

Automated Evaluation of RF-based Indoor Localization Systems

Stephan Wagner, Hugues Smeets, Marcus Handte,
Ngewi Fet, Chia-Yen Shih, and Pedro José Marrón

Networked Embedded Systems Group
University of Duisburg-Essen, Germany

{stephan.j.wagner,hugues.smeets,marcus.handte,ngewi.fet,chia-yen.shih,
pjmarron}@uni-due.de

Abstract. An important basis of many smart city applications is knowledge about the location of persons and objects. In outdoor environments, this knowledge can be acquired reliably on a global scale using the well-known Global Positioning System (GPS). In indoor environments, both the availability and reliability of GPS is significantly limited. This, in turn, has led to active research on approaches and systems to enable indoor localization using various cooperating objects technologies. A key challenge during the development of any of these indoor localization approaches and systems is the systematic evaluation of their performance. To do this, developers have to perform extensive and time-consuming measurements at different locations over an extended period of time. In this paper, we discuss how the time requirements can be reduced by means of automation. Furthermore, based on our experiences with both, manual and automatic evaluation, we discuss the achievable benefits and possible limitations.

Keywords: Localization, Cooperating Objects, Evaluation, TrainSense

1 Introduction

An important basis of many smart city applications is knowledge about the location of persons and objects. In outdoor environments, this knowledge can be acquired reliably on a global scale using the well-known Global Positioning System (GPS) [5]. In indoor environments, both the availability and reliability of GPS is typically limited. Over the past decade, researchers and practitioners have spend a significant amount of resources to develop various indoor localization systems and approaches with different technologies. Example technologies include camera systems [8], infrared light [13], (ultra-) sound [3] and a broad spectrum of RF technologies such as Bluetooth [9] or WLAN [1] to name a few. Due to its scalability and wide availability, RF-based indoor localization is considered to be favorable in many application scenarios.

As discussed in depth in [7], existing approaches for RF-based indoor localization can be broadly classified into four categories:

- *Proximity analysis*: Proximity analysis uses connectivity information as a basis for localization. However, to provide a sufficiently high accuracy, proximity analysis requires a comparatively dense deployment of devices. Consequently, it cannot be considered cost effective for large scale deployments.
- *Angulation*: Angulation uses the angle of arrival of a particular signal in order to determine a location. For this, however, it is necessary to use specifically designed antennas that are able to determine the angle and thus, it cannot be realized with off-the-shelf hardware that is readily available.
- *Lateration*: Lateration approaches such as TOA, TDOA or RTOF rely on estimates of the signals flight time in order to determine a location. As a result, they typically require precise synchronization or time measurements. Usually, this results in expensive hardware setups that include special wiring. To avoid such costs, lateration can also be done on the basis of signal strength. However, due to the multi-path effects of most indoor environment, this is likely to cause significant inaccuracy.
- *Scene analysis*: To avoid both, the increased hardware cost of time-based lateration approaches as well as the error caused by distance estimation, it is possible to rely on scene analysis. The basis for scene analysis is calibration in the target environment that is used to overcome the non-linear signal behaviour caused by multi-path effects.

Especially, when considering the popular category of systems based on scene analysis, a key challenge during the development is the systematic evaluation of their performance. To do this, developers have to perform extensive and time-consuming measurements in order to calibrate and test the system. Typically, this involves the repeated positioning of objects at different locations. In addition, to gather meaningful metrics, it is often necessary to repeat the process over an extended period of time. Consequently, researchers have used simplified processes to study the performance of their systems. However, this limits the insights that can be gathered from their results.

In this paper, we discuss how the time required to evaluate the performance of RF-based indoor localization systems can be reduced by means of automation. To do this, we outline the typical evaluation process for such systems in Section 2. Thereafter, in Section 3, we introduce two localization systems that we have built and evaluated as part of WebDA [11], one of our ambient assisted living projects. Based on our experiences with the manual and automated evaluation of these localization systems, we discuss the achievable benefits and possible limitations in Section 4. Finally, we conclude the paper with a summary and an outlook in Section 5.

2 Process and Challenges

Typically, the design of an RF-based indoor localization system starts with the selection of an appropriate RF technology and a suitable localization approach. In the past, researchers have studied a multitude of technologies including mainstream communication technologies such as FM-Radio [10], DECT [6], Bluetooth

[9] or WLAN [1, 14, 4] as well as identification technologies such as RFID [12] or specialized technologies such as UWB [17]. As hinted in the introduction, there are a number of different localization approaches with varying strengths and weaknesses. Due to its simplicity, low cost and potentially high accuracy a widely adopted approach is scene analysis.

