

# Platform for Benchmarking of RF-Based Indoor Localization Solutions

Tom Van Haute, Eli De Poorter, Filip Lemic, Vlado Handziski, Niklas Wirström, Thiemo Voigt, Adam Wolisz, and Ingrid Moerman

## ABSTRACT

Over the last few years, the number of indoor localization solutions has grown exponentially, and a wide variety of different technologies and approaches are being explored. Unfortunately, there is currently no established standardized evaluation method for comparing their performance. As a result, each solution is evaluated in a different environment using proprietary evaluation metrics. Consequently, it is currently extremely hard to objectively compare the performance of multiple localization solutions with each other. To address the problem, we present the EVARILOS Benchmarking Platform, which enables automated evaluation and comparison of multiple solutions in different environments using multiple evaluation metrics. We propose a testbed-independent benchmarking platform, combined with multiple testbed-dependent plug-ins for executing experiments and storing performance results. The platform implements the standardized evaluation method described in the EVARILOS Benchmarking Handbook, which is aligned with the upcoming ISO/IEC 18305 standard “Test and Evaluation of Localization and Tracking Systems.” The platform and plug-ins can be used in real time on existing wireless testbed facilities, while also supporting a remote offline evaluation method using precollected data traces. Using these facilities, and analyzing and comparing the performance of three different localization solutions, we demonstrate the need for objective evaluation methods that consider multiple evaluation criteria in different environments.

Tom Van Haute, Eli De Poorter, and Ingrid Moerman are with Ghent University.

Filip Lemic and Vlado Handziski are with Technische Universität Berlin.

Niklas Wirström and Thiemo Voigt are with the Swedish Institute of Computer Science.

<sup>1</sup> <http://www.ict-fire.eu/home.html>.

## INTRODUCTION

This article addresses one of the major problems of indoor localization research: the lack of comparability between existing localization solutions, due to the fact that most of them have been evaluated under individual, thus not comparable and not repeatable, conditions. This situation is partially the result of the complexity required for the evaluation of an indoor localization solution, which requires technical expertise to efficiently set up large-scale experiments, control the exper-

imental environment, gather the necessary performance data, and calculate the output metrics using standardized methods. All these steps are time consuming, and more theoretically inclined researchers typically lack the necessary technical skills to perform these steps efficiently and accurately. We address these deficiencies by providing a platform that allows simple evaluation of indoor localization solutions. The main contributions of the presented article are as follows.

We describe a generic benchmarking platform that implements the standardized evaluation method described in the EVARILOS Benchmarking Handbook (EBH), and is aligned with the upcoming International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) 18305 standard “Test and Evaluation of Localization and Tracking Systems.”

We further describe plug-ins that are available for instantiating the components of the EVARILOS Benchmarking Platform on multiple future Internet research and experimentation (FIRE) facilities.<sup>1</sup>

Finally, we provide open datasets that help in simplifying the process of benchmarking and evaluation of indoor localization solutions.

The rest of this article is structured as follows. The next section provides an overview of the related work. Then the EVARILOS Benchmarking Platform (EBP) is explained in detail. The integration of the EBP in a wireless test facility and the public datasets are then discussed. We then demonstrate the usage of the EBP in an experimental validation of multiple RF-based indoor localization solutions. Finally, we conclude the work.

## RELATED WORK

As the number of indoor localization solutions is growing, a more thorough procedure for evaluating and comparing them is necessary. As already observed in other fields [1], a well defined objective evaluation methodology needs to take into consideration a wide range of metrics. Some metrics are important from a theoretical point of view, and are well suited for analyzing and improving proposed algorithms, whereas others

focus on the performance of end solutions, and are more important for industry and end users. If only accuracy is taken into account, the results can give a distorted view. Such considerations have motivated M. Ficco *et al.* [2] to evaluate indoor localization solutions with respect to deployment metrics. They compare and calibrate the deployment and usage of access points (APs), and show that the quality of the radiomap has a direct influence on the accuracy. Furthermore, Hui Liu *et al.* state in [3] that precision, complexity, scalability, robustness, and cost should be included if a comprehensive performance analysis is required. Additionally, they also recognize the lack of an objective methodology for the evaluation of indoor localization solutions. Motivated by these circumstances, a number of organizations are trying to develop comprehensive standardized evaluation approaches for indoor localization solutions.

**EVARILOS Project:** In the scope of the FP7 EVARILOS project, focused on objective evaluation of RF-based indoor localization solutions, the EBH [4] has been published. The handbook describes a set of evaluation metrics that are important for the evaluation of indoor localization, including different notions of accuracy, functional metrics such as response delays, and deployment metrics such as setup time and required infrastructure. Furthermore, the handbook contains a set of scenarios that describe how to adequately evaluate an indoor localization solution. The project is also the first one to systematically address the effect of interference on indoor localization solutions, although interference is expected to be present at most sites where these solutions are deployed. The EBH includes a wide range of evaluation metrics, including functional metrics, such as response delays, and deployment metrics, such as setup time and required infrastructure.

