluation environments for indoor localisation*

It is a well-known fact that environmental conditions significantly influence propagation characteristics. Table 1 gives an overview of a number of recent research papers evaluating localisation solutions and describes the environments they have been evaluated in.

It is clear from Table 1 that most existing indoor localisation solutions have been evaluated in office environments since these are the buildings, which are most readily available for researchers. Owing to the time-consuming nature of performing localisation experiments, most localisation solutions are evaluated only in a single environment. However, as will be shown in Section 5, office environments can have very different characteristics. Based on existing literature, it is not clear how these differences in environment influence the reported accuracy results. Therefore, this paper will analyse the performance of multiple localisation solutions in three different environments: an office environment with brick walls, an office environment with plywooden walls and an industrial-like open environment.

## 2.4 *Evaluation points for indoor localisation*

In terms of which points to use in an environment to evaluate the performance of a localisation solution, two

**Table 1** Overview of a few existing indoor localisation solutions with the related environment, testbed and metrics

| *Solution* | *Environment* | *Used testbed* | *Used metrics* |
|---|---|---|---|
| Energy efficient solution (Dezhong et al., 2014) | Office | building on campus | Point acc. & energy cons. |
| GSM fingerprinting (Otsason et al., 2005) | Office / home | university, research lab, house | Point accuracy |
| WiFi Bayesian (Ladd et al., 2004) | Office | Their own hallway | Point accuracy |
| EZ localisation (Chintalapudi et al., 2010) | Office | Office floor, Call Centre | Point accuracy |
| Smartphone localisation (Martin et al., 2010) | University | Berkeley campus | Point accuracy |
| WiFi in tunnel (Sunkyu et al., 2011) | Mining | Tunnel in Guangzhou MTR | Point accuracy |
| Fingerprinting (Stella et al., 2014) | University | Fourth floor of university building | Point accuracy |
| UWB fingerprinting (Steiner and Wittneben, 2011) | Office / testroom | Anechoic chamber, office floor | # of multipath components |

main approaches are possible. For industry-related testing, an evaluation track can be created that mimics typical operations in a building. For example, the path of a person can be recreated and only evaluation points on this path can be used (Dezhong et al., 2014). For more generic, application-independent testing, ideally the evaluation points should be randomly chosen. Unfortunately, most research papers manually select a number of evaluation points based on subjective criteria such as accessibility. As will be shown in Section 6, the accuracy of localisation solutions can strongly depend on the used evaluation points, e.g., points near a wall vs. open spaces. As a result, the performance of localisation solutions can artificially be 'improved' by selecting mostly evaluation points, which perform well for the evaluated solution.

As such, it is clear that future evaluations of indoor localisation solutions should use standardised evaluation methods. To remedy this, future benchmarking methodologies such as EVARILOS and ISO/IEC 18,305 are creating standardised methods for generating evaluation points. For this paper, all evaluated localisation solutions use the same evaluation points in each testbed.

## 3 Evaluated localisation solutions

To evaluate how different test environments influence typical localisation solutions, we selected three localisation solutions that use different wireless technologies and that use different processing approaches for estimating positions. The following localisation solutions were selected and implemented:

- an IEEE 802.15.4 based time-of-arrival solution
- an IEEE 802.11 based fingerprinting solution
- an IEEE 802.15.4 based RSSI triangulation solution.

Although more accurate solutions exist, these solutions represent the most popular RF-based technologies described in literature.

### *3.1 Particle filter using ToA and RSSI measurements*

The first solution is designed by Pettinato et al. (2012). The basic concept behind this localisation solution is the following: measurements are performed by letting a stationary node transmit packets to the anchors that reply with a hardware ACK (acknowledgement). The initiating node measures both the time between the transmission of the packet and the reception of the ACK, and stores the RSSI values associated with the ACK. These measurements are then processed using *Spray* (Wirström et al., 2014), a particle filter based platform.

The basic idea of the ToF ranging is to estimate the distance between two nodes by measuring the propagation time that is linearly correlated to the distance between the nodes when they are in LoS. Two-way ToF ranging, as opposed to one-way, does not require tight time synchronisation between sender and receiver. This is an advantage since tight time synchronisation is hard to achieve in wireless sensor networks (WSNs) (Elson and Römer, 2003).

The distance between nodes can be calculated according to equation (1) where $c$ is the speed of light, $t_{rmToF}$ is the round-trip-time measurements, and $t_{\text{off}}$ is an offset time accounting for all processing delays in the system. This includes the time for the sender to transmit the packet, the time the receiver needs to process it, and send the acknowledgement.

$$d = \frac{c}{2}(t_{\text{ToF}} - t_{\text{off}}) \tag{1}$$

The measurements $t_{ToF}$ are computed as $t_{\text{ToF}} = \frac{n_{\text{cycles}}}{f_{\text{timer}}}$, were $n_{\text{cycles}}$ is the number of measured clock ticks, and $f_{\text{timer}}$ is the frequency of the radio's internal crystal oscillator. In this case $f_{\text{timer}} = 12$ MHz. A single measurement is not sufficient, however. The resolution of a single clock allows for a spatial precision equal to $\Delta_d = \frac{c}{2f_{\text{timer}}}$. For a 12 MHz clock, the resulting spatial resolution is 12.5 m. To achieve higher resolution, one can average over a series of measurements, as proposed by Mazomenos et al. (Mazomenos et al., 2011). This way, sub-clock precision can be achieved.

#### *3.1.1 Range computation methods*

Once the range measurements are collected, they have to be transformed into actual distance measurements. For this, a wide range of computation methods are available. We have applied five different methods to the measurements. Four of these use ToF measurements as input, and one use RSSI measurements. The following subsections describe the methods.

*Mazo*: This model builds directly on equation (1). This is the model used by Mazomenos et al. (2011). The calibration step consists of estimating the constant offset $t_{\text{off}}$ by averaging over various ToF measurements according to equation (2).

$$\hat{t}_{\text{off}} = \frac{1}{N}\sum_{i=1}^{N} t_{\text{ToF},i} - \frac{2d_i}{c} \tag{2}$$

*k-sigma*: This method was proposed by Pettinato et al. (2012). It uses the variance between measurements taken on different channels to improve range estimations. The idea is that when two nodes are in line-of-sight, most packets will travel the shortest path between the nodes, regardless of the channel being used.

If the two nodes are not in the LoS, however, the different frequencies of the different channels will cause slightly different propagation paths, and result in different ToF measurement values. The concept is captured in equation (3), where $\sigma$ is the inter-channel standard deviation. Calibration consists of estimating $t_{\text{off}}$ and $k$ using linear regression.

$$d = \frac{t_{\text{ToF}}}{2} - t_{\text{off}} - k\sigma \tag{3}$$

*Least squares*: For this method, the calibration phase consists simply of fitting data to the equation (4), where $a$ and $b$ are estimated using linear regression. This method is model-free in the sense that it does not rely on a physical model.

$$d = a + bt_{\text{ToF}} \tag{4}$$

*Free space RSSI*: This method uses the free space propagation model in equation (5), to transform RSSI measurements to range estimations. In the equation, $P_r$ and $P_t$ are the received and transmitted power, respectively. $G_r$ and $G_t$ are the receivers and the transmitters antenna gains, respectively. $\lambda$ is the wavelength and $L$ is called the system loss factor.

$$P_r = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 d^2 L} \tag{5}$$

However, instead of determine these constants individually we combine them into on single constant $K$ as in equation (6), and estimate $K$ using least squares approximation.

$$P_r = K\frac{1}{d^2} \tag{6}$$

### *3.1.2 Using spray to estimate location*

Once the raw range measurements are transformed to distance estimations, the final location estimations are obtained from *Spray*, a particle filter based localisation system that can be used to fuse multiple types of measurements simultaneously.

In this case, i.e., using a single range based modality, *Spray* generates *particles* that have both a position and a weight, in a ring-shaped cloud (an annulus) around each testbed node that has an associated range measurement to the node that is to be localised. The distance between each particle and its associated testbed node, is determined by the sum of the range measurement and a zero-mean normally distributed random variable with a given variance.

Each particle is then evaluated using measurements from all the other testbed nodes, on the basis of how the particle's position fits their measurements. This is done by assigning a weight between 0 and 1 to the particle. The more coherent the particle's position is with the measurement, the higher the weight. A final weight for each particle is then computed by multiplying the weights assigned in the evaluation phase.

## *3.2 Fingerprinting based localisation using WiFi beacon packets RSSI measurements*

Another solution is provided by Lemic (2014) and Lemic et al. (2014d), which is based on WiFi fingerprinting. Fingerprinting methods in the indoor localisation are generally divided in two phases. The first phase is called the training or *offline* phase. In this phase, the localisation area is divided in a certain number of cells. Each cell is scanned a certain number of times for different signal properties, and using a methodology for processing the received data a fingerprint of each cell is created. By using the obtained training fingerprints the training database is created and stored on the localisation server. In the second phase, known as the runtime or *online* phase, a number of scans of the environment are created using the user's device. From the scanned data, using the same predefined data processing method, the runtime fingerprint is created and sent to the localisation server. At the server's side the runtime fingerprint is compared with the training dataset using the matching method. The training fingerprint with the most similarities to the runtime fingerprint is reported as the estimated position. In the section below a general notion of the WiFi fingerprinting is given using beacon packets RSSI values.

Let $K_t$ and $M$ be respectively the number of WiFi APs used for a localisation procedure and the number of training points in a given localisation area. Furthermore, let $N_t$ be the number of scans of the area taken at a training point $m$ ($m \in 1, ..., M$). During each scan the vector of RSSI measurements from each visible AP used for localisation is collected. This vector has at most $K_t$ elements, but it is possible that it will have less elements if the user's device is not in the range of a number of APs or because beacon packets are lost owing to interference. After collecting $N_t$ measurement vectors from different APs at training point $i$ the training matrix $\boldsymbol{S}_i^t$ is created. The matrix $\boldsymbol{S}_i^t$ has $K_t$ rows and $N_t$ columns ($\boldsymbol{S}_{K_t \times N_t}^t$). The matrix of the training measurements from each training cell is preprocessed training data. Based on the method that each localisation algorithm uses for creating the fingerprint, from the matrices $\boldsymbol{S}^t$ $M$ training fingerprints are created.

A similar procedure, with different parameters, is used for creating the runtime scan of the RSSI measurements. Let $K_r$ be the number of WiFi AP used in the localisation procedure and visible to the user's device at a given location. The number of measurements taken by the user's device is equal to $N_r$. A runtime fingerprint is a matrix of RSSI values $\boldsymbol{S}_{K_r \times N_r}^r$. A fingerprint is created using a method defined in the fingerprinting based localisation algorithm.

The principle of fingerprint based localisation algorithms is to accurately detect the similarities between training dataset and runtime fingerprints. Owing to the time and energy constrains of a (usually wireless) user's device, the number of measurements in the runtime fingerprint $N_r$ is usually smaller than the number of measurements taken while collecting training fingerprints $N_t$. For this reason, the number of measurements given as an input to a localisation algorithm is equal to $N_r$. Furthermore, only a subset of RSSI measurements from the APs that are common to both training and runtime fingerprint is given to the second phase of the localisation algorithm.

For the evaluation, we use three fingerprint based indoor localisation algorithms, which have been proposed in previous research work.

*ED of averaged RSSI vectors*: The Euclidean distance (ED) of the averaged RSSI vectors is one of the most basic and well known algorithms used for fingerprint-based indoor localisation algorithms (Milioris et al., 2001). The input to the matching method is an average value of RSSI measurements obtained from each AP used for localisation in both training and runtime phase, where $K_{r,t}$ is the length of the vector. Let $\boldsymbol{\mu}_{t,m} = [\overline{\mathrm{RSSI}}_{t,1}, ..., \overline{\mathrm{RSSI}}_{t,k}, ..., \overline{\mathrm{RSSI}}_{t,K_{r,t}}]$ be the vector of averaged RSSI values from each AP obtained during the training phase at cell $m \in 1, ..., M_t$, i.e., the training fingerprint. In the same manner, let $\boldsymbol{\mu}_r = [\overline{\mathrm{RSSI}}_{r,1}, ..., \overline{\mathrm{RSSI}}_{r,k}, ..., \overline{\mathrm{RSSI}}_{r,K_r}]$ be the vector of averaged RSSI values from each AP obtained during the runtime phase, i.e., the runtime fingerprint. The distance between the training fingerprint at the cell $m$ and the runtime fingerprint is given as:

$$D_E(\boldsymbol{\mu}_{t,m}, \boldsymbol{\mu}_r) = |\overline{\boldsymbol{\mu}}_{t,i} - \overline{\boldsymbol{\mu}}_{r,i}| \tag{7}$$

The distance $D_{EU}(\boldsymbol{\mu}_{t,m}, \boldsymbol{\mu}_r)$ is the ED distance between the vectors of averaged RSSI values of the cell $m$ and runtime point. The cell with the smallest distance (also called smallest weight) is reported as the estimated position.