To implement this approach for a given technology, it is necessary to design and fine-tune signal aggregation and comparison functions based on the underlying hardware characteristics. This involves the collection of base measurements, for example, to determine the radiation pattern of antennas, the impact of distance between a sender and receiver on the signal characteristics or the dampening factors of certain materials. Once the design and fine-tuning is complete, it is necessary to evaluate the performance of the resulting system. Again, this involves the collection of measurements using the chosen hardware and the computation of the desired performance metrics on top of the tuned algorithms.

During both, the initial measurements for fine-tuning as well as the final measurements for evaluation, the following four factors play an important role:

- *Number of positions*: Since approaches based on scene analysis can degrade based on the complexity of the underlying scene, it is usually necessary to collect measurements for a large number of different positions. This ensures that the initial algorithm design and tuning is not hampered by specific effects of the test environment. Furthermore, it ensures that the final evaluation results provide a strong indication of the system’s achievable performance.
- *Number of samples*: Since most RF technologies are not able to measure signal characteristics perfectly, it is usually necessary to collect multiple samples for each position. This ensures that both the tuning as well as the final evaluation are not significantly distorted through measurement outliers.
- *Positioning precision*: Since scene analysis requires a training phase in which the system is adapted to a particular scene, it is usually necessary to perform 2 consecutive measurements, i.e. one for training and one for validation. When performing these measurements, it is necessary to reposition the signal sources and sinks precisely at the same location twice. Intuitively, in order to ensure that the quality of the data is high, it is necessary to ensure that the repositioning is accurate.
- *Experiment time frame*: Last but not least, it is usually not advisable to perform the training and validation measurements in a short time frame. Instead, it is necessary to stretch out the experiments over a longer period of time. This ensures that temporal effects, e.g. signal fluctuations due to temperature or humidity changes, are represented adequately in the collected measurements.

When trying to minimize the impact of all four factors, it becomes apparent that in many cases, both, the design as well as the evaluation of an RF-based indoor localization system require a considerable amount of precise measurements. Consequently, the measurement process consumes a significant part of

the system developer’s time. While it may seem possible to reduce the time consumption by not considering some of the factors, the impact of doing so is hard to judge and may complicate the design, e.g. due to non-representative measurements, or worsen the expressiveness of the performance evaluation, e.g. due to highly variable and non-reproducible results. In this paper, we argue that it is preferable to reduce the time required by the developer by automating the collection process instead. To clarify this, we present two localization systems that we have built as part of WebDA, one of our ambient assisted living projects, in the following. Thereafter, we describe our experiences with the manual as well as the automatic evaluation of the systems.

3 Indoor Localization in WebDA

The goal of the WebDA [11] project is to prolong the duration during which elderly persons suffering from dementia can be treated at home. For this, the project has developed a platform that provides a number of web-based services to both, the elderly persons as well as their family members or care takers. The goal of these services is to compensate for the ongoing loss of capabilities by providing assisting services that reduce the need for frequent personal (emergency) visits which are often cited as a main reason for moving the elderly to a nursing home and to improve the comfort of the elderly by reducing stressful situations. Among the most important services provided by WebDA are the following two:

- *Context-aware notifications*: By interpreting the movement patterns of the elderly, this service is able to issue context-aware reminders to the elderly, for example, to drink some water when going to the kitchen. Furthermore, it can notify the family members about abnormal and potentially dangerous behavior, for example, the elderly being on the balcony at night.
- *Misplaced object finder*: By automatically gathering the location of important objects in the home, the service allows the elderly to search for misplaced objects. The elder can, for example, make a phone call to his care taker who then performs a remote search in response to the request via a web-interface provided by the object finder service. Also, a key finder module can remind the dement person to take the key when leaving home.

Obviously, both services require accurate and detailed location information. To gather this information in an unobtrusive and automated fashion, we have been working on the development of two low-cost localization systems that rely on active and passive RFID technology. In the following, we briefly describe the requirements on these systems, the details on the hard- and software that we used to realized them as well as their integration with the remaining software. In the next section, we describe our evaluation experiences.