**ISO:** Recently, the ISO and IEC established a joint technical committee, ISO/IEC JTC 1, focused on proposing a new ISO/IEC 18305 standard, “Test and Evaluation of Localization and Tracking Systems.”<sup>2</sup> Current drafts include evaluation methodologies for a single technology (e.g., Bluetooth), as well as methodologies for the evaluation of full localization solutions, which is in line with the methodology proposed in the EVARILOS project. While this effort is more general in that it also pertains to a wide range of non-RF-based technologies such as motion sensors, thus far it does not include non-accuracy-related metrics such as ease of use or energy consumption. At the time of writing, none of the drafts were publicly available.

**EvAAL:** Until now, the only attempts at direct comparison of different indoor localization solutions were indoor localization competitions. One popular series of indoor localization competitions has been organized by Microsoft as part of the Information Processing in Sensor Networks (IPSN) conference. During the 2014 edition of the competition [5], 22 different indoor localization solutions were evaluated (organized in two categories: infrastructure-free and infrastructure-

based). The evaluation process uses only a single metric: average localization error across 20 test points. The errors are measured manually using a handheld laser distance meter. In 2015, the evaluation process for the 23 competing solutions took more than one day. In 2014 we shadowed the official evaluation process using the EBP presented in this article, and demonstrated the viability and the benefits of a full automation of this process. The Evaluating AAL Systems through Competitive Benchmarking (EvAAL) project<sup>3</sup> uses a set of metrics as part of the evaluation process for its competition series. In addition to the accuracy of indoor localization, usability metrics are defined such as installation complexity, user acceptance, availability, and interoperability with AAL systems. The evaluation process is not automated, and involves deploying physical devices in the environment of interest.

Most scientific papers evaluate the solution they propose in an easily accessible environment in the development area of the authors. Typically, these are office environments with brick walls [6, 7]. Since evaluation is rather time consuming, most localization solutions are evaluated only in a single environment. Both the EVARILOS project and ISO/IEC JTC 1 refer to the fact that this evaluation is not representative for other environments. Therefore, our platform offers developers the possibility to evaluate their localization solutions using input datasets collected in multiple environments: an office environment with brick walls, an office environment with plywood walls, and finally, an industrial-like open-space environment. Since the accuracy strongly depends on the used evaluation points, for example, points near a wall vs. in the middle of a room or in an open space, our public datasets contain data measured at a wide range of measurement points.

## EVARILOS BENCHMARKING PLATFORM

This section describes the EBP.<sup>4</sup> The EBP has been created to address the fact that, although numerous experimental testbed facilities are available [8, 9], evaluating the performance of a localization solution under controlled conditions using standardized performance metrics has proven to be very complicated, in particular for researchers who have limited experience with experimental research. The EBP addresses this issue by providing an open software solution that implements user friendly methods to support the full performance evaluation cycle. The developed software components are independent of any experimental facilities and use open source principles, allowing researchers to download and modify any of the components.

An overview of the EBP architecture is shown in Fig. 1.

**The Rectangles:** Represent components that are available as web services. These components run on a cloud platform where they can be accessed remotely or downloaded to be modified and/or run locally.