*KL distance of MvG distributions of RSSIs*: The second fingerprinting based indoor localisation algorithm uses the Kullback-Leibler (KL) distance between the Multivariate Gaussian distributions of RSSI measurements from each AP used in the localisation procedure (Milioris et al., 2001). The algorithm assumes that the RSSI values from each AP are distributed according to the Multivariate Gaussian distribution. In other words, the distribution of the RSSI values from each AP at one cell can be written as $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. In the same manner as in the previously presented algorithm, let $\boldsymbol{\mu}_{t,m}$ and $\boldsymbol{\mu}_r$ be the vectors of the averaged RSSI values from each AP in training phase at the cell $m$ and in the running phase, respectively. Furthermore, let the $\boldsymbol{\Sigma}_{t,m}$ and $\boldsymbol{\Sigma}_r$ be the covariance matrices of the RSSI measurements at training cell $m$ and running point respectively. The Multivariate Gaussian distributions of the training point $m$ and running point can then be written as $\mathcal{N}_{t,m} = \mathcal{N}(\boldsymbol{\mu}_{t,m}, \boldsymbol{\Sigma}_{t,m})$ and $\mathcal{N}_r = \mathcal{N}(\boldsymbol{\mu}_r, \boldsymbol{\Sigma}_r)$ respectively.

$$\begin{aligned} D_{KL}(\mathcal{N}_{t,m}, \mathcal{N}_r) = \frac{1}{2}&((\boldsymbol{\mu}_{i,T}^S - \boldsymbol{\mu}_R^S)^T (\boldsymbol{\Sigma}_{i,T}^S)^{-1} \\ &\times(\boldsymbol{\mu}_{i,T}^S - \boldsymbol{\mu}_R^S) \\ &+ tr(\boldsymbol{\Sigma}_{i,T}^S (\boldsymbol{\Sigma}_{i,T}^S)^{-1} - \boldsymbol{I}) \\ &- \ln|\boldsymbol{\Sigma}_R^S (\boldsymbol{\Sigma}_{i,T}^S)^{-1}|) \end{aligned} \tag{8}$$

where $tr(\cdot)$ denotes the trace of a matrix (sum of its diagonal elements) and $I$ is the identity matrix. The matching method reports the cell with the smallest KL distance as the estimated position.

*PH distance of RSSI quantiles*: Finally, as the third fingerprinting method, we propose a new approach using quantiles of the RSSI values from each AP for creating fingerprints and the Pompeiu-Hausdorff (PH) distance for estimating the similarities between the training and runtime fingerprints. Using the quantiles for indoor localisation purposes is frequently used in robotics, where robots are using quantiles of images of the environments to localise themselves (Chambers et al., 2006). PH distance is usually used in image processing for pattern recognition and measuring the dissimilarities between shapes. As far as we know, using a combination of quantiles of RSSI distributions and PH distance for location estimation has not been proposed and examined in literature. We find this approach promising because a higher amount of information is provided to the matching method. In other words, in our opinion using only the vector of averaged RSSI values and the covariance between measurements between different APs may not be sufficient for precise localisation. In our case the $q$-quantile of the RSSI measurements from each AP is calculated in two steps. The first one computes the cumulative distribution functions (CDFs) of the RSSI measurements from each AP. The second step calculates the quantiles, i.e., the RSSI values with probabilities $k/(q-1)$, where $k = 0, 1, ..., q-1$. The result of the quantile calculation in both training and runtime phase is a quantile matrix $Q_{K,q}$, where $K$ is the number of APs visible at the given location and $q$ is a number of quantiles. The similarities between the RSSI quantiles from the training fingerprints and the runtime fingerprint are computed using the PH distance metric. The PH distance between two sets of quantiles is given as follows:

$$D_{PH}(Q_1, Q_2) = \max_{q_{1,k} \in Q_1} \big( \min_{q_{2,k} \in Q_2} (d(q_{1,k}, q_{2,k})) \big) \tag{9}$$

where $d(q_{1,k}, q_{2,k})$ is the Euclidean distance (ED) measurement. The training cell with the smallest PH distance is reported as an estimated location.

### 3.3 Hybrid model: proximity and weighted RSSI

A final localisation solution (Van Haute et al., 2014) that has been implemented and evaluated is a hybrid combination of a range-based and a range-free algorithm. It includes a range-based location estimator based on weighted RSSI values. The main idea of RSSI is that the transmission power $P_T$ directly affects the received power $P_R$ of a signal. Using the Friis transmission equation, the linear relationship can be stated as follows.

$$P_R = P_T \times G_T \times G_R \left(\frac{\lambda}{4\pi d}\right)^2 \tag{10}$$

In the equation $G_T$, $G_R$ are the gains of transmitter and receiver, respectively. $\lambda$ is the wavelength of the signal and $d$ is the distance between sender and receiver. The RSSI can be defined as the ratio of the received power to the reference power $P_{\mathrm{Ref}}$.

$$\mathrm{RSSI} = 10 \times \log \frac{P_R}{P_{\mathrm{Ref}}} \tag{11}$$

Each RSSI value can be matched with a certain distance. The proposed algorithm in Van Haute et al. (2014) not only uses the RSSI values to measure the distance between a fixed and mobile node, but also the distance between the fixed nodes. These values function as weight factors for the distance calculation between the fixed and mobile node. These weight factors are shown in Figure 2 as $w_{12}, w_{13}$ and $w_{23}$. The distance from $M$ to, for example, $B_1$ can be calculated as follows:

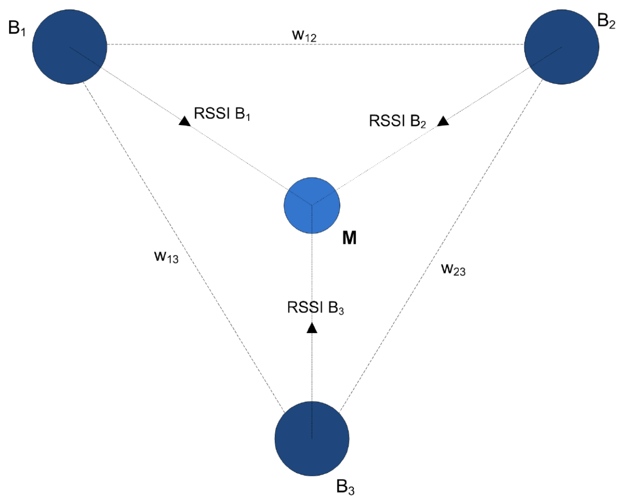
$$\text{Distance}(M, B_1) = \frac{\text{RSSI}(M, B_1) \times w_{12} + \text{RSSI}(M, B_1) \times w_{13}}{2} \tag{12}$$

whereby $w_{ij}$:

$$w_{ij} = \frac{\text{Dist}(B_i, B_j)}{\text{RSSI}(B_i, B_j)} \tag{13}$$

Previous results prove that these weight factors add value to the accuracy. A drawback of the RSSI technique is that these measurements are very sensitive to the environment and any changes in it. The relationship between the distance and RSSI is room dependent. For example, signals in a long corridor propagate much further because they reverberate through the long walls.

**Figure 2** Weighted algorithm: schema (see online version for colours)




In contrast to the technique above, range-free algorithms do not take RSSI-values into account. If a mobile sensor node has a range of 10 m, then a fixed node can only receive his messages if the mobile node is maximum 10 m away. This is the only information that is used to calculate the position of a mobile node. For this approach, it is important that the transmission power is well configured. If the power is too low, the mobile node could be out of range between two anchors. On the other hand, if the power is too high, too many fixed nodes will receive the beacon and a wrong estimation could be made.

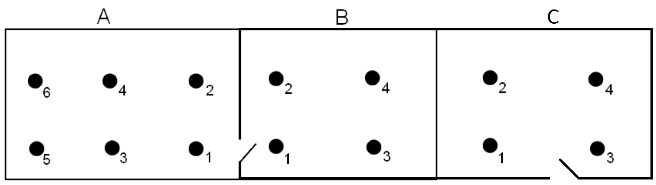
The latter problem can be solved by using a centroid algorithm. This is only useful if there is a set of fixed nodes with an overlapping coverage area. The beacon of the mobile node is received by multiple fixed nodes. To determine the position, the centroid of all the receiving fixed nodes is calculated:

$$x_M = \frac{\sum_{n=0}^{k} x_n}{k} \qquad y_M = \frac{\sum_{n=0}^{k} y_n}{k} \tag{14}$$

In theory, this algorithm would give a 100% guarantee that room-accuracy is possible. However, experiments have shown that this is not always the case. If the walls are small enough and/or do not strongly attenuate the signal, signals can go through and a fixed node in a different room can receive up the beacon. To prevent incorrect location estimation, extra logic can be added to the algorithm.

The extra logic takes the form of additional environmental metadata. Suppose we have the exact coordinates of all the walls, doors and nodes inside a building. Knowing that every beacon has an index number, the direct path could be checked between the two fixed nodes who received the consecutive beacons. If the mobile node goes from one room to another, without using a door, then the last beacon can be dismissed. For example (Figure 3) when node $A_2$ receives a beacon and the next beacon is received by node $B_2$. It is impossible to move directly from $A_2$ to $B_2$ without passing nodes $A_1$ and $B_1$. So the message that was received by beacon $B_2$ will be rejected.

**Figure 3** Three neighbouring offices




With this optimisation room-accuracy can be guaranteed. Still, this solution has the drawback that a lot of fixed infrastructure sensor nodes are necessary to retrieve good results. If the network is sparse distributed, then the algorithm would not work properly.

Finally, for the evaluation of this solution, experiments were performed using four different Tx power levels (Tx3, Tx7, Tx19 and Tx31), as shown in Table 2.

**Table 2** Used Tx power levels for the weighted RSSI localisation experiments

| *Tx power level* | *Output power* [dBm] |
|---|---|
| 3 | –25 |
| 7 | –15 |
| 19 | –5 |
| 31 | 0 |

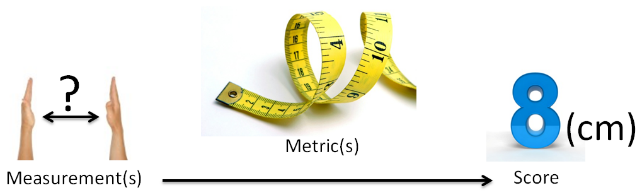
## 4 The benchmarking methodology

### *4.1 Introduction*

One of the major problems of indoor localisation is the challenge of reproducing research results in real life scenarios and the inability to compare their performance owing to evaluation under individual, not comparable and not repeatable conditions. Therefore, contrary to previous approaches, our benchmarking approach does not focus exclusively on the accuracy of the evaluated localisation approach, but also considers other performance measures that are relevant from the point of view of practical deployment of localisation solutions such as energy efficiency and response time.

Owing to variation in the sensibility of different use-case scenarios on the individual metrics, the methodology cleanly decouples between evaluating individual metrics and calculation of a final score used for ranking. As illustrated on Figure 4, after collecting a set of measurements necessary for the calculation of the individual metrics, the methodology envisions the use of weighting factors and thresholding for the calculation of the final ranking score, reflecting the different impact of the individual metrics for the particular application scenario of interest.

**Figure 4** Transform measurements to scores using metrics (see online version for colours)




## 4.2 Used metrics

The metrics that will be used for the evaluation of the solutions will have a critical impact on the final score. A classical mistake by other comparison and evaluation tools is only using the point accuracy as a reference for a good or bad working solution. In this paper, we will take others metrics into account as well, which are defined in the EVARILOS Benchmarking Handbook (Van Haute et al., 2013a).

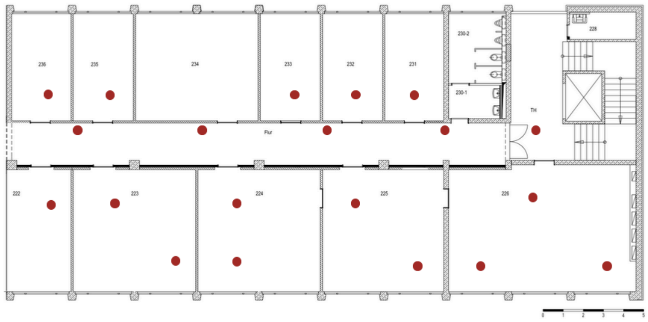
## 4.3 Used scenarios

Each solution is evaluated using a predefined scenario in each testbed. These are based on the generic scenario descriptions of the EVARILOS Benchmarking Handbook. In the next paragraphs, we will describe each scenario of each testbed. A detailed overview of each testbed is given in Section 5.