3.1 Requirements

The two primary application areas of indoor localization in WebDA are the real-time tracking of the elderly persons as well as the on-demand localization of

objects. When putting these application areas in the context of ambient assisted living, we can identify the following general requirements which are also often targeted by other systems in this application domain:

- *Low cost*: To be suitable for many elderly persons, the localization system must be inexpensive to deploy and operate. Beyond the cost of the hardware, this also includes factors such as installation cost, the configuration cost as well as the cost for maintaining both, the hard- and the software configuration over an extended period of time.
- *Low obtrusiveness*: To increase the acceptance of the overall system, the localization system should be as unobtrusive as possible. This has a major impact on the hardware form factors that can be used. Moreover, it also has an impact on the actual technology that is used to enable localization, especially when dealing with less technically versed persons.
- *High accuracy*: In order to be usable for the intended purposes, the localization system must be highly accurate. Due to the structure of typical home environments achieving room-level accuracy will not be sufficient for many scenarios. As a simple example, consider that in order to return useful results, the object finder service should be able to differentiate between different locations such as "the shelf" or "the coffee table" even if they are in a single room.
- *High reliability*: Last but not least, the localization system must also be reliable. This does not merely refer to reliability with respect to the localization results, but also implies high reliability with respect to small short-term changes in the environment. Such changes may include the movement of furniture such as chairs or the opening and closing of windows, for example.

3.2 Hardware

In order to support the localization of persons while fulfilling the requirements described above, we chose to use an active RFID system that consists of battery-powered tags and readers. The reason for this choice is the fact that when compared to passive RFID systems, active systems typically exhibit a significantly higher range and thus, this technology requires a lower number of readers. As the readers are typically more expensive than the tags and we are aiming at localizing a single person with dementia, this reduces the overall hardware cost.

To support the localization of objects, however, we chose to use a passive RFID system that consists of passive tags which are powered through an active reader. The rationale for this decision is that passive tags are significantly smaller and cheaper than active tags. Consequently, it is possible to tag a large number of objects of different sizes with little additional cost. Furthermore, due to the fact that passive RFID tags receive their energy through the reader, there is no need to replace or recharge the batteries of the tags that are attached to objects.

Figure 1 shows a picture of the main components of the active as well as the passive RFID system. Both systems consist exclusively of comparatively low-cost off-the-shelf components. In the following, we briefly describe each of the main components of these systems in more detail.



Fig. 1: WebDA Localization Hardware

- *Active RFID System:* For person localization, we use the LogiSphere system from Sensite Solutions [15] which consists of readers (HBL100) and active tags (BN208). The system operates at a frequency of 868 MHz which provides coverage for a typical home environment with more than $100m^2$ easily. Given the overall cost of the system components as well as the range that can be covered by each reader, we expect that a typical installation can consist of four readers per home. The readers can automatically form a wireless multi-hop network which drastically reduces the deployment cost and time since only one RFID reader must be connected via RS-232 to a PC. The remaining readers solely need to be connected to power. Besides from identifying individual tags, the readers are also able to estimate the power of the received signals broadcast periodically by the tags whereby the periodicity can be configured.
- *Passive RFID System:* For object localization, we use a long range passive RFID system from Feig Electronic [2] consisting of OBID i-scan UHF LRU 3500 readers that are connected to up to 4 antennas. Similar to the active system, the passive system operates at a frequency of 868 MHz. Based on the specification, the readers can read tags at a distance of up to $16m$ and when reading the tag, the reader can estimate the received signal strength. However, with our combination of antennas we can typically cover a range of up to $7m$ with one antenna. For a typical mid-sized room of about $50m^2$, this usually results in one reader with 4 antennas per room. Each reader can be equipped with a USB stick to connect itself to a wireless LAN. Just like with the active readers, this allows us to reduce the deployment cost and time since there is no need to run cables from one room to another.

3.3 Software

The localization approach used for persons and objects in WebDA is based on scene analysis. To be more specific, we use a fingerprinting technique that is similar to RADAR [1]. However, in contrast to RADAR which is based on IEEE 802.11 we use RFID technology. As a consequence, there are several significant differences between the approach presented in [1] and the one taken by WebDA.

In the following, we briefly explain the overall approach as well as these differences.