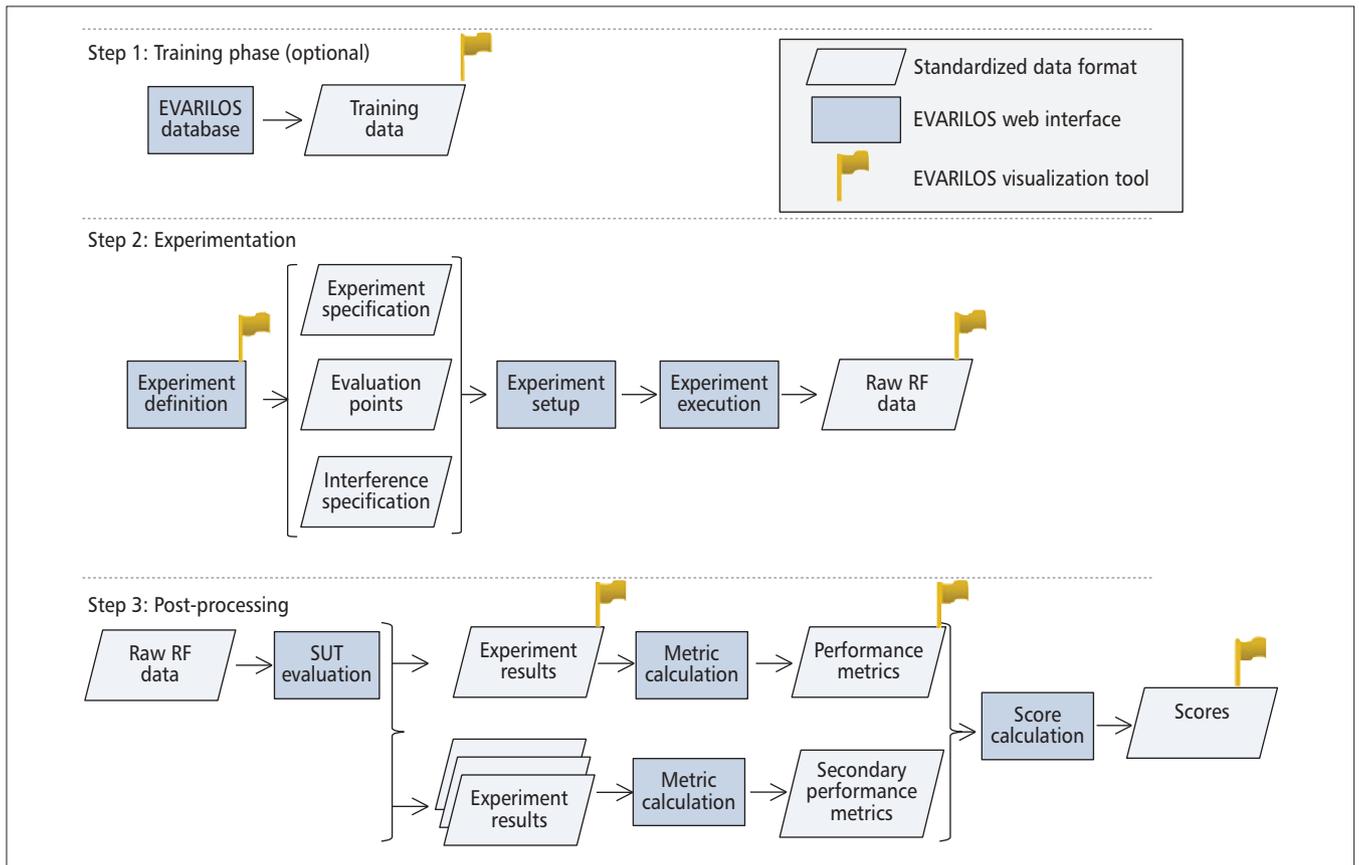
**The Parallelograms:** Represent data struc-

*Our platform offers developers the possibility to evaluate their localization solutions using input datasets collected in multiple environments: an office environment with brick walls, an office environment with plywood walls, and finally, an industrial-like open-space environment.*

<sup>2</sup> <http://www.iso.org>

<sup>3</sup> <http://evaal.aalooa.org>

<sup>4</sup> <http://ebp.evarilos.eu/>



**Figure 1.** Overview of the components of the EBP and the data structures used to exchange information between the components.

tures that are used to exchange data between the web services.

**The Flags:** Represent the tools that can be used to analyze and visualize the different steps of the process.

The architecture consists of a set of components that, when used sequentially, implement a workflow which represents three experimentation steps. A summary can be found below, while in the next subsections each step is discussed in detail.

**Pre-Experimentation Phase:** During a pre-experimentation phase, users can download environment-specific training datasets from public repositories. These datasets are typically used for training the localization solution.

**Experimentation Phase:** In the experimentation phase, all the components required for the experimentation are orchestrated, and the experiments are executed. The platform offers the possibility for automated generation of experiment configurations, including specifications of the used evaluation points, the interference patterns that will be generated, and so on. Based on these descriptions, experiment executables are created using testbed-specific tools with the Control and Management Framework (OMF),<sup>5</sup> which is used in many recent wireless testbeds,<sup>6</sup> and are automatically executed. Note that this step can be omitted if the next step utilizes pre-collected input (e.g., WiFi beacons) for a localization solution.

**SUT:** Finally, the environmental RF data is fed to the system under test (SUT), either in

real-time or using precollected measurements, depending on the experiment configuration. The estimated locations are stored together with additional performance metrics such as the response delay. It is also possible to combine results from multiple experiments to observe how certain evaluation metrics evolve.