### 4.3.1 TWIST testbed

The scenario is instantiated on the 2nd floor of the TWIST testbed, and can be characterised as a 'small office environment' according to the EVARILOS Benchmarking Handbook. The evaluation points used to evaluate the localisation solutions are shown in Figure 5. These points were selected based on the Latin Hypercube principle, taking into account that there are limitations owing to unreachable places.

**Figure 5** TWIST evaluation points utilised for the first benchmarking scenario (see online version for colours)

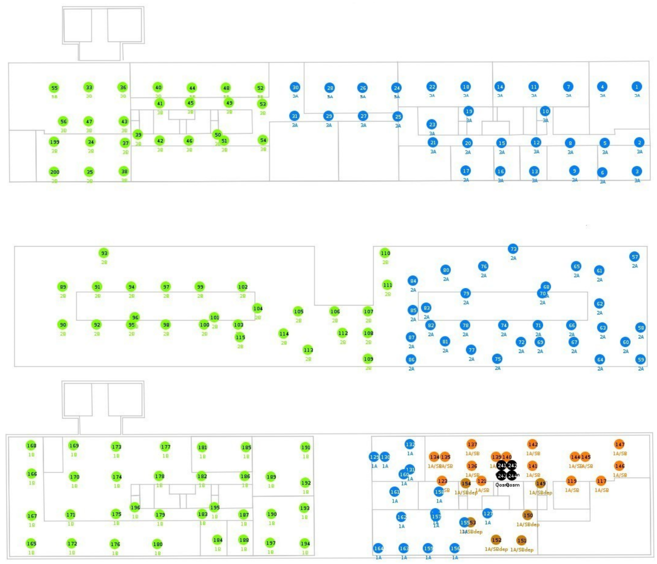


### 4.3.2 w-iLab.t I testbed

At w-iLab.t I testbed, we use the third floor to execute the experiments. On this floor, 57 nodes are available for the experiments. An overview of the third floor is given in Figure 6. There is no actual difference between the green and the blue dots, it is for reservation purposes only.

Unfortunately, not the whole floor can be considered as a test area. Some private offices, technical staff room, etc. are not available for measuring. The unreachable zones are marked with a red layer. to define the measurement points, we used a grid (see Figure 7). The decision has to be taken without premeditation. Therefore, a randomiser is used. To avoid measurement points close to each other, making an unbalanced distribution, the principle of the Latin Square is applied.

**Figure 6** The w-iLab.t I wireless testbed: map (see online version for colours)



### 4.3.3 w-iLab.t II testbed

In this testbed, the 26 measurement points are well spread over the area. These are selected by randomness in each sub-area. This is shown in Figure 8. Special in this setup is that there is no physical person present in the building. Everything is controlled remotely using robots. In this way, the repeatability of the measurement point is very high. On the other hand, this 'open environment' is made of metal walls and contains a lot of metal objects, making it very challenging for accurate localisation.

**Figure 7** Measurement points in the w-iLab.t I testbed (see online version for colours)

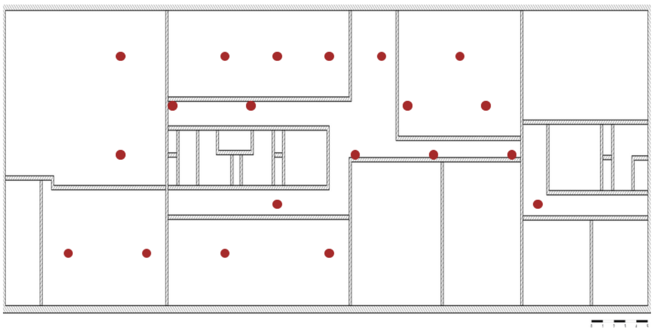
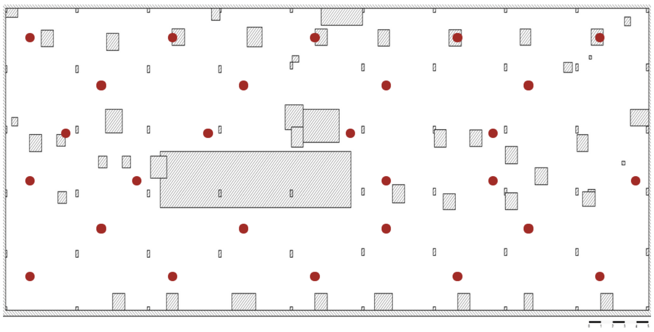


**Figure 8** The measurement points of w-iLab.t II testbed (see online version for colours)



## 5 Test environments

### *5.1 TWIST testbed in Berlin*

The TKN wireless indoor sensor network testbed (TWIST) is a multiplatform, hierarchical testbed architecture developed at the Technische Universität Berlin. The TWIST instance at the TKN office building is one of the first and most popular remotely accessible testbeds (Handziski et al., 2006). It has 204 SUT sockets, currently populated with 102 eyesIFX and 102 Tmote Sky nodes (Figure 9, with location of TMote Sky nodes indicated with red and locations of WiFi access points with purple dots.). Tmote Sky nodes use a Tiny OS development environment. They consist of a TI MSP430 processor running at 8 MHz, 10 KB of RAM, 1 Mbit of flash memory and an IEEE 802.15.4 compliant Chipcon CC2420 radio operating at 2.4 GHz with a maximum indoor range of approximately 100 m. Each node includes sensors for light, temperature, and humidity. The hardware setup is extendible with a large variety of other radios (e.g., Software Defined Radio, sensing engine), as long as the radio has a USB or RS232 serial interfaces. The nodes are deployed in a 3D grid spanning 3 floors of an office building at the TUB campus, resulting in more than 1500 m$^2$ of instrumented office space. In small rooms ($\sim$14 m$^2$), two nodes of each platform are deployed, while the larger ones ($\sim$28 m$^2$) have four nodes (Figure 10). This setup results in a fairly regular grid deployment pattern with intra node distance of 3 m. Within the rooms the sensor nodes are attached to the ceiling.

**Figure 9** TWIST testbed: nodes (see online version for colours)

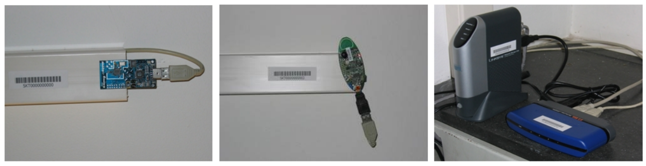
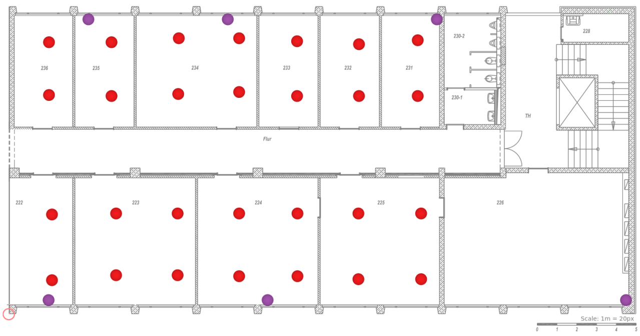


**Figure 10** TWIST testbed: map (2nd floor) (see online version for colours)



For specific purpose of benchmarking of RF-based indoor localisation, in addition to the described sensor network, the TWIST infrastructure consists of multiple other devices, as described in Lemic et al. (2014c,b). Deployed WiFi access points are commercial of-the-shelf TL-WDR4300 routers (Figure 11). The WiFi routers can serve two functions. They can be used as a part of the localisation solution, if particular solution requires WiFi anchor points. At the same time, some routers can also be used for creating different types and amounts of IEEE 802.11 traffic to generate controlled WiFi interference.

For supporting mobility and automation of the localisation measurements multiple TWISTbot robotic platforms (based on the TurtleBot design from Willow Garage), are used. Their function is to carry nodes that need to be localised through the measurement points and report the ground truth position.

**Figure 11** TWIST testbed: hardware components (see online version for colours)

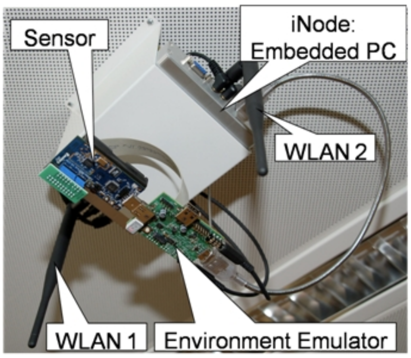
Furthermore, the TWIST infrastructure is complemented by several WiSpy sensing devices: these are low-cost spectrum scanners that monitor activity in the 868 MHz, 2.4 and 5 GHz spectrum, and output the measured RF energy and the quality of the received signals. Also, for more precise sensing of the wireless environment spectrum analysers are used. Except for the before mentioned WiFi routers and TWIST platform, for generating interference Rohde & Schwarz signal generator is used. Signal generator can be used for generating arbitrary RF signals, and its usual usage is for generating microwave interference.

## *5.2 w-iLab.t I Testbed at De Zuiderpoort*

The w-iLab.t I testbed is located at 'De Zuiderpoort' in Ghent, Belgium. The infrastructure is distributed on three floors (18 $\times$ 90 m) of the iMinds office (Figure 6). The network consists of 200 nodes. Every w-iLab.t node is generic and is equipped with one or more sensor nodes, an intermediate node with 2 WiFi 802.11 radios, the environment emulator (EE) and a Bluetooth interface. As in TSIT testbed, the sensor nodes are Tmote Sky motes.

The intermediate nodes (called iNodes, Figure 12) are Alix 3C3 devices running Linux. These are mini PCs equipped with Ethernet, USB, serial, VGA, audio and two IEEE 802.11 a/b/g interfaces. All the iNodes are connected to the management backbone using Power-over-Ethernet switches, making it possible to power up/down the iNodes as needed without physical interaction with the iNodes. The iNodes can become an active member of the experiment as it is possible to adjust the kernel, the driver, to add click router code or to add java-based applications.

**Figure 12** iNode mounted to the ceiling of the w-iLab.t I wireless testbed (see online version for colours)




Finally, the EE is located in between the iNode and the sensor node. Using the EE, it is possible to emulate the behaviour of any type of sensor or actuator without the need for real sensoractuator hardware or the development of a full-blown sensor application. It is possible to emulate the battery depletion, depending on the real life power consumption of the sensor node. When the node's battery is depleted or the node is destroyed (e.g., in an explosion), the node can be switched off. The EE can be programmed to emulate a sensor event (e.g., temperature rise, motion detection), an actuator event or to support voice streams. Further, the EE can be used to monitor the energy consumption of each individual sensor. Altogether, this means that it is possible to assess the complete usability of a certain wireless sensor and actuator network application or protocol in a real-life environment. The initial core of w-iLab.t was based on the widely used MoteLab testbed from Harvard University. According to the EVARILOS Benchmarking Methodology This building belongs to the category 'Plywooden walls' and the size is 'Big'.

This is a classic office environment where multiple devices communicate wirelessly with each other. Laptops using WiFi and Bluetooth, smartphones using the 3G network. Here we consider typical office applications like email, file transfer, videoaudio conferencing and web surfing. The office environment is a live environment. Meaning the interference in this testbed is uncontrolled. During daytime several people are working in these buildings. So the w-iLab.t I is a testbed with very realistic office interference. The cost of this realistic office environment is the uncontrollable interference.

The w-iLab.t I testbed is centrally managed for control and monitoring purposes. It supports easy configuration and deployment, including installation of new software, protocols and middleware components via an intuitive web-based interface. Registered users can upload executables, associate those executables with the nodes (both sensor nodes and iNodes) to create a job, and schedule the job to be run on w-iLab.t I. During the job, all messages and other data are logged to a database, which is presented to the user upon job completion, and then can be used for processing and visualisation.

All the possibilities of the complete testbed, the EE scenarios and events, a visualisation and a graphical analysis tool, are accessible through a web interface. The visualisation too can visualise any type of node status and/or link information on a map of the building, while the graphical analyser plots out the data. The information for both tools is gathered from the database through the use of user customisable MySQL statements, making it extremely flexible. External users can access the testbed over a secured OpenVPN connection.