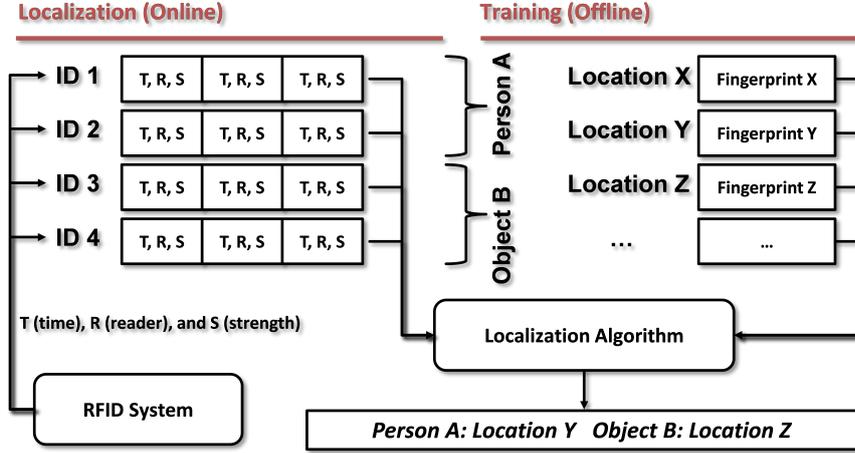


Fig. 2: WebDA Localization Approach

In general, fingerprinting-based localization is done in two phases that are depicted in Figure 2. During the first phase - the so-called training phase, measurements are made at different locations in the target environment. These measurements are then processed - resulting in so-called fingerprints - and the fingerprints are stored together with their respective locations. During the second phase - the actual localization phase, runtime measurements are made and these are then compared to the fingerprints made during the training phase. The location of the persons or objects is then determined by choosing the location of the most similar fingerprint captured during the training phase. Thus, using the environment-specific calibration done in the training phase, it is possible to account for the variances induced by multi-path effects.

However, in order for this approach to work in our setting using RFID technology for person and object localization, we introduce the following modifications to traditional fingerprinting approaches:

- *Time-based aggregation:* As the RSS values computed by our RFID readers exhibit a considerable variance (due to measurement imprecisions) we aggregate several readings over time. To account for outliers, we use the 80th percentile measurement as the true value for the interval. To be more specific, we aggregate readings in a 20 second interval as we experimentally determined this to be a suitable trade-off between accuracy and latency. Intuitively, for the real-time tracking of persons, 20 seconds of latency is already high. To further reduce this, we use a sliding window during the

localization phase which typically results in a latency of about 10 seconds. This latency can be tolerated by our services.

- *Measurement comparisons:* To compare measurements of the localization phase with the measurements made during the training phase, RADAR proposes the use of the Euclidean distance. To do this, the readings of a single tag made by each reader are represented in a vector R_k where k denotes one of the readers. We rely on the same basic approach. However, to reduce the number of comparisons, RADAR proposes to further aggregate multiple measurements made at the same location using the component-wise maximum. Since we did not find this to be necessary in order to achieve a satisfying performance, we are refraining from using this step.
- *Multiple tags:* In the past, researchers have showed that by introducing more readers, it is possible to improve the resulting localization accuracy. However, as the readers are typically much more expensive than tags, we are proposing to use multiple tags instead. In [18], we have shown that this can dramatically improve the localization accuracy when applied to person localization, so we are taking the same approach for object localization as well. The idea thereby is to attach multiple tags to the same person or object. The fingerprints are then constructed as a so-called wide fingerprint by combining the measurements of all tags belonging to the same person or object from all readers. This results in a vector R_{ik} where i denotes one of the tags and k denotes one of the readers. Then we apply the comparison as described above. Through experimentation with our concrete hardware, we found that 4 tags per person and 3 tags per object provide a suitable balance between hardware cost and accuracy.
- *Specific placement:* The last specific component to our localization approach is tag placement. As discussed in [18], for person localization, tag placement has a significant impact on the resulting accuracy. Consequently, for person localization, we attach the 4 tags to a belt of the person so that their orientation stays relatively stable. For object localization, the same argument can be made due to the fact that objects can be rotated. Consequently, in order to account for different rotations, the three tags are attached to each object. When performing the comparison between training fingerprints and the localization phase fingerprints, we can then construct different rotations by swapping the individual components in the vector R_{ik} . As the resulting location, we then use the location that exhibits the minimum Euclidean distance among all locations and all rotations.

3.4 Integration

In order to make the previously described localization hardware and algorithms usable by the user-facing services implemented in WebDA, we have integrated them. Furthermore, to speed up the adaptation of these services to different home environments, we have developed a number of tools. In the following, we briefly describe the resulting architecture as well as the provided tool support.

Architecture The web-based services implemented by the WebDA project rely on the OSGI component model for modularization. Consequently, we used the same framework to implement the localization system. The building blocks of the resulting localization system are depicted in Figure 3a.