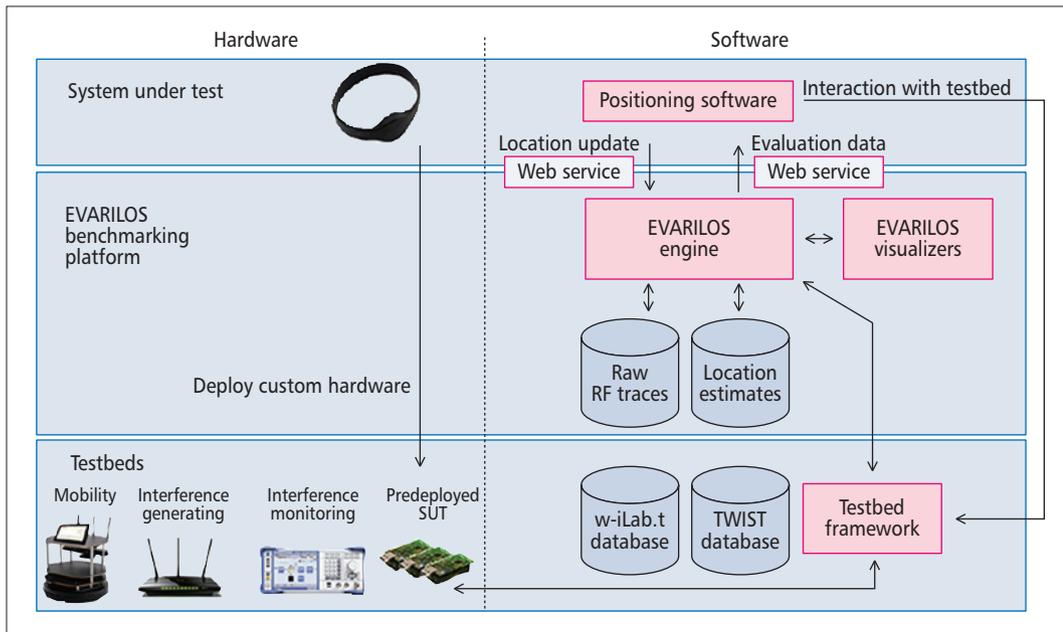
### TRAINING PHASE

The training phase offers experimenters the possibility to train their localization solutions based on measurements that are performed in advance on a representative location. The measurements currently offered represent raw data that can be used as input into an RF-based indoor localization solution, such as received signal strength indicator (RSSI), link quality indicator (LQI), or time of arrival (ToA). Measurements for training purposes are captured in an area that is representative for the experimentation phase. Typically, the data is captured in the same environment where the SUT will be evaluated. To prevent aliasing problems, the training data should not exactly correspond to the data that is used during the evaluation phase. Otherwise, the performance evaluation of step 2 of the evaluation process will be biased. To this end, users can use data that is:

- Captured at a different time
- Captured using devices from a different manufacturer
- Captured at evaluation points other than the one used during the performance evaluation

<sup>5</sup> <http://omf.mytestbed.net/projects/omf6/wiki/Wik>

<sup>6</sup> <http://mytestbed.net/projects/omf/wiki/DeploymentSite>



**Figure 2.** Deployment of the EBP.

The platform offers researchers a database to access previously measured environmental information relevant for their localization solution. Users can either download the data directly from the EVARILOS data repository or can access an EVARILOS application programming interface (API) that encapsulates the data and can serve the data at a finer granularity.

### EXPERIMENTATION PHASE

The experimentation phase offers experimenters the possibility to define setups for raw RF data collection or full localization experiments in FIRE facilities, as well as an interface for automatic execution. The user will start with an “experiment definition” (Fig. 1). The role of the experiment definition component is to configure all aspects of the experiment that will be used to evaluate a SUT. To this end, the experiment definition component requires the following input: the *experiment specification* (e.g., which nodes will be used as anchor points, when will the experiment be scheduled, which binary files to use), the *evaluation points* (at which locations is a SUT evaluated), and the *type of (artificial) interference* that should be generated. To assist with this process, a fully automated web service is available, where users can select among different preconfigured options. Of course, it is possible to modify any of the default settings to adjust the experiment behavior. This information is also stored in a standardized data format.

Next, the “experiment creation” component is executed, which is a fully automated step, whereby the testbed-independent information is translated into testbed-dependent executables using the appropriate plug-ins. The final step is the actual execution of an experiment. In this step, the executables are executed on the corresponding testbed, and the result of the execution is stored in an appropriate data structure together with additional metadata, describing a whole experiment in detail. The result of the execution

is raw data, such as WiFi or IEEE 802.15.4 beacon information, which is collected by a SUT at different locations in an environment.

### POST-PROCESSING PHASE

In this step the obtained raw data traces can be fed to the evaluated SUT, and location estimates can be produced. Furthermore, the metrics characterizing the performance of a SUT can be calculated. The experiment results are stored in an appropriate data structure, which consists of a set of ground truths and estimates for different measurement locations and a set of metrics characterizing the performance of a SUT for a given experiment.

Experiment results from multiple experiments can be combined to observe how certain evaluation metrics evolve, for example, for different scenarios or different parametrization of a SUT. These results are stored in a secondary metrics data structure. For comparability purposes, a final score can be assigned to the performance of each SUT. This score is an abstraction of the performance of a SUT in a specific environment and necessarily hides many intrinsic trade-offs. Finally, it is worth mentioning that the full post-processing phase can also be applied to location estimates from non-EBP-compliant solutions. As long as the experiment results are provided in the correct data format, the same tools can be used to analyze and rank the outcome of any localization solution.