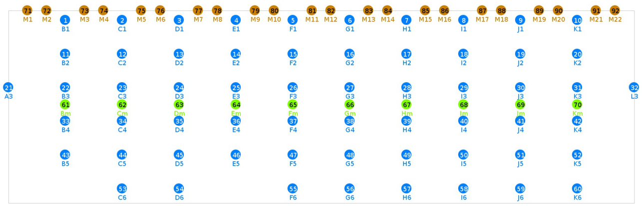
## *5.3 w-iLab.t II in Zwijnaarde*

The w-iLab.t II testbed is located in 'Zwijnaarde', above a cleanroom. At this location, there is (almost) no interference. It is one open space where 60 fixed nodes are distributed over an area of 70 $\times$ 25 m (Figure 13). In this environment there are also 20 mobile nodes. These nodes are based on a vacuum cleaning robot and are extended with a radio for remote control and accurate positioning algorithms (with rasters on the floor). Owing to the fact that the movement of these robots is controlled, mobility is reproducible. The fixed nodes are marked with blue spots while the mobile nodes have orange spots on the map in Figure 13. Every node location contains

- a Zotac embedded PC
- an EE (see w-iLab.t I)
- an iMinds Rmoni sensor node
- a Bluetooth dongle and some of them have a web-camera.

These nodes are remotely powered by Racktivity PDUs.

**Figure 13** The w-iLab.t II wireless testbed: map (see online version for colours)



## 6 Results in TWIST testbed

### *6.1 RSSI and ToA with particle filter*

In this section we evaluate the particle filter localisation approach outlined in Section 3.1. Table 3 presents the summarised results. In this table, 'Min.' stands for Minimum, 'Max.' for Maximum and 'RMS' for root mean square. These abbreviations are also used in the other tables.

**Table 3** Statistical information about the performance of the particle filter algorithm in TWIST testbed

| *Metric* | *RSSI* | *ToA* |
|---|---|---|
| *Average error* [m] | 4.35 | 5.56 |
| *Min. error* [m] | 0.62 | 0.68 |
| *Max. error* [m] | 12.99 | 22.47 |
| *Median error* [m] | 3.22 | 3.91 |
| *RMS error* [m] | 5.28 | 7.11 |
| *Room accuracy* [%] | 45.00 | 30.00 |
| *Response time* [ms] | 14,285 | 14,282 |

Measurements are collected for over a minute at each measurement point. Most of this time is spent trying to reach nodes that are not within reach, and finding the channel that a reachable testbed node currently is using. The response time could be decreased significantly by dedicating a single channel for communication to be used before starting the ranging phase. Moreover, range estimations do not improve significantly after 1300 measurements, as shown in Figure 14. Therefore, we use only the first 1300 collected values for our range estimations. This also helps limiting the response time because each measurement takes on average 4 ms. We use the same approach for the RSSI measurements, although the figure shows that 50 measurements are likely to be enough. The figure also shows that after approximately 500 measurements the ToA based method performs better.

Figure 15 shows the CDFs for the absolute range errors and the localisation errors. The RSSI based range estimation performs better than the ToA based estimation, although Figure 14 shows that ToA should give better results for a high number of measurements. The reason for this is that only about 50% of the pair-wise ranging procedures result in 500 measurements or more, and only about 10% result in 1300 measurements or more.

The power consumption of both the target node and the testbed nodes is approximately 105 mW. It is computed as the mean of the transmission and reception power consumptions. The energy consumption per node is especially important when battery powered devices will be used, since it directly impacts the lifetime of a battery-powered localisation solution. The node energy consumption can also be used to calculate the overall energy consumption. The infrastructure nodes are always on, and a total of 68 testbed nodes are used. As a result, the continuous total power consumption can be calculated to be 7.1 W for the infrastructure. The mobile node is only on during the response time, which is in the order of 15 s, resulting in an average energy consumption of 1.5 J per measurement.

**Figure 14** The absolute range error for ToA decreases with the number of measurements until approximately 1300 measurements. The RSSI error fluctuates about the same value, and is not improved by additional measurements (see online version for colours)


### *6.2 Fingerprinting*

This section evaluates the fingerprinting localisation approach described in Section 3.2. The accuracy results are shown in Table 4. The results show that the PH Distance of RSSI Quantiles give comparable results with the ED Distance of Averaged RSSI Vectors in office scenarios (see also Figure 16). The results also show that, when more beacons are collected and thus the response time increases, the PH has a slightly better overall performance in terms of accuracy. This improvement will be more emphasised in the open space scenario (see results in Section 8). The minimum error of all solutions equals zero, which is possible because some of the fingerprints taken during the training set are at the same locations that were used for the evaluation of the algorithm.

Since a localisation solution, in general, contains many configurable parameters, we expect that they will typically be offered to end users using predetermined configuration setting. As such, it is important to be aware of the inherent trade-offs that are made by the developer of the solution. This is especially important when considering also additional metrics such as the response time. For this solution, the time during which fingerprints are collected (e.g., the time needed before a location estimate could be generated) was set to 35 s (excluding the off-line time required for fingerprinting). Figures 17 and 18 show the trade-offs between response and point and room level accuracies for fingerprinting based solutions in TWIST testbed. As more fingerprints are collected, a better match can be made to better estimate the position. Lower response times are possible, at the cost of decreased accuracy. Especially more complex algorithms (such as the PH distance) require more samples to estimate the distributions of the RSSI values. As such, when comparing different localisation solutions, the targeted response time has an important influence on the selection of the best algorithm.

**Figure 15** CDFs for the absolute range error (top) and the localisation error (bottom) (see online version for colours)

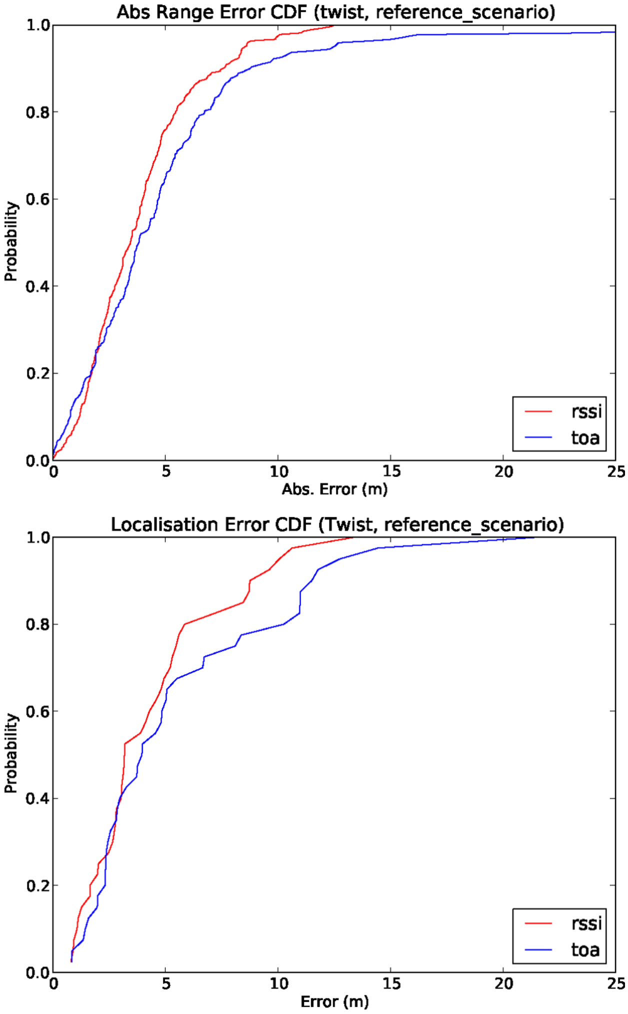
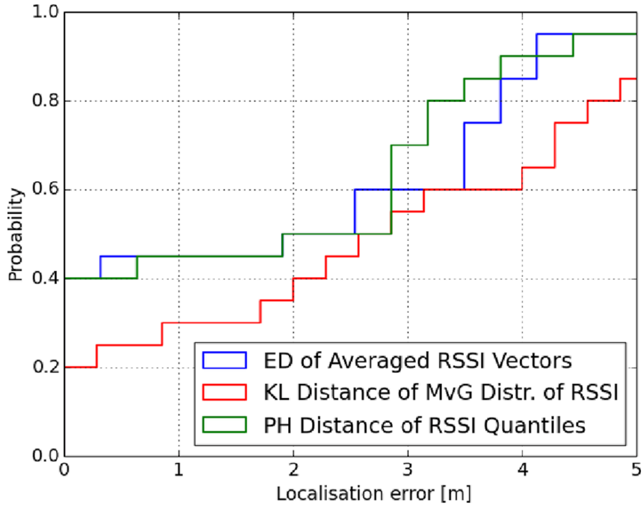



**Table 4** Statistical information about the performance of fingerprinting algorithms in TWIST testbed

| *Metric* | *KL* | *ED* | *PH* |
|---|---|---|---|
| *Average error* [m] | 2.77 | 2.16 | 2.02 |
| *Min. error* [m] | 0.00 | 0.00 | 0.00 |
| *Max. error [m]* | 5.71 | 6.35 | 6.35 |
| *Median error* [m] | 2.98 | 2.48 | 2.52 |
| *RMS error* [m] | 3.39 | 2.95 | 2.79 |
| *Room accuracy* [%] | 50.00 | 80.00 | 85.00 |
| *Response time* [s] | 35.67 | 35.11 | 35.12 |

**Figure 16** CDF of the localisation error of fingerprinting based solutions in TWIST testbed (see online version for colours)




Finally, the energy consumption of the infrastructure nodes (TPLINK 4300 router) is on average 0.5 W, whereas the energy consumption of the used mobile devices (MacBook Pro AirPort Extreme NIC) was on average 7 W.

## 6.3 *Proximity and weighted RSSI*

This section evaluates the RSSI based localisation approach described in Section 3.3. The obtained accuracy is summarised in Table 5 for different transmission powers. The average accuracy is relatively low: the concrete walls in the building cause unpredictable signal attenuation, resulting in less accurate estimations of the true location. Using lower transmission powers causes less signals to propagate to multiple rooms, hence the better performance of low transmission powers. A CDF of the errors is shown in Figure 19.

**Figure 17** Fingerprint collection delay vs. point accuracy (see online version for colours)

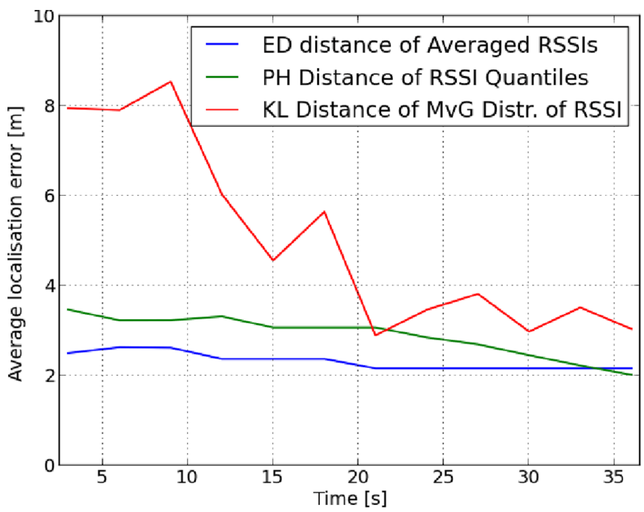


**Figure 18** Fingerprint collection delay vs. room level accuracy (see online version for colours)

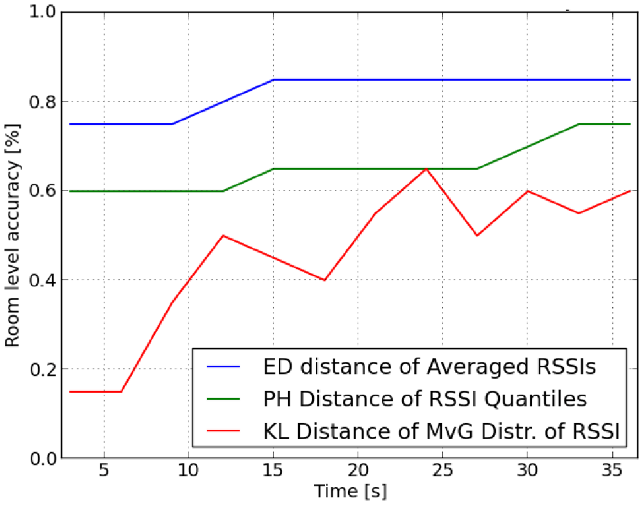



**Figure 19** Distribution of the RMS localisation error in TWIST testbed (see online version for colours)

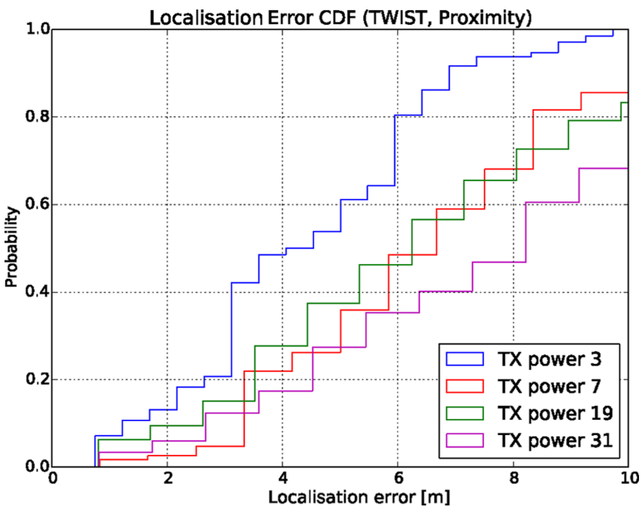


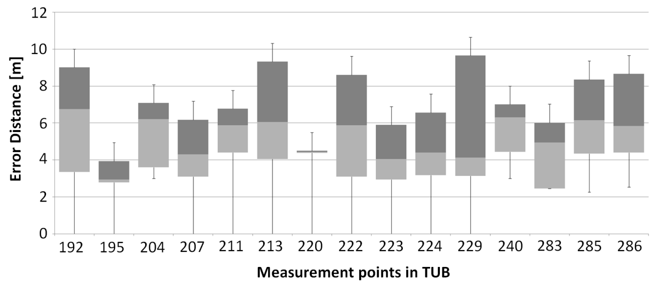

**Table 5** Statistical information about the performance of the hybrid algorithm in TWIST testbed

| *Metric* | *Tx3* | *Tx7* | *Tx19* | *Tx31* |
|---|---|---|---|---|
| *Average error [m]* | 4.63 | 7.08 | 6.93 | 8.31 |
| *Min. error* [m] | 0.75 | 0.83 | 0.80 | 0.82 |
| *Max. error* [m] | 10.20 | 17.52 | 18.93 | 19.31 |
| *Median error [m]* | 4.39 | 6.81 | 6.68 | 8.63 |
| *RMS error* [m] | 5.13 | 7.75 | 7.82 | 9.24 |
| *Room accuracy* [%] | 26.67 | 6.70 | 13.45 | 9.56 |
| *Response time* [ms] | 1503 | 1507 | 480 | 460 |