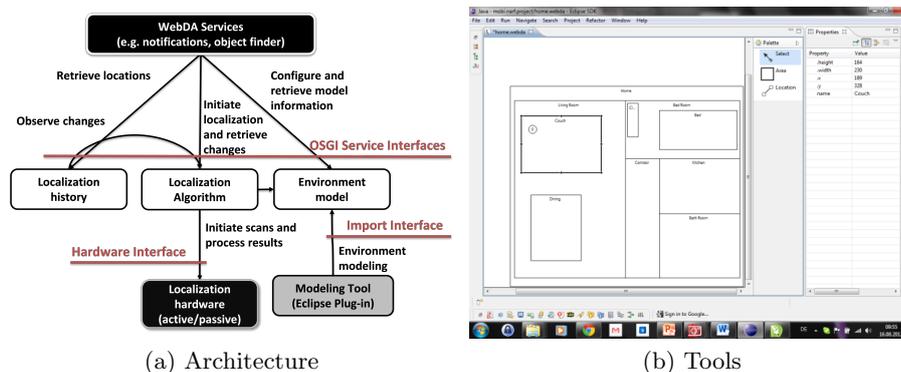


Fig. 3: WebDA Architecture and Tools

At the lowest level, a number of drivers provide access to the active and passive RFID systems. On top of that, the localization algorithm is implemented as an OSGI service. Using this service, other services can subscribe to localization events which indicate the location of persons and objects. Furthermore, they can trigger passive scans in order to start searching for objects. By subscribing to all localization events, a localization history service can store all location information persistently over time. Furthermore, it can make the history of locations available to other services, for example, to perform long-term pattern recognition or to facilitate testing and debugging. In order to abstract from the concrete geometric properties of a particular home environment and in order to generate user interfaces that are meaningful for the users of the system, the localization events are expressed in terms of a symbolic location model. The location model encompasses relevant "zones" such as rooms and other areas, e.g., closets, tables, etc. Furthermore, it models the geometrical containment relationships between the zones. This enables the user interface related services to display, for example, hierarchical menus and refined outputs such as "the object is in the living room close to the couch". Intuitively, this model has to be adapted specifically for each environment as the geometrical properties change between environments.

Tools In order to configure the localization system for a particular home environment, it is necessary to a) model the environment and b) perform the calibration required as part of the training phase of the localization algorithm. To support both tasks, we developed a simple tool on the basis of the Eclipse platform. A screenshot of this tool is depicted in Figure 3b.

Using a graphical editor, a user can model different areas which constitute the environment. Based on the Eclipse Graphical Editor Framework, the editor supports all common actions such as insertion, movement by dragging, deletion, etc. and it is also possible to specify exact coordinates by means of an Eclipse Property View. The modeled areas may refer to different concepts such as different types of rooms or zones within them. By providing an approximate geometry of the areas, the tool can automatically gather the containment relationship. By clicking an export button, the modeled environment can be exported to a database format which is then used by the localization system. In order to calibrate the system, the Eclipse tool can also be used to pinpoint the current location at which a measurement for a fingerprint is taken. The resulting fingerprints and their locations can then be fed into the localization algorithm which can then a) compute the location and b) attach the location to an area in the location model. From our experience, this tool significantly speeds up the overall deployment process as the visual representation of areas and locations speeds up the database configuration and reduces the potential for attaching false locations to the measurements during calibration.

4 Experiences

To evaluate WebDA's person localization system, we have performed an extensive manual evaluation. Furthermore, to evaluate the object localization system, we have recently started with an automated evaluation using our TrainSense testbed. In the following, we describe our experiences during both types of evaluations. Thereafter, we highlight the benefits and limitations of each approach.

4.1 Manual Evaluation

In order to measure the performance of WebDA's person localization system, we performed a manual evaluation the results of which are described in [18]. The goal was to show the impact of antennas on the localization result when localizing a human. Since we wanted to test the system under real conditions it was necessary to attach the active RFID tags to a person since the absorption of the human body also has a high influence on the signals. For this evaluation we performed a measurement at 34 different positions in four directions for one minute each. This means that the pure measurement time is 138 minutes. Since the person wearing the tags needs to turn and move between measurements this time is prolonged. Also due to human nature the person executing the experiment needs some rest during the experiments. From performing these experiments, we learned that for such a measurement it is realistic to assume that it takes about five to six hours. Since we wanted to collect a training and a testing set we needed to repeat the experiment two times. Overall the experiments therefore took about two days.

4.2 Automated Evaluation

To test the accuracy of object localization, we have performed further experiments. However, in order to reduce the time requirements induced by manual evaluation, we tried to automate the data collection process. As the experiments required us to place the tags at a number of positions in a repeatable manner, we decided to use our TrainSense testbed for automation. TrainSense uses miniaturized model trains to simulate the nomadic movement of cooperating objects. In the following, we briefly outline this testbed before we describe our experiences.