## INTEGRATION OF EBP IN WIRELESS EXPERIMENTATION FACILITIES

The EBP is designed to simplify the evaluation of RF-based localization solutions. The components of the platform can be used “as is” by utilizing precollected data traces as input. However, as already mentioned, the platform components can also be used to *facilitate the evaluation of*

*It is worth mentioning that the full post-processing phase can also be applied to location estimates from non-EBP-compliant solutions. As long as the experiment results are provided in the correct data format, the same tools can be used to analyze and rank the outcome of any localization solution.*



**Figure 3.** Two examples of the testbeds (w-iLab.t I and II) where experiments can be executed.

*localization solutions in new environments.* The available deployment options for indoor localization benchmarking are presented in Fig. 2. Three main components can be identified:

- The bottom layer represents a wireless experimentation facility or testbed. The testbed-specific tools are installed on a server in a given test facility.
- The EBP includes services that facilitate testbed-independent definition of experimentation and the evaluation of localization solutions.
- Finally, the upper layer represents a SUT, which can include both hardware and/or software components.

As mentioned, the EBP is integrated in existing FIRE facilities. This integration is part of the experiment execution component illustrated in Fig. 1. Automatic conversion from experiment descriptions to testbed-dependent scripts is supported, thereby integrating and simplifying the complex steps that otherwise need to be taken for objective experimentation. Building on top of the CREW Cognitive Radio testbeds,<sup>7</sup> the infrastructure leverages a robotic mobility platform, which serves as a reference localization system and can transport the localized device in an autonomous and repeatable manner. In addition, the platform uses the capabilities of the CREW testbed infrastructure to generate typical interference scenarios in a reproducible manner. This further improves benchmarking of indoor localization solutions by testing the performance of a SUT under realistic and repeatable interference conditions.

The interaction between a SUT and the EBP is designed to be as simple as possible: at most two REST interfaces [10, 11] are required, depending on the requirements of an experiment. One interface provides location estimates and ground truth information to the EBP, and the other stores the raw data from a SUT or uses the precollected raw data as input to a SUT.

**During an Experiment:** The EBP can issue a request for the location estimate from a SUT through the first REST interface. As such, the minimum requirement for a SUT to comply with the EBP is to provide the location estimate over HTTP upon request.

**The EBP Can Also Request the Real-Time Environmental Data:** (RSSI values, ToA, etc.) from a SUT, which is then stored through a second REST interface. This data can be collected and at a later time be offered to future experimenters as an open data set.

This architecture allows experimenters to choose among different utilization options.

**Option 1:** The evaluation of a localization algorithm using precollected raw data traces that can be used as input to a SUT. In this scenario, the localization algorithms can be evaluated remotely using the EBP.

**Option 2:** The evaluation of a localization solution using software running on an existing wireless testbed. In this scenario, the localization algorithms can run on local hardware that is available at the experimentation facilities.

**Option 3:** The evaluation of localization hardware using a testbed. In this scenario, experimenters can install custom hardware at the experimentation facility while still using the EBP for the evaluation of their solution.

One of the major advantages of the EBP is that all three approaches make use of the same common components.

The feasibility of these options has been demonstrated through the EVARILOS Open Challenge [12], as well as during the Microsoft Indoor Localization Competition (IPSN 2014) [5].

## PUBLIC DATASETS

One of the features of the EBP is the capability to reuse previously collected RF data for offline evaluation of RF-based indoor localization solutions. This feature addresses one of the important challenges for the indoor localization research community: the complex and expensive process of obtaining relevant measurements of RF features from multiple environments. The EBP offers a wide range of available precollected RF data sources through its user interface. However, for those researchers who prefer to download full annotated datasets, the EBP also offers the possibility to download the datasets for research purposes. Two types of datasets are currently available: raw RF traces and performance information.

### RAW RF TRACES

Environmental RF data can be used as a basis for either training an algorithm (e.g., by creating propagation models) or offline evaluation of a SUT. The EBP makes available the measured raw RF traces from multiple environments, including a plywood office environment (w-iLab.t I [8]), a brick office environment (TWIST [13]), an industrial-like environment (w-iLab.t II [8]), a hospital environment, and an underground mine. A view of w-iLab.t I and II is available in Fig. 3. The details about the structure of the raw RF data, exact descriptions of the currently available datasets, and an overview of the services available for using the raw RF data for the evaluation of RF-based indoor localization algorithms can be found in [14].