To estimate the position, the anchor points collect RSSI values from the beacons transmitted by the mobile node. All these RSSI values are collected and merged in the position calculator. There, a translation from RSSI values into coordinates is made. For low transmission powers, the corresponding response delay is about 1.5 s (exact values are given in Section 6.4), with an energy consumption of about 31 mW for the mobile node.

To analyse the spatial distribution of the errors, a box-plot of the accuracy per measurement point is shown in Figure 20. The overall performance is for each measurement point the same, there are no obvious outliers. Noticeable, the worst minimum values are obtained in the corridor (measurement points 283, 285 and 286) and in the room where no LoS nodes are available (measurement point 240). If the results of these rooms are excluded in the room accuracy calculation, then the results are marginally better (e.g., for a Tx power = 7 the room accuracy increases to 21.3% instead of to 10.3%), but even in the rooms where the nodes were available, the average error distance of almost 5 m is not enough to guarantee room accuracy: only 33.8% of all the measurement points are in the same room.

**Figure 20** Proximity and weighted RSSI solution – spatial distribution of accuracy error in TWIST testbed, including maximum error, minimum error, quartile 1, quartile 2 and median error




The main reason for these results is that proximity requires extremely low transmission powers: even using the lowest transmission powers from the TMoteSky nodes, signals still easily penetrated the walls. Finally, the box-plot of measurement point 220 is also remarkable. The most logical explanation for this result is that only a few fixed nodes received the beacons of the mobile node. As a result, the calculator does not have much data to process. This makes the result very stable, but not necessarily more accurate.

## *6.4 Conclusions from the TWIST experiments*

An overview of the performance of the different localisation solutions is given in Table 6. In terms of accuracy, the best performing solutions are the fingerprinting localisation techniques. Since the TWIST building represents typical office buildings with concrete and/or brick walls, different rooms are very diverse in terms of their wireless characteristics. Model based localisation solutions (such as RSSI based solutions) suffer from degraded performance owing to unexpected obstacles. In contrast, localisation solutions that exploit this diversity, such as fingerprinting based approaches, obtain the highest accuracy.

When also considering other metrics, these conclusions need to be nuanced. In terms of response time, fingerprinting performs worst, owing to the need to collect a minimal number of beacons. Since the beacon interval is not always configurable on already deployed access points, it is not always possible to decrease the response time when using existing off-the-shelf access points. In contrast, the ToA solutions can give location estimates in only halve of the response time (about 15 s vs. 35 s), and as shown earlier in Figure 14, the response time of the RSSI based solutions can theoretically be reduced to about 200 ms.

**Table 6** TWIST testbed: summarised benchmarking results

| *Algorithm* | *Average error* [m] | *Room accuracy* [%] | *Response time* [ms] | *Energy efficiency* [mW] *Mobile* | *Fixed* |
|---|---|---|---|---|---|
| *Particle filter solution* | | | | | |
| Spray RSSI | 4.35 | 45.00 | 14,285 | ∼105 | ∼105 |
| Spray ToA | 5.56 | 30.00 | 14,282 | ∼105 | ∼105 |
| *Fingerprinting solution* | | | | | |
| KL distance | 2.7 | 50.0 | ∼35,000 | ∼7000 | ∼500 |
| ED distance | 2.2 | 80.0 | ∼35,000 | ∼7000 | ∼500 |
| PH distance | 2.0 | 85.0 | ∼35,000 | ∼7000 | ∼500 |
| *Hybrid solution* | | | | | |
| TX Power = 3 | 4.6 | 26.7 | 1503.1 | ∼30.9 | ∼47.4 |
| TX Power = 7 | 7.1 | 6.7 | 1507.6 | ∼35.1 | ∼47.4 |
| TX Power = 19 | 7.9 | 13.4 | 480.6 | ∼47.1 | ∼47.4 |
| TX Power = 31 | 8.7 | 9.5 | 460.9 | ∼57.6 | ∼47.4 |

In terms of energy consumption, the devices used in the fingerprinting solution consume most energy, which means that battery-powered solutions will have a low network lifetime when using IEEE 802.11 based fingerprinting. The energy consumption of the IEEE 802.15.4 devices is significantly lower. However, owing to the large number of measurements required, the ToA still consumes twice the energy of the hybrid solution. It is clear that the energy consumption could be further optimised, albeit at the cost of longer response times.

Finally, the fingerprinting approach, although the most accurate, has one other disadvantage, which is not taken into account by considering only the shown metrics. More specifically, the need for an off-line training phase and the need for retraining if the environmental conditions change can significantly impact the accuracy over time in realistic conditions. This clearly shows the need for an objective comparison method that takes into account multiple evaluation criteria when comparing localisation solutions.

## 7 Results in w-iLab.t I testbed

### 7.1 RSSI and ToA with particle filter

The results for the particle filter based localisation using RSSI and ToA range measurements are presented in Table 7. Measurements are collected and processed in the same way as in the TWIST testbed.

Figure 21 shows how the number of measurement affects the accuracy. As in TWIST testbed, ToA benefits from more measurements, while RSSI based ranging does not. A major difference to TWIST, also shown by the CDFs to the left in Figure 22, is that both types of range measurements have much larger errors. A reason for this can be that more testbed nodes far away from the measurement points are reachable, and that the far travelling signals are subject to multi-path effects to a greater extent, resulting in unpredictable attenuation that is not captured by the free-space model. The response time is also much higher than in TWIST testbed. This is also because more testbed nodes are used for each measurement. The response time can be reduced if no measurements are collected after a certain number of testbed nodes have been used.

**Figure 21** The absolute range error for ToA decreases with the number of measurements until approximately 2500 measurements. The RSSI error fluctuates about the same value, and is not improved by additional measurements (see online version for colours)

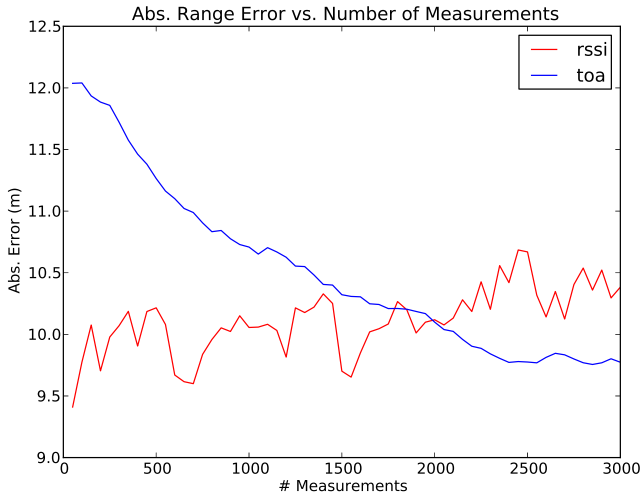



**Table 7** Accuracy of the particle filter in the w-iLab.t I testbed

| *Metric* | *RSSI* | *ToA* |
|---|---|---|
| *Average error* [m] | 7.79 | 7.16 |
| *Min. error* [m] | 3.59 | 1.51 |
| *Max. error* [m] | 14.04 | 14.31 |
| *Median error* [m] | 7.09 | 6.09 |
| *RMS error* [m] | 8.43 | 7.92 |
| *Room accuracy* [%] | 30.00 | 20.00 |
| *Response time* [ms] | 55.45 | 55.44 |

The right graph in Figure 22 shows the CDF for the localisation error. The median error is about the same as that for the range measurements in the left graph (8 and 7 m, respectively), but has lower errors above the median.

The power consumption for the mobile node is the same as in the TWIST experiments, i.e., 105 mW. Although less testbed nodes are used in this experiment, we consider the total infrastructure power consumption to be of the same magnitude as in TWIST, i.e., 7 W.

### 7.2 Fingerprinting

This section evaluates the fingerprinting localisation approach described in Section 3.2 in the w-iLab.t I testbed. The accuracy results are shown in Table 8. In general, the accuracy is lower than in the TWIST testbed environment. The CDF of localisation error is shown in Figure 23. The decrease in accuracy can be explained by the fact that the w-iLab.t I testbed is an office environment that uses plywood walls, which attenuate the signals less than the concrete walls in the TWIST testbed.

**Figure 22** CDFs for the absolute range error (top) and the localisation error (bottom) (see online version for colours)

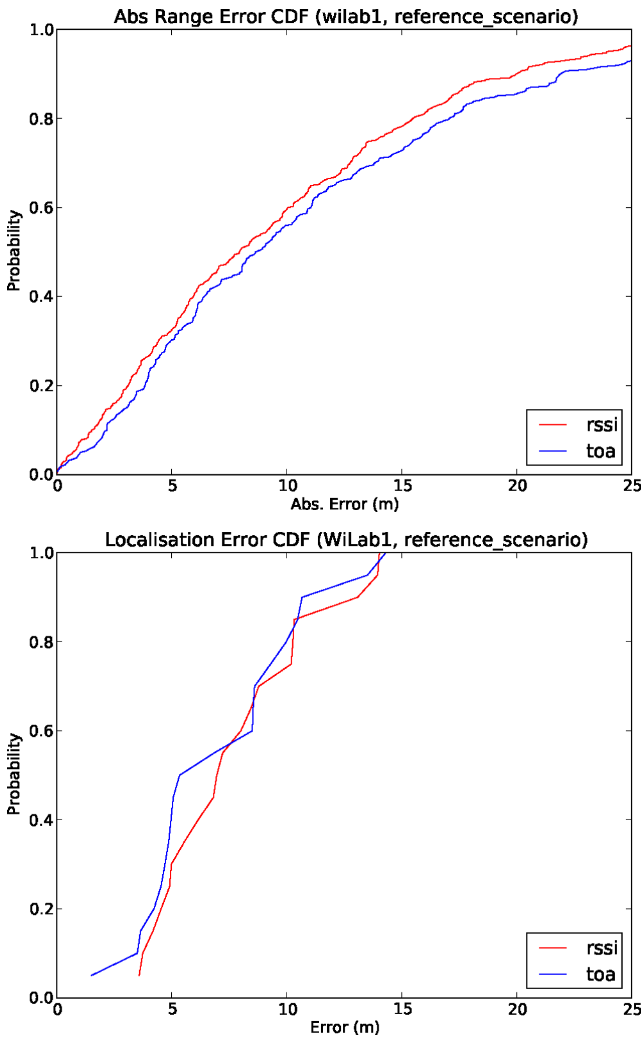
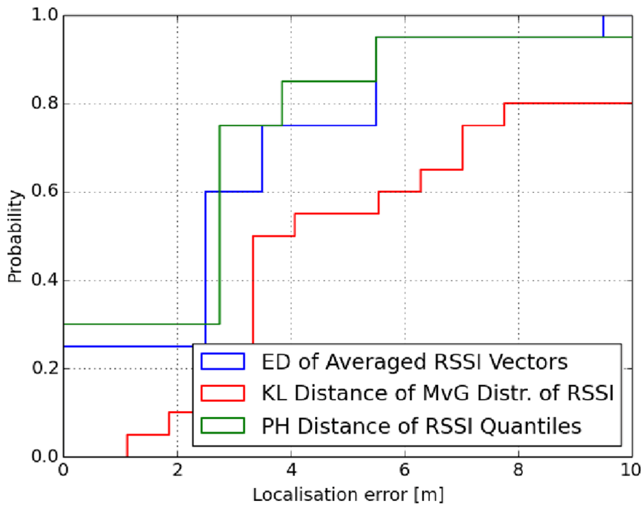



**Table 8** Statistical information about the performance of the fingerprinting algorithms in w-iLab.t I testbed

| *Metric* | *KL* | *ED* | *PH* |
|---|---|---|---|
| *Average error* [m] | 6.15 | 2.37 | 2.75 |
| *Min. error* [m] | 1.12 | 0.00 | 0.00 |
| *Max. error [m]* | 15.86 | 5.50 | 11.0 |
| *Median error* [m] | 4.37 | 2.75 | 2.75 |
| *RMS error* [m] | 7.25 | 3.34 | 3.76 |
| *Room accuracy* [%] | 50.0 | 80.0 | 85.0 |
| *Response time* [s] | 24.98 | 24.16 | 24.36 |

As a result, the evaluated locations have less diversity in terms of the received signal strengths, and are thus more difficult to uniquely characterise in a fingerprint. This effect will have an even greater influence in the results of open space environment of w-iLab.t II (Section 8). The point and room level accuracy vary with the performance delay, as presented in Figures 24 and 25.