TrainSense The TrainSense testbed [16] uses digitally controlled model trains to carry on various kinds of wireless experiments involving stationary as well as mobile elements. However, for the evaluation of our indoor localization system, we used it as an automated system to position the RFID tags at several locations in a precise and repeatable manner. For our experiments, the passive RFID tags are attached to a train that is moved automatically to each desired location. This system allows us to position a tag within a range of less than 2 centimeters, which is precise enough to obtain valid measurements in an automated fashion.

Figure 4a illustrates the basic hardware architecture of TrainSense. The system is based on off-the-shelf model trains and it consists of a computer host, a controller, train detectors, tracks and locomotives. This setup is similar to what is commercially available. The tracks, detectors and trains are bought off-the-shelf. The tracks are mounted on wooden plates which can be combined arbitrarily in order to create different layouts for different experiments. The wiring required among the different components of the train control is exactly the same as the one required for standard model trains and it is built into the modules. Similarly, the communication protocols used in our system are standard, to allow the reuse of existing, low-cost hardware components (Maerklin/Motorola I for the train control and S88-N for the train detection).

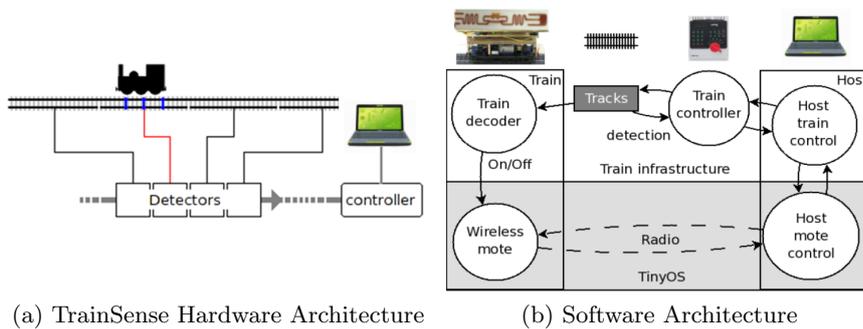


Fig. 4: TrainSense

However, the hardware and software of the controller and the detection mechanism have been re-designed to satisfy the needs of precise and repeatable wireless experiments. The controller is a real-time system, which means it reacts to the perceived events, such as the detection of a train at a particular location, within a bounded amount of time. Beyond providing the ability to detect the passage of the train at a detector, this also enables us to let it run for a precisely determined duration and to stop the train quickly. If we do not consider extremely rare mishaps like stalling or derailment, this allows us to place the train with a guaranteed precision.

For most of our evaluations, we use basic dead reckoning to position the tag-carrying train. For this, the train is set to run at a constant speed. When it crosses a detector, the controller is instructed to let it run for a programmed amount of time (proportional to the distance from the detector to the targeted position). After that time has elapsed, the controller stops the train. Once the train has reached its destination, the host starts the code that performs a measurement for the current position.

The resulting overall software architecture is depicted in Figure 4b. The figure shows the three main components of the testbed, which are the train that is able to move to a position (left), the real-time controller that controls the movement of the train and the host computer (right) that runs the experiment by performing the following steps:

1. Move the train to the first position and signal the arrival to the software on the host computer.
2. Once the arrival is signaled, start the measurement software and wait for its completion.
3. If there are more positions, go to the next one, signal its arrival and repeat from step 2.
4. Else signal the end of experiment and run the measurement software to compute the results.

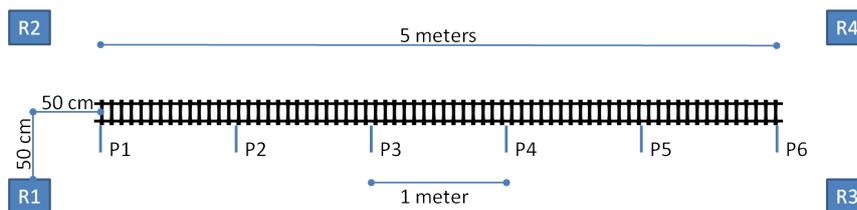


Fig. 5: Experimental setup

Experiences For the evaluation of the passive RFID system we were interested in how the radio signals behave over time for certain positions. We therefore

set up the experiment with a five meter long testbed shown in Figure 5 making measurements every meter. This results in six different positions marked as P1-P6. TrainSense offers the possibility to send the Trains to fixed positions in the testbed with a precision of less than 2 cm. Due to the nature of the track layout this error can also only occur in one dimension. In our experiment this gives an upper bound for misplacements by 2%. For the capturing of the signals we used four antennas connected to the passive RFID reader. Two of those antennas (R1,R2) were placed 50cm from the beginning of the tracks and 50 cm to the side while the other two were placed in a similar fashion on the other side of the tracks. In order to be able to measure the signals of the tags at different positions we attached three RFID tags to the train. These tags were arranged in a triangle to expose different angles to the readers antennas. The photo in figure 6 shows the real setup as we have taken the measurements.