To evaluate a solution for a wide range of conditions, the raw RF traces contain significantly more data than would be used in a typical operational environment. The datasets are rich in terms of number of collected samples per evaluation point (over 1000 samples per evaluation point), the captured data types (including WiFi beacons, sensor RSSI, and sensor time-of-flight information), the used configuration set-

<sup>7</sup> <http://www.crew-project.eu/>

tings (multiple frequencies, multiple transmission powers), and the used anchor points (data is collected from up to 60 anchor points per evaluation point). This richness of the dataset makes the data relevant for a wide range of interested researchers and allows investigation of how changing any of these parameters influences the performance of the solution. Transforming the over-dimensional dataset into a set that is more sparse (and more realistic from an operational point of view) can easily be done by removing any unnecessary information (sub-sampling). In addition, the available environment data is annotated with metadata describing the exact conditions in which the data was captured. This metadata describes characteristics including the used hardware, type of collected raw data, timestamps, measurement frequency, environment description, and so on.

### PERFORMANCE INFORMATION

The EBP gives a ranked overview of evaluated solutions on its web page. However, these performance indicators necessarily hide a number of low-level statistics. Researchers interested in also evaluating the temporal or spatial behavior of different solutions can analyze the performance datasets. EBP makes available the results from its own localization solutions, as well as of those solutions that participated in the EVARILLOS Open Challenge [12]. Each of these datasets also has its associated experiment configuration settings, allowing detailed analysis not only of the performance but also of the conditions in which the solutions were evaluated.

## EXPERIMENTAL VALIDATION

In [14] we illustrate the benefits of leveraging the presented platform for the evaluation of RF-based indoor localization, in terms of time and complexity of usage, in comparison to using an infrastructure or performing a manual evaluation. In the following we demonstrate the need for a standardized evaluation method by showing that the performance of localization solutions depends strongly on its parametrization and can only be done objectively by considering multiple evaluation metrics.

### THREE INDOOR LOCALIZATION SOLUTIONS

In order to develop, test, and optimize our platform, three different types of indoor localization solutions were used as SUTs. The basic concept behind the first localization solution [15] is the following: measurements are performed by requesting a stationary node to transmit packets to the testbed nodes that then reply with a hardware acknowledgment (ACK). The initiating node measures the time between transmission of the packet and reception of the ACK, and stores the RSSI values associated with the ACK. These measurements are then processed using Spray, a particle-filter-based platform [15]. The basic idea of the ToF ranging is to estimate the distance between two nodes by measuring the propagation time, which is linearly correlated to the distance when the nodes are in the line of sight (LoS).

A second solution [16] is based on fingerprinting. Fingerprinting methods for indoor

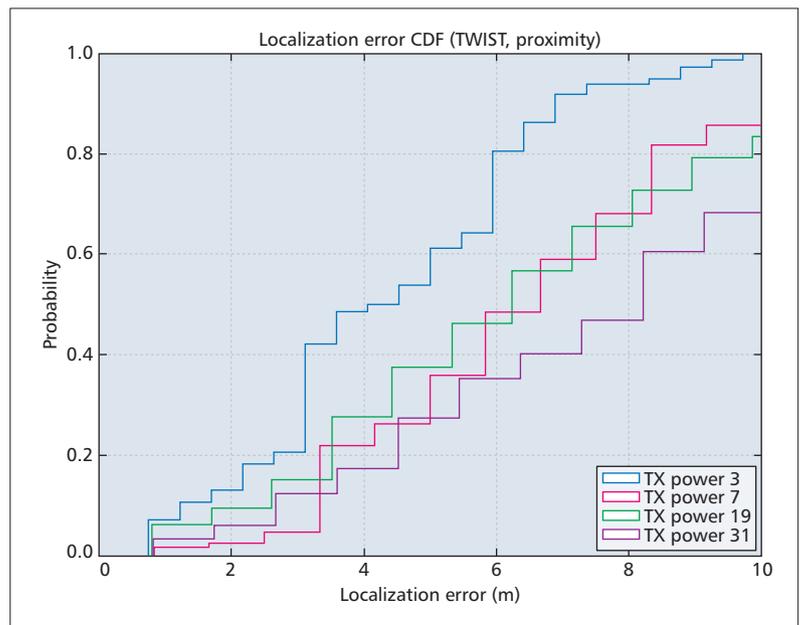


Figure 4. CDFs for the hybrid solution in the TWIST testbed.

localization are generally divided in two phases. The first phase is called training or offline phase. In this phase, the localization 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 representative fingerprint of each cell is created. Using the obtained training fingerprints, the training database is created and stored on a localization 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 methodology, a runtime fingerprint is created and sent to the localization server. At the server's side, the runtime fingerprint is compared to the training dataset using a matching method. The training fingerprint with the most similarities to the runtime fingerprint is reported as the estimated position.

A third localization solution [17] that has been implemented and evaluated is a hybrid combination range-based and range-free algorithm. It includes a range-based location estimator based on weighted RSSI values. Each RSSI value can be matched with a certain distance. The proposed algorithm in [17] not only uses the RSSI values to measure the distance between a fixed node and a mobile node, but also the distance between fixed nodes. These values function as weight factors for the distance calculation between the fixed and mobile nodes. Once the distances are known, triangulation can be applied in order to determine the final position of the person/object that needs to be localized. This approach is combined with a range-free algorithm, which does not take RSSI-values into account. If a mobile sensor node has a range of 10 m, a fixed node can only receive its 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 fixed nodes. 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.