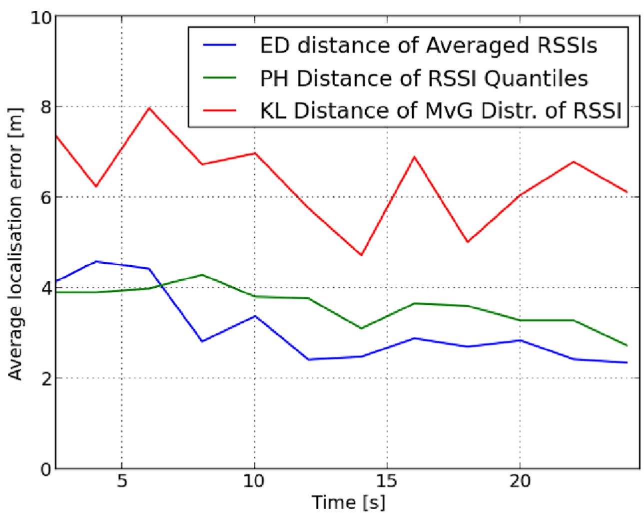
**Figure 23** CDF of the localisation error of the fingerprinting based solutions in w-iLab.t I testbed (see online version for colours)




## 7.3 *Proximity and Weighted RSSI*

The location accuracy of the weighted RSSI based localisation solution in w-iLab.t I testbed is shown in Table 9. The CDF of the localisation error can be found in Figure 26. Because the plywood walls do not attenuate the signals significantly, locations need to be determined based on weighted RSSI values (rather than proximity) even when using low transmission powers. As a result, in contrast to the experiments in the TWIST testbed, where the localisation accuracy depends strongly on the transmission power, the results in w-iLab.t I testbed are less dependent on the transmission power.

**Figure 24** Fingerprint collection delay vs. point accuracy (see online version for colours)




The spatial spread of the accuracy is shown in Figure 27 using a box-plot. The measured points more in the centre of the testbed have a higher accuracy then those at the edges, because the edge evaluation points are outside the grid of used anchor points. For example, in Figure 28, a clear bias in the estimated locations can be observed caused by the fact that all anchor nodes are located at the same side of the evaluation point. This highlights the importance of using anchor nodes outside the area that is evaluated, which is a requirement that is not found for the fingerprinting solutions. Finally, Figure 29 shows the room accuracy of location estimation. It is interesting to note that it is not possible to predict the room accuracy based only on the point accuracy, because the room accuracy depends strongly on random factors such as the direction of the inaccuracies.

**Figure 25** Fingerprint collection delay vs. room level accuracy (see online version for colours)

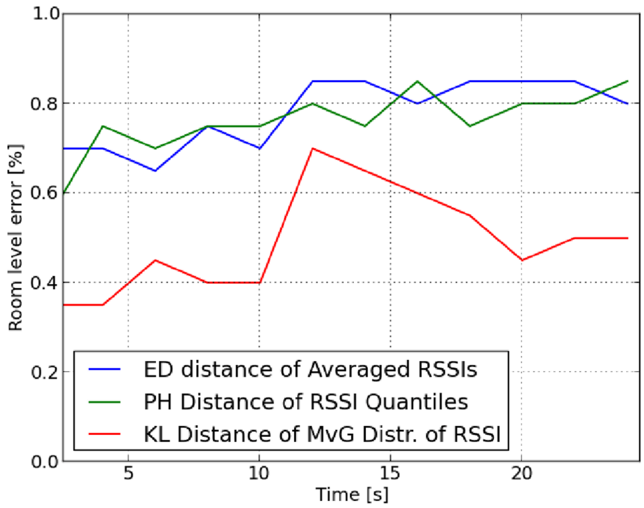



**Table 9** Statistical information about the performance of the hybrid algorithm in w-iLab.t I testbed

| *Metric* | *Tx3* | *Tx7* | *Tx19* | *Tx31* |
|---|---|---|---|---|
| *Average error* [m] | 7.64 | 8.86 | 7.47 | 8.21 |
| *Min. error* [m] | 0.07 | 0.04 | 0.17 | 0.04 |
| *Max. error* [m] | 48.77 | 65.98 | 45.23 | 45.23 |
| *Median error [m]* | 5.87 | 7.63 | 6.18 | 7.21 |
| *RMS error* [m] | 9.35 | 10.15 | 8.83 | 9.44 |
| *Room accuracy* [%] | 18.62 | 16.27 | 12.60 | 9.46 |
| *Response time* [ms] | 2100 | 113 | 108 | 110 |

## *7.4 Conclusions from w-iLab.t I experiments*

An overview of the performance of the different localisation solutions is given in Table 10. As a representative of a typical building with plywood walls, signals are less attenuated than in the TWIST testbed, resulting in less unique wireless features per room. As a result, fingerprinting solutions perform worse than in the TWIST environment. Also the ToA and RSSI based solutions have significantly degraded performance. This can be explained by the fact that, owing to testbed limitations, anchor nodes are not installed on the corner points, meaning that several evaluation points are outside the grid of anchor nodes. In addition, although the walls are made from plywood that have a very small attenuation factor, signal propagation still behaves very unpredictable owing to the presence of large metal cupboards and metal ceilings. This demonstrates that the performance of localisation solutions in typical environments is influenced by many factors besides the building construction materials, and highlights the fact that localisation performances measured in a empty building should not be considered representative for the performance of said solutions when the buildings are actively used.

**Figure 26** Distribution of the RMS localisation error in w-iLab.t I testbed (see online version for colours)

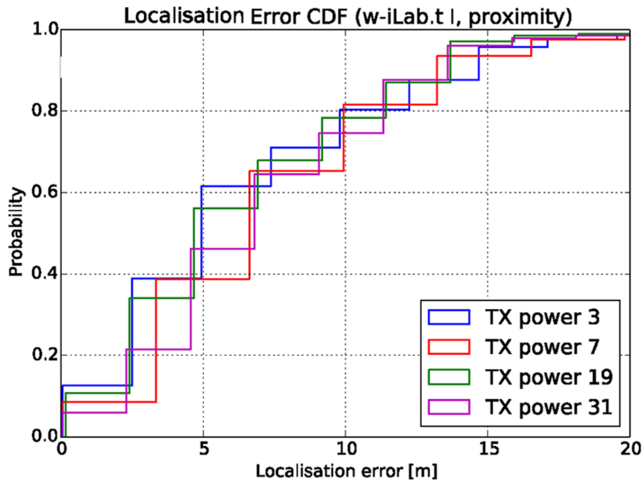



**Figure 27** Proximity and weighted RSSI solution – spatial distribution of accuracy error in w-iLab.t I, including maximum error, minimum error, quartile 1, quartile 2 and median error

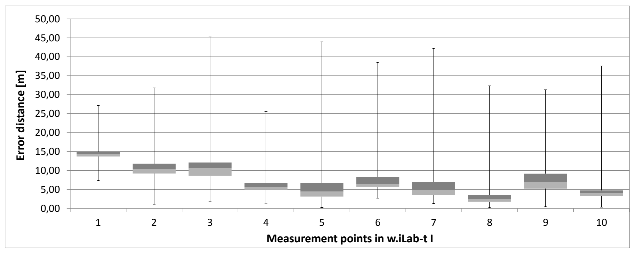



**Figure 28** Biased spread of the location estimates resulting from evaluating measurement locations outside the grid of anchor nodes (see online version for colours)

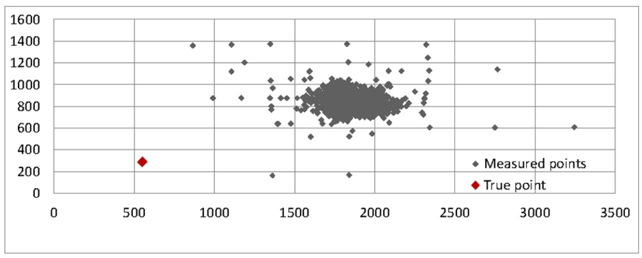


**Figure 29** The results of the room accuracy in each measurement point in the w-iLab.t I testbed (see online version for colours)

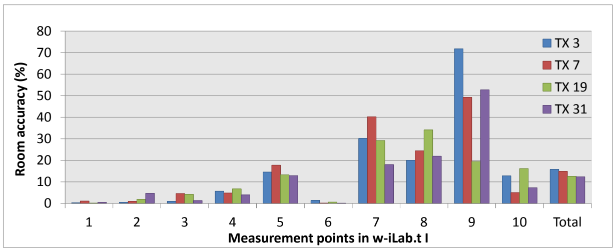



**Table 10** w-iLab.t I testbed: summarised benchmarking results

| *Algorithm* | *Average error* [m] | *Room accuracy* [%] | *Response time* [ms] | *Energy efficiency* [mW] *Mobile* | *Fixed* |
|---|---|---|---|---|---|
| *Particle filter solution* | | | | | |
| Spray RSSI | 7.79 | 30.00 | 55,448 | ∼105 | ∼ 105 |
| Spray ToA | 7.16 | 20.00 | 55,444 | ∼105 | ∼105 |
| *Fingerprinting solution* | | | | | |
| KL distance | 6.15 | 50.00 | ∼24000 | ∼7000 | ∼500 |
| ED distance | 2.37 | 80.00 | ∼24000 | ∼7000 | ∼500 |
| PH distance | 2.75 | 85.00 | ∼24000 | ∼7000 | ∼500 |
| *Hybrid solution* | | | | | |
| TX Power = 3 | 7.64 | 18.62 | 2100.74 | ∼30.9 | ∼47.4 |
| TX Power = 7 | 8.86 | 16.27 | 113.15 | ∼35.1 | ∼47.4 |
| TX Power = 19 | 7.47 | 12.60 | 107.99 | ∼47.1 | ∼47.4 |
| TX Power = 31 | 8.21 | 9.46 | 110.17 | ∼57.6 | ∼47.4 |

## 8 Results in w-iLab.t II testbed

### *8.1 RSSI and ToA with particle filter*

The results for the particle filter based localisation using RSSI and ToA range measurements are presented in Table 11.

**Table 11** Statistical information about the performance of the particle filter algorithm in w-iLab.t II testbed

| *Metric* | *RSSI* | *ToA* |
|---|---|---|
| *Average error* [m] | 6.41 | 6.66 |
| *Min. error* [m] | 0.90 | 0.99 |
| *Max. error* [m] | 20.22 | 27.06 |
| *Median error* [m] | 5.68 | 5.50 |
| *RMS error* [m] | 8.05 | 8.59 |
| *Response time* [ms] | 59633 | 59620 |

In this testbed, the measurements are collected in a slightly different way than in TWIST and w-iLab.t I testbeds. A single channel is used owing to a limitation of the testbed nodes. Moreover, instead of collecting multiple measurements from a specific testbed node before switching to the next testbed node, a single message is exchanged with each infrastructure node and when all nodes have been tried, the process starts over again with the first node. As a result, measurements are collected from more testbed nodes, but each having fewer measurements. For this reason a maximum of approximately 900 measurements are collected from a single node in this testbed. Figure 30 shows that, at least for within this range of collected measurements, the accuracy is not affected for any of the ranging methods. Although few measurements are collected from each node, we observe a high response time owing to the fact that many testbed nodes are used at each measurement point. As in the w-iLab.t I testbed, the response time can be reduced by limiting the amount of testbed nodes used at each measurement point.