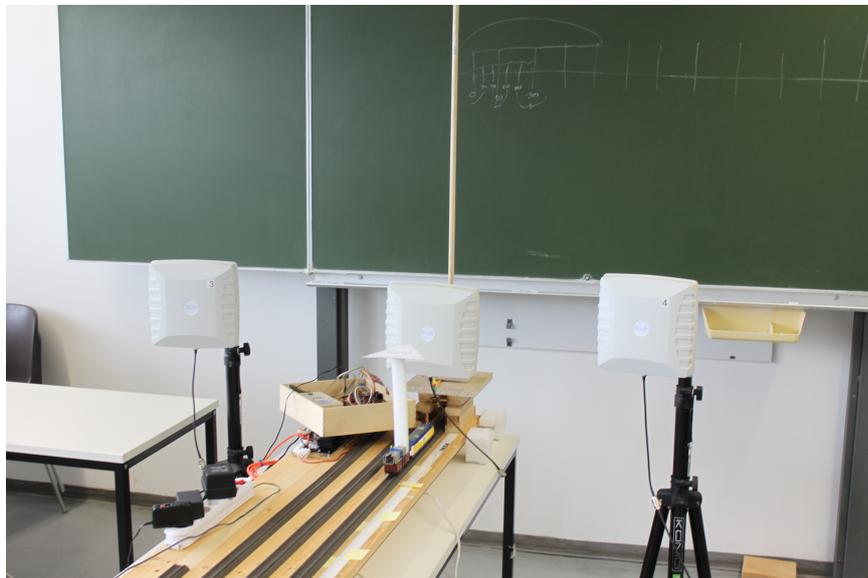


Fig. 6: TrainSense with RFID technologies

While setting up the experiment certainly takes a few hours because of the placement of the rails, antennas and other hardware factors this is only needed once before starting the real experiment. After the sensing software is connected with the TrainSense platform and the required positions are programmed, the experiment can run unattended. This is especially handy when things go wrong. So when for example the cable of an antenna is not properly connected it is easy to repeat the whole process. This also allows for repeating the same experiment again later without wasting the system developer's time.

In our case we ran the experiment twice with some hours in between. We then analysed the results of both sets as follows. In each set signals we collected for a period of two minutes per position. We divided each of these sets in 8 consecutive bins resulting in 15 seconds long parts of the data. For each of these sets we calculated the average signal strength and standard deviation for each tag - antenna combination. We found that for these intervals the standard deviation is close to zero in almost all cases so we continued our evaluation only considering the average signal strength. We then did a cross-validation on the 8 sets per position measuring the differences in the average signal strength. Figure 7 shows the result of this evaluation depicting the number of occurrences of certain signal differences within the single measurement runs and between the two runs. The graph also depicts the number of sets where one tag was read at one antenna in one set but not in the other set.

It can clearly be seen that when using only one measurement a cross-validation the signal differences are only minimal and for the majority of the data well below 1 db. While this is true for both single sets the same evaluation using both sets shows very different results. Here it can clearly be seen that the majority of differences are around 3db. When considering the number of sets that have one tag-antenna combination in one bin of the first set but not in the bin to which it is compared it becomes clear that the likelihood of such an event increases with time. This experiment shows that for the evaluation of fingerprint based systems it is essential to have two disjunct sets of data that are used for training and testing. When such a system is tested with only one data set which is then split into different bins for a cross-validation the results may not be accurate.

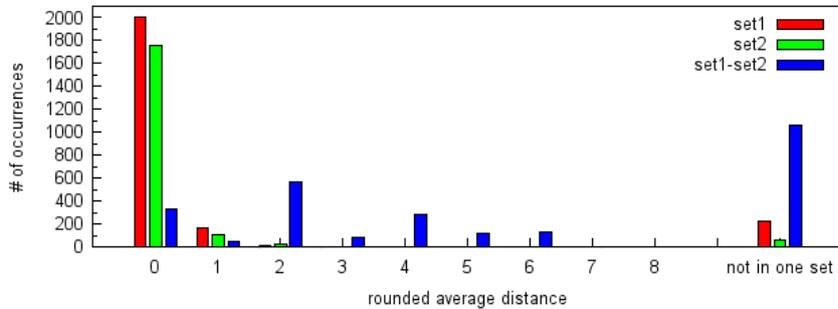


Fig. 7: Differences in signal space within and across different measurements

4.3 Discussion

Evaluating localization systems with the TrainSense platform has shown high potential. While it might require some time to set up the testbed and connect the measuring software to the testbed controller the amount of time that can be