### ANALYSIS OF A SINGLE SOLUTION

An important feature of the EBP is its capability to streamline the process of obtaining better insight on the evaluated localization solution. Every solution contains a set of adjustable parameters, which can considerably influence the overall performance, implying that optimizing this set of parameters can be a hard task. Therefore, the EBP can easily compare the same solution using multiple values of a single parameter.

This can be demonstrated with an example. The hybrid solution [17] described in the section above states that the transmission power is an important value that needs to be configured well

in order to receive acceptable results. Therefore, the solution was evaluated using the EBP using multiple transmission powers, the outcome of which is shown using a cumulative distribution function (CDF) (Fig. 4) and a table with multiple metrics (Table 1). Based on these results, it is clear that this solution obtains the lowest average error when the transmission power equals three. But it also illustrates inherent trade-offs that are present in the solution: suppose the response time is the most important criteria; then a transmission power of 31 would be the best option. This example illustrates the advantages of the EBP for fast and efficient identification of an optimal operating point depending on adjustable parameters, and demonstrates the need for considering multiple metrics to identify trade-offs.

### COMPARISON BETWEEN MULTIPLE SOLUTIONS

Table 2 compares the performance of three different solutions evaluated using the EBP by considering multiple evaluation criteria. By utilizing the same evaluation points, objective comparisons are possible. Again, the results illustrate the presence of trade-offs that can only be observed by comparing multiple metrics. More specifically, it demonstrates that the approach taken in most current scientific papers, wherein point accuracy is considered as the only relevant metric, fails to take into account the associated costs in response time and energy consumption.

## CONCLUSION

The proliferation of RF-based indoor localization solutions raises the need for testing systems that enable objective evaluation of their functional and non-functional properties. Although a significant number of localization solutions are available, different approaches are used for the evaluation of these solutions in terms of used performance metrics and evaluation methodology. This article tries to address these shortcomings by providing tools for evaluating and comparing localization solutions using standardized evaluation methods, as described in the EBH.

We introduce a testbed-independent benchmarking platform for automatized benchmarking of RF-based indoor localization solutions. Using a well defined interface, the infrastructure obtains location estimates from the SUT, which are subsequently processed in a dedicated metrics computation engine. The components can be accessed through web services that are available for external users or can be downloaded for custom modifications. The benchmarking platform has proven to be useful for locations where no testbed facilities are available. Multiple components of the platform were extensively used during the Microsoft Indoor Localization Competition (IPSN 2014) as well as the EVARILOS Open Challenge. In these events, the components of the benchmarking platform improve the time efficiency and ease of use of the experiments, and also resulted in more objective comparability.

Finally, to accommodate the need for wider accessibility of experimental data, open datasets are provided. These datasets include both annotated localization data from multiple environments, as well as detailed descriptions of the

Metric	TX 3	TX 7	TX 19	TX 31
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
Room accuracy (%)	26.67	6.70	13.45	9.56
Response time (ms)	1503	1507	480	460

**Table 1.** Statistical information about the performance of the hybrid solution in the TWIST testbed.

Algorithm	Mean error (m)	Room acc. (%)	Latency (ms)	Energy eff. (mW)	
				Mobile	Fixed
<b>Particle filter solution</b>					
Using RSSI	4.35	45.00	14,285	~105	~105
Using ToA	5.56	30.00	14,282	~105	~105
<b>Fingerprinting solution</b>					
Using ED distance	2.2	80.0	~35,000	~7000	~500
Using PH distance	2.0	85.0	~35,000	~7000	~500
<b>Hybrid solution</b>					
TX Power = 3	4.6	26.7	1503	~30.9	~47.4
TX Power = 7	7.1	6.7	1507	~35.1	~47.4

**Table 2.** TWIST testbed: summarized results.

setup and outcome of the performed localization experiments from earlier experiments. These repositories can be used to quickly evaluate a SUT in different environments, analyze the effects of changing configuration settings, analyze the setup of different experiments, and compare the performance of a wide range of localization solutions.