**Figure 30** The absolute range error is not affected for neither ToA nor RSSI measurements (see online version for colours)

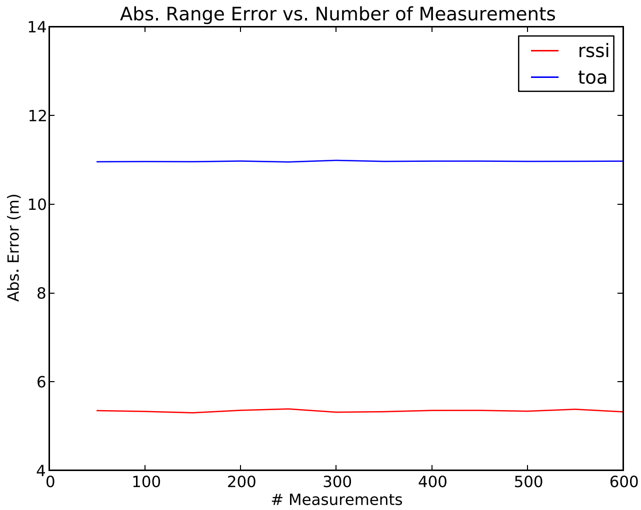



Figure 31 shows the CDFs for the range measurements and the localisation estimations. Although the ToA range measurements (left graph) are less accurate than that of the RSSI based method, the final localisation estimations (right graph) for the two methods are more or less equal. The power consumptions stated for TWIST and w-iLab.t I testbeds are also applicable here.

### *8.2 Fingerprinting*

Table 12 contains the accuracy statistics of the fingerprinting localisation solutions described in Section 3.2. No room accuracy is reported because the testbed consists of a single large open space. The accuracy is significantly lower than the accuracy obtained in the other testbeds (Figure 32). This degradation is mainly caused by two physical characteristics of the environment:

- no separate rooms are present, which makes it difficult to create unique fingerprints for each location
- owing to the metal walls, random reflections result in signal strengths that vary strongly from packet to packet.

Figure 33 shows the influence of collecting additional data before creating fingerprints. It demonstrates the importance of using robust fingerprinting creation methods (e.g., PH distance of RSSI quantiles) and demonstrates that these robust fingerprinting creation methods can be used to generate more accurate results, on the condition that more data is collected (at the cost of higher response delays).

**Figure 31** CDFs for the absolute range error (top) and the localisation error (bottom) (see online version for colours)

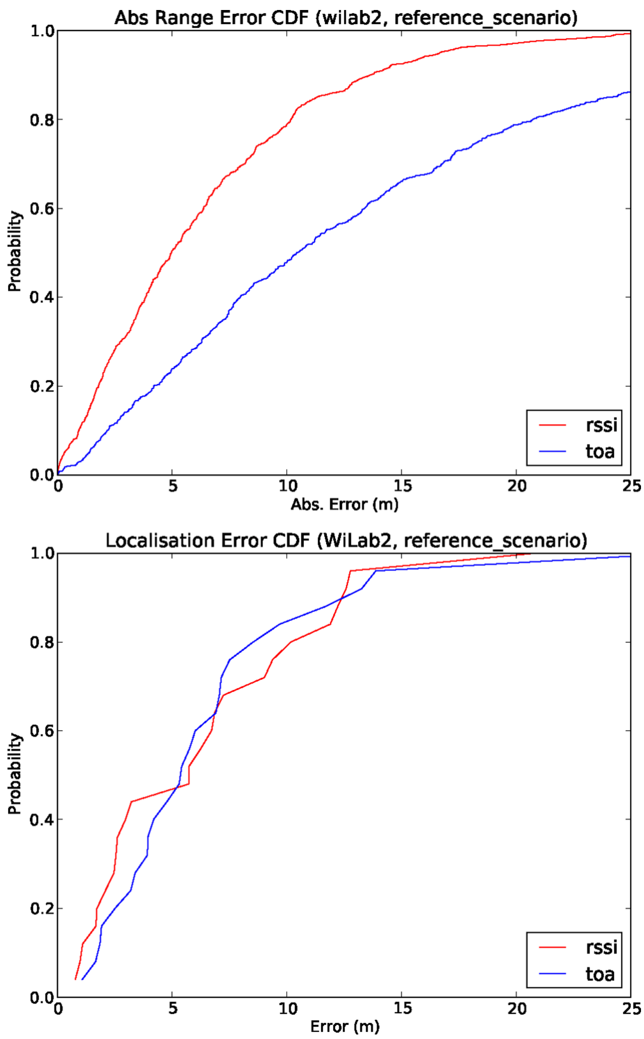



**Table 12** Statistical information about the performance of fingerprinting algorithms in w-iLab.t II testbed

| *Metric* | *KL* | *ED* | *PH* |
|---|---|---|---|
| *Average error* [m] | 24.76 | 19.08 | 8.13 |
| *Min. error* [m] | 3.00 | 3.00 | 0.00 |
| *Max. error* [m] | 47.43 | 39.00 | 15.10 |
| *Median error* [m] | 21.0 | 18.97 | 6.70 |
| *RMS error* [m] | 28.09 | 20.76 | 8.97 |
| *Response time* [s] | 24.78 | 24.37 | 24.12 |

## *8.3 Proximity and weighted RSSI*

The location accuracy of the weighted RSSI based localisation solution in w-iLab.t II testbed is shown in Table 13.

The CDF of the localisation error can be found in Figure 34. The average accuracy is significantly lower than in the previous environments, mainly owing to the many reflections in the environment thereby causing self-interference. The spatial spread of the accuracy is shown in Figure 35 using a box-plot.

**Figure 32** CDF of the localisation error of fingerprinting based solutions in w-iLab.t II testbed (see online version for colours)

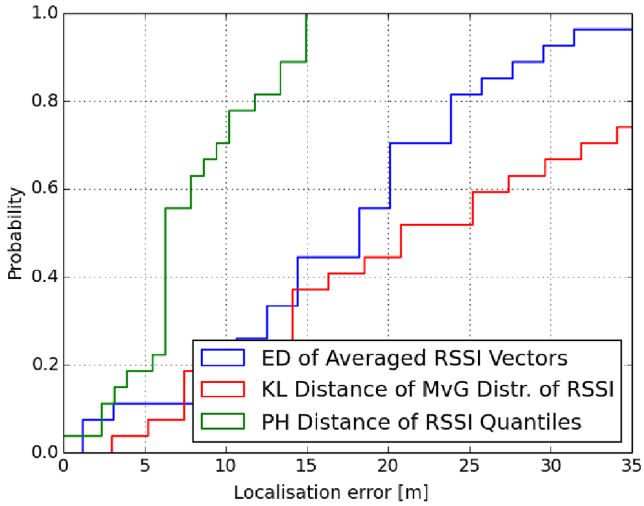



**Figure 33** Fingerprint collection delay vs. point accuracy (see online version for colours)

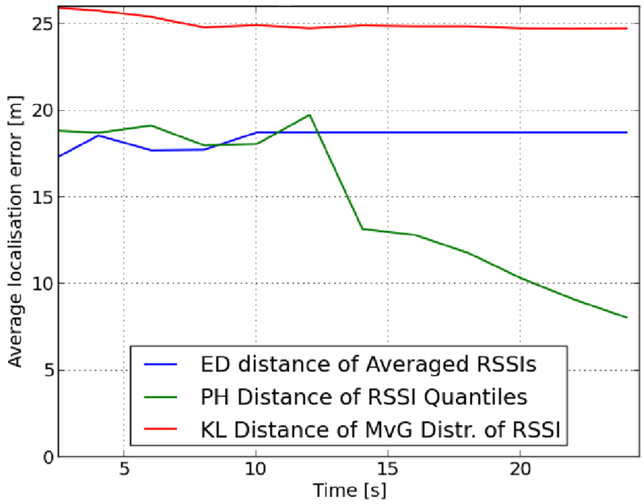



**Table 13** Accuracy of the hybrid localisation algorithm in the w-iLab.t II testbed (Tx = 31)

| | |
|---|---|
| *Average error* [m] | 17.16 |
| *Min. error* [m] | 1.57 |
| *Max. error* [m] | 52.15 |
| *Median error* [m] | 16.24 |
| *RMS error* [m] | 19.73 |
| *Response time* [ms] | 15.7 |

**Figure 34** Distribution of the RMS localisation error in w-iLab.t II testbed (see online version for colours)

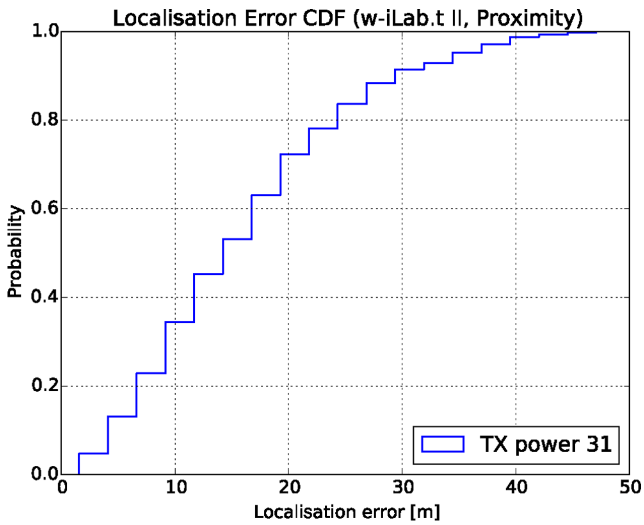
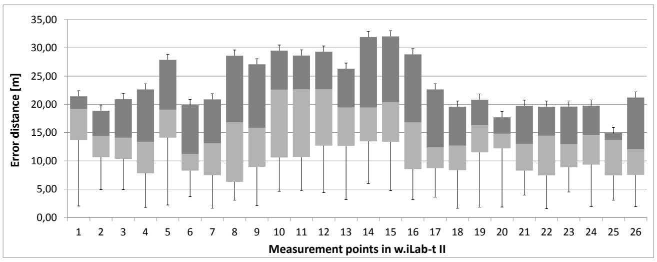



**Figure 35** Proximity and weighted RSSI solution – spatial distribution of accuracy error in w-iLab.t II, including maximum error, minimum error, quartile 1, quartile 2 and median error




## 8.4 *Conclusions from w-iLab.t II experiments*

The environment from w-iLab.t II testbeds exhibits characteristics which are typical for many large-size industrial indoor environments, namely open spaces surrounded by metal obstacles, walls and ceilings.

The results, summarised in Table 14, clearly indicate that all tested types of RF-based localisation solutions degrade significantly in these environments. All the signals have a lot of reflections with the metal construction, causing a lot of multipath effects. This indicates that accurate indoor localisation in industrial open environments is a difficult task. In this industrial environment, the ToA and RSSI based ranging solutions using the particle filter contains the best results. An average error distance of 6–7 m instead of 11 m for fingerprinting and 17 m for the hybrid technique.

# 9 Discussion

Over the course of the performed experiments, several lessons were learned. First of all, the experiments clearly show the importance of choosing representative measurement locations. Several locations have consistent lower accuracy results. For example the hallways (which are narrow and as such have very low room accuracy). As such, it is clear that the localisation points should include a representative mix of 'easy to locate positions' and more challenging ones. Ideally, a fine-grained grid-like approach should be used, in which the positioning accuracy is evaluated every X meter. All experiments in this paper use the same evaluation points.

**Table 14** w-iLab.t II testbed: summarised benchmarking results

| *Algorithm* | *Average error* [m] | *Room accuracy* [%] | *Response time* [ms] | *Energy efficiency* [mW] *Mobile* | *Fixed* |
|---|---|---|---|---|---|
| *Particle filter solution* | | | | | |
| Spray RSSI | 6.41 | – | 59,633 | ~105 | ~105 |
| Spray ToA | 6.66 | – | 59,620 | ~105 | ~105 |
| *Fingerprinting solution* | | | | | |
| KL distance | 24.76 | – | ~24 000 | ~7000 | ~500 |
| ED distance | 19.08 | – | ~24 000 | ~7000 | ~500 |
| PH distance | 8.13 | – | ~24 000 | ~7000 | ~500 |
| *Hybrid solution* | | | | | |
| TX Power = 31 | 17.16 | – | 15.7 | ~57.6 | ~47.4 |

Another lesson learned is that the location accuracy differs strongly between different testbeds. Environment specifics (such as metal ceilings) strongly influence propagation behaviour. The highest accuracy was obtained in more 'traditional' brick-wall office scenarios, such as represented by the TWIST testbed. The w-iLab.t I testbed, which has plywood walls and metal ceilings, has a lower accuracy. Finally, w-iLab.t II testbed consists of a fully shielded environment, in which the walls and ceilings are from metal, and contains a number of metal obstructions. Performing localisation in this testbed, i.e., in a confined and strongly reflecting environment, proves to be very challenging.