saved during the actual experiments will pay off very fast. This is especially true for large scale experiments where a lot of different positions need to be measured. Running such an experiment manually would block the system developer for the whole time whereas TrainSense makes it possible to run the experiment unattended. This makes TrainSense a good choice for measuring signal-strengths in the case of object localization or when trying to simulate the behaviour of the senders without any interferences – which is a typical measurement during the initial system design. On the other side TrainSense is not capable of simulating dampening factors such as a human body. If such factors are needed for the system it will still be necessary to do the measurements manually. As discussed in detail above this can be a very time-consuming task and therefore it might seem reasonable to use cross-validation on only one dataset to generate the needed results. As shown in Figure 7 this might not lead to reasonable results.

5 Conclusion

Knowledge about the location of persons and objects is an important basis of many smart city applications. The unavailability of GPS in indoor environments has led to active research on approaches and systems to enable indoor localization using various cooperating objects technologies. A key challenge during the development of any of these indoor localization approaches and systems is the systematic evaluation of their performance. In this paper, we described how the time requirements can be reduced by means of automation. Furthermore, based on our experiences with both, manual and automatic evaluation, we discussed the achievable benefits and possible limitations. Based on this, we argue that in cases where it can be applied, automation is preferable to simplified experiments as these may lead to less expressive results. At the present time, we are completing our automated evaluation of the WebDA object localization system using our TrainSense testbed. Simultaneously, we are working on several improvements to the software as well as the hardware of the testbed in order to further increase its applicability to other scenarios.

Acknowledgements

The work presented in this paper has been partly funded by the German Federal Ministry of Education and Research as part of the WebDA project under grant number 16SV4023 and the European Center for Ubiquitous Technologies and Smart Cities (UBICITEC).

References

1. P. Bahl and V.N. Padmanabhan. Radar: an in-building rf-based user location and tracking system. In *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 2, pages 775 –784 vol.2, 2000.

2. Feig Electronics. Feig obid rfid system. <http://www.feig.de/home.html>.
3. S. Holm. Hybrid ultrasound-rfid indoor positioning: Combining the best of both worlds. In *IEEE International Conference on RFID*, pages 155–162, april 2009.
4. V. Honkavirta, T. Perala, S. Ali-Loytty, and R. Piche. A comparative survey of wlan location fingerprinting methods. In *6th Workshop on Positioning, Navigation and Communication, 2009. WPNC 2009.*, pages 243–251, march 2009.
5. Elliott Kaplan and Christopher Heygarty. *Understanding GPS: Principles and Applications*. Artech House Inc., 2005.
6. M. Kranz, C. Fischer, and A. Schmidt. A comparative study of dect and wlan signals for indoor localization. In *IEEE International Conference on Pervasive Computing and Communications*, pages 235–243, 29 2010-april 2 2010.
7. Hui Liu, H. Darabi, P. Banerjee, and Jing Liu. Survey of wireless indoor positioning techniques and systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 37(6):1067–1080, nov. 2007.
8. Diego Lopez de Ipina, Paulo R. S. Mendonca, and Andy Hopper. Trip: A low-cost vision-based location system for ubiquitous computing. *Personal Ubiquitous Computing*, 6:206–219, January 2002.
9. Anil Madhavapeddy and Alastair Tse. Study of bluetooth propagation using accurate indoor location mapping. In *The Seventh International Conference on Ubiquitous Computing*, pages 105–122, 2005.
10. A. Matic, A. Papliatseyeu, V. Osmani, and O. Mayora-Ibarra. Tuning to your position: Fm radio based indoor localization with spontaneous recalibration. In *2010 IEEE International Conference on Pervasive Computing and Communications*, pages 153–161, April 2010.
11. Yehya Mohamad, Henrike Gappa, Jaroslav Pullmann, Gaby Nordbrock, Carlos Velasco, Marcus Handte, Stephan Wagner, and Marcel Schweda. Context-aware support for people with dementia and their families. In *VDE 5. Deutscher AAL Kongress*, January 2012.
12. Lionel M. Ni, Yunhao Liu, Yiu Cho Lau, and Abhishek P. Patil. Landmarc: Indoor location sensing using active rfid. *Wireless Networks*, 10:701–710, 2004. 10.1023/B:WINE.0000044029.06344.dd.
13. Tom Pfeifer and Dirk Elias. Commercial hybrid ir/rf local positioning system. In *Kommunikation in Verteilten Systemen*