### ACKNOWLEDGMENT

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

- [1] M. Seltzer et al., "The Case for Application-Specific Benchmarking," *Proc. 7th Wksp. Hot Topics in Op. Sys.*, 1999.
- [2] A. N. M. F. C. Esposito, "Calibrating Indoor Positioning Systems with Low Efforts," *IEEE Trans. Mobile Computing*, vol. 13, no. 4, 2014.
- [3] H. Liu et al., "Survey of Wireless Indoor Positioning Techniques and Systems," *IEEE Trans. Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 37, no. 6, 2007.
- [4] T. V. Haute et al., "The EVARILOS Benchmarking Handbook: Evaluation of RF-based Indoor Localization Solutions," *MERMAT 2013*, May 2013.
- [5] D. Lymberopoulos et al., "A Realistic Evaluation and Comparison of Indoor Location Technologies: Experiences and Lessons Learned," *IPSN '15*, 2015.
- [6] K. Chintalapudi et al., "Indoor Localization without the Pain," *Proc. 16th Int'l. Conf. Mobile Computing and Networking*, ACM, 2010.
- [7] E. Martin et al., "Precise Indoor Localization Using Smartphones," *Proc. Int'l. Conf. Multimedia*, ACM, 2010.
- [8] S. Bouckaert et al., "The w-ilab. t testbed," *Testbeds and Research Infrastructures, Development of Networks and Communities*, Springer, 2011.
- [9] F. Lemic et al., "Infrastructure for Benchmarking RF-Based Indoor Localization under Controlled Interference," *Proc. UPINLBS'14*, 2014.
- [10] F. Lemic, "Service for Calculation of Performance Metrics of Indoor Localization Benchmarking Experiments," tech. rep. TKN-14-003, 2014.
- [11] F. Lemic and V. Handziski, "Data Management Services for Evaluation of RF-Based Indoor Localization," tech. rep. TKN-14-002, 2014.
- [12] F. Lemic et al., "Experimental Evaluation of RF-Based Indoor Localization Algorithms under RF Interference," *Proc. ICL-GNSS'15*, 2015.
- [13] V. Handziski et al., "TWIST: A Scalable and Reconfigurable Testbed for Wireless Indoor Experiments with Sensor Network," *Proc. RealMAN'06*, 2006.
- [14] F. Lemic et al., "Web-Based Platform for Evaluation of RF-Based Indoor Localization Algorithms," *Proc. IEEE ICC Wksp.*, 2015.
- [15] N. Wirstrom, P. Misra, and T. Voigt, "Spray: A Multi-Modal Localization System for Stationary Sensor Network Deployment," *Proc. IEEE Wireless On-Demand Network Systems and Services*, 2014.
- [16] F. Lemic, "Benchmarking of Quantile based Indoor Fingerprinting Algorithm," Tech. Rep. TKN-14-001, 2014.
- [17] T. Van Haute et al., "A Hybrid Indoor Localization Solution Using a Generic Architectural Framework for Sparse Distributed Wireless Sensor Networks," *Proc. IEEE Comp. Sci. and Info. Systems*, 2014.

### BIOGRAPHIES

TOM VAN HAUTE is a doctoral researcher at Ghent University. He received his M. Sc. degree (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 in the Internet Based Communication Networks and Services research group (IBCN). His research is focused on wireless sensor networks combined with indoor localization and indoor navigation in particular.

ELI DE POORTER is a postdoctoral researcher at Ghent University. He received his Master's degree in computer science engineering from Ghent University, Belgium, in 2006. He received his Ph.D. degree in 2011 from the Department of Information Technology at Ghent University through a Ph.D scholarship from the Institute for Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen). After obtaining his Ph.D., he received an FWO postdoctoral research grant and is now a postdoctoral fellow in the same research group.

FILIP LEMIC is a junior researcher and Ph.D. candidate in the Telecommunication Networks Group at the Technical University of Berlin. He finished his Bachelor's and Master's studies with the Faculty of Electrical Engineering and Computing at the University of Zagreb. His main scientific interests are in context awareness, with an emphasis on indoor localization.

VLADO HANDZISKI is a senior researcher in the Telecommunication Networks Group at Technische Universität Berlin, where he coordinates the activities in the areas of sensor networks, cyber-physical systems, and the Internet of Things. He is currently also serving as interim professor at the chair for Embedded Systems at Technische Universität Dresden. He received his doctoral degree in electrical engineering from TU Berlin (summa cum laude, 2011) and his M.Sc. degree from Ss. Cyril and Methodius University in Skopje (2002).

NIKLAS WIRSTRÖM is a researcher in the Networked Embedded Systems (NES) group at SICS and a Ph.D. student at Uppsala Universitet, Sweden. His research focus is on machine learning techniques for localization in WSNs and other resource constrained systems.

THIEMO VOIGT is a professor at Uppsala University. He also leads the NES 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.

ADAM WOLISZ received his degrees (Diploma 1972, Ph.D. 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.

INGRID MOERMAN received her degree in electrical engineering (1987) and her Ph.D. degree (1992) from 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 ([www.ibcn.intec.ugent.be](http://www.ibcn.intec.ugent.be)), where she leads the research on mobile and wireless communication networks. In 2006 she joined iMinds, where she coordinates several interdisciplinary research projects.

*The proliferation of RF-based indoor localization solutions raises the need for testing systems that enable objective evaluation of their functional and non-functional properties. Although a significant number of localization solutions are available, different approaches are used for the evaluation of these solutions in terms of used performance metrics and evaluation methodology.*