Thirdly, the experiments and benchmarking results that are executed illustrate the need for evaluating a broad set of the metrics. Although the accuracy of the fingerprinting solutions in office environments is shown to be very good, these solutions require significantly more time to collect beacons for fingerprinting, which strongly influences the response delay. Similarly, the WiFi based fingerprinting localisation solutions perform very well, but have higher energy requirements than the solutions using sensor nodes.

Moreover, the experiments indicate that the performance of localisation solutions strongly depends on several algorithmic and deployment aspects, such as the used technology, the ranging approach, the location estimation approach, post- and preprocessing, anchor positions, etc. Making even minor changes to one of these aspects can have a profound influence on several performance metrics. It was also shown that the internal configuration of the algorithms, such as preprocessing the data (such as removing the 10% highest and lowest outliers) or setting the minimum number of beacons that is collected for location estimation can significantly influence the performance.

As such, to allow objective comparison of localisation solutions, it is clear that independent evaluation procedures should be defined by an impartial third party and that such evaluation procedures should include at least the following aspects:

- definition of a wide set of evaluation metrics
- clear definition of the evaluation environments, in which the results are valid
- an objective method for generating a representative set of evaluation points.

## 10 Conclusion

Although many indoor localisation solutions exist, this paper pointed out that the scientific evaluation methods and praxis for RF based solutions are currently limited in scope. Localisation solutions are evaluated mainly based on location accuracy and are evaluated in a single testbed environment. As a result, it is not clear to which level the results from existing scientific literature can be compared to each other.

To evaluate how these different conditions can influence the localisation performance, three localisation solutions were selected that represent typical approaches for indoor localisation, including multiple technologies (IEEE 802.11 and IEEE 802.15.4), multiple localisation approaches (fingerprinting, time-of-arrival and RSSI-based) and multiple processing methods. To allow objective comparisons, the same evaluation methodology was used to evaluate the performance of these localisation solutions in three different environments: an office environment with brick walls, an office environment with plywood walls and an open environment with metal walls and metal obstacles.

The main conclusion of these experiments were the following.

- Several inherent trade-offs between different metrics have been identified, which are typically ignored when reporting only on the accuracy of the solutions. More specifically, the results show a very clear trade-off between the collected number of measurements (which are directly translated into energy consumption and response delay) and the point-level accuracy.
- The accuracy of localisation solutions depends strongly on the characteristics of the environment. Owing to the presence of concrete and/or brick walls in the office testbeds, different rooms are very diverse in terms of their wireless characteristics. Localisation solutions that exploit this diversity, such as fingerprinting based approaches, obtain the highest accuracy. In contrast, in more industrial-like open environments time-of-arrival solutions performed better. These results show that future scientific literature describing performance results of localisation solutions should include detailed descriptions of the used evaluation environment(s), including information such as propagation characteristics, typical room sizes and a description of the materials used in walls and ceilings.
- We have shown that the choice of evaluation points strongly influence the reported accuracy. As such, papers that use self-selected evaluation points can significantly influence their reported accuracy by artificially selecting those evaluation points that outperform other locations.
- Owing to testbed constraints, one evaluation environment contained evaluation points outside the grid of anchor points. It was shown that this set-up had a negative influence on some of the evaluated solutions (mainly the RSSI-based and ToA solutions) but not on the fingerprinting solution. As such, when considering which is the best localisation solution for an industrial deployment, building layout constraints should also be included.
- The accuracy can decrease significantly when evaluating an environment for which the localisation solution is not specifically tweaked. For example, all tested solutions suffered from degraded accuracy in the open industrial-like environment, up to a factor 10 lower. Since most existing solutions have been optimised for office environments, these results hint that many existing localisation solutions might not be ready for use in industrial environments or other challenging environments, such as underground mines.

The above findings reveal several weaknesses in the evaluation methods used in the majority of existing scientific literature of indoor localisation solutions. As such, there is a clear need for a standardised evaluation methodology to objectively compare different localisation solutions in multiple conditions, as developed within the EVARILOS project and pursued by standardisation agencies like ISO.

## Acknowledgements

The research leading to these results has received funding from the European Union's Seventh Framework Program (FP7/2007-2013) under grant agreement no 317989 (STREP EVARILOS). The author Filip Lemic was partially supported by DAAD (German Academic Exchange Service).

## References

Bulusu, N., Heidemann, J. and Estrin, D. (2000) 'GPS-less low-cost outdoor localization for very small devices', *Personal Communications*, IEEE, Vol. 7, No. 5, pp.28–34.

Chambers, J.M., James, D.A., Lambert, D. and Wiel, S.V. (2006) 'Monitoring networked applications with incremental quantile estimation', *Statistical Science*, Vol. 21.4, pp.463–475.

Chintalapudi, K., Padmanabha Iyer, A. and Padmanabhan, V.N. (2010) 'Indoor localization without the pain', *Proceedings of the Sixteenth Annual International Conference on Mobile Computing and Networking*, ACM, September, pp.173–184.

Dezhong, Y., Chen, Y., Anind, K.D., Christian, K., Geyong, M., Laurence, T.Y. and Hai, J. (2014) 'Energy efficient indoor tracking on smartphones', *Future Generation Computer Systems*, Vol. 39, pp.44–54.

Elson, J. and Römer, K. (2003) 'Wireless sensor networks: a new regime for time synchronization', *ACM SIGCOMM CCR*, Vol. 33, No. 1, pp.149–154

Handziski, V., Köpke, A., Willig, A. and Wolisz, A. (2006) 'TWIST: a scalable and reconfigurable testbed for wireless indoor experiments with sensor network', *Proceedings of the 2nd International Workshop on Multi-Hop Ad Hoc Networks: From Theory to Reality*, ACM, May, pp.63–70.

Hightower, J. and Borriello, G. (2001) 'Location systems for ubiquitous computing', *Computer*, Vol. 8, pp.57–66.

Ladd, A.M., Bekris, K.E., Rudys, A.P., Wallach, D.S. and Kavraki, L.E. (2004) 'On the feasibility of using wireless ethernet for indoor localization', *IEEE Transactions on Robotics and Automation*, Vol. 20, No. 3, June, pp.555–559.

Lemic, F., Handziski, V., Behboodi, A., Wolisz, A., Wirström, N., Voigt, T., Van Haute, T. and De Poorter, E. (2014a) *D2.3: Report on Experiments with Controlled Interference*, The EVARILOS Project. EVA-D23-TUB-V1.1, August.

Lemic, F., Büsch, J., Chwalisz, M., Handziski, V. and Wolisz, A. (2014b) 'Demo abstract: Testbed infrastructure for benchmarking rf-based indoor localization solutions under controlled interference', *Proc. of 11th European Conference on Wireless Sensor Networks (EWSN'14)*, February, pp.1–5.

Lemic, F., Büsch, J., Chwalisz, M., Handziski, V. and Wolisz, A. (2014c) 'Infrastructure for benchmarking RF-based indoor localization under controlled interference', *Ubiquitous Positioning Indoor Navigation and Location Based Service (UPINLBS), 2014*, IEEE, November, pp.26–35.

Lemic, F., Behboodi, A., Handziski, V. and Wolisz, A. (2014d) 'Experimental decomposition of the performance of fingerprinting-based localization algorithms', *Proc. of 5th International Conference on Indoor Positioning and Indoor Navigation (IPIN'14)*, October, Busan, South Korea.

Lemic, F. (2014) *Benchmarking of Quantile-based Indoor Fingerprinting Algorithm*, TKN Technical Report TKN-14-001.

Liu, H., Darabi, H. and Banerjee, P. (2007) 'Survey of wireless indoor positioning techniques and systems', *IEEE Transactions on Man, and Cybernetics, Part C: Applications and Reviews*, Vol. 37, No. 6, pp.1067–1080.

Martin, E., Vinyals, O., Friedland, G. and Bajcsy, R. (2010) 'Precise indoor localization using smart phone', *MM '10 Proceedings of the International Conference on Multimedia*, pp.787–790.

Mazomenos, E.B., De Jager, D., Reeve, J.S. and White, N.M. (2011) 'A two-way time of flight ranging scheme for wireless sensor networks', *Wireless Sensor Networks*, Springer Berlin Heidelberg, pp.163–178.

Milioris, D., Tzagkarakis, G., Papakonstantinou, A., Papadopouli, M. and Tsakalides, P. (2014) 'Low-dimensional signal-strength fingerprint-based positioning in wireless LANs', *Ad Hoc Networks*, Vol. 12, pp.100–114.

Otsason, V., Varshavsky, A., LaMarca, A. and de Lara, E. (2005) 'Accurate GSM indoor localization', *UbiComp 2005: Ubiquitous Computing*, pp.141–158.

Pahlavan, K., Li, X. and Makela, J. (2002) 'Indoor geolocation science and technology', *IEEE Commun. Mag.*, Vol. 40, No. 2, pp.112–118.

Pettinato, P., Wirström, N., Eriksson, J. and Voigt T. (2012) 'Multi-channel two-way time of flight sensor network ranging', *Proceedings of the 9th European Conference on Wireless Sensor Networks (EWSN'12)*, 978-3-642-28168-6, pp.163–178.

Steiner, C. and Wittneben, A. (2011) 'Efficient training phase for ultrawideband-based location fingerprinting systems', *IEEE Transactions on Signal Processing*, Vol. 59, No. 12, pp.6021–6032.

Stella, M., Russo, M. and Begušić, D. (2014) 'Fingerprinting based localization in heterogeneous wireless networks', *Expert Systems with Applications*, pp.6738–6747.

Sunkyu, W., Seongsu, J., Esmond, M., Linyuan, X., Changsu, C., Muwook, P. and Joon H. (2011) 'Application of WiFi-based indoor positioning system for labor tracking at construction sites: A case study in Guangzhou MTR', *Automation in Construction*, pp.3–13.

Van Haute, T., De Poorter, E., Rossey, J., Moerman, I., Handziski, V., Behboodi, A. and de las Heras, J.J. (2013) *The EVARILOS Benchmarking Handbook: Evaluation of RF-based Indoor Localization Solutions*, EVARILOS Project, Mermat Workshop, pp.1–5.

Van Haute, T., De Poorter, E., Moerman, I., Lemic, F., Handziski, V., Behboodi, A., Wolisz, A., Wirström, N., Voigt, T. and Crombez, P. (2013) *D2.1: EVARILOS Benchmarking Handbook*, The EVARILOS Project, EVA-D21-iMinds-V1.9, 30 April, 2013.

Van Haute, T., De Poorter, E., Lemic, F., Handziski, V., Behboodi, A., Wolisz, A., Wirström, N., Voigt, T. (2014) *D2.2: Report on Experiments without Interference*, The EVARILOS Project, EVA-D22-iMinds-V2.0, 8 August, 2013.

Van Haute, T., Rossey, J., Becue, P., De Poorter, E., Moerman, I., Demeester, P. (2014) 'A hybrid indoor localization solution using a generic architectural framework for sparse distributed wireless sensor networks', *Federated Conference on Computer Science and Information Systems, WSN2014*, pp.1015–1021.

Wirström, N., Misra, P. and Voigt, T. (2014) 'Spray: a multi-modal localization system for stationary sensor network deployment', *11th Annual Conference on Wireless On-demand Network Systems and Services (WONS), 2014*, pp.25–32.

## Notes

[1]The results described in this paper originate from the project and have first been described in the public EVARILOS deliverable D2.2 "Report on experiments without interference" (Van Haute *et al*., 2013c).

[2]The outcome of initial studies on the influence of interference on the localisation solutions evaluated in this paper can be found in EVARILOS deliverable D2.3 "Report on experiments with interference" (Lemic *et al*., 2014a).

[3]ISO/IEC 18305 is being prepared by Joint Technical Committee ISO/IEC JTC 1, Information technology, Subcommittee SC 31, Automatic identification and data capture techniques, Working Group 5, Real time locating systems. The committee is currently referred to as ISO/IEC JTC1/SC31/WG5.

## Websites

EvAAL – Evaluating AAL Systems through Competitive Benchmarking – http://evaal.aaloa.org/index.php?option=com_content&view=article&id=187:technical-annexes-localization2013&catid=15&Itemid=261

ISO Standard – ISO/IEC CD 18305 – Information technology – Real time locating systems – Test and evaluation of localization and tracking systems – http://www.iso.org/iso/home/store/catalogue_htc/catalogue_detail.htm?csnumber=62090

'The EVARILOS project', http://www.evarilos.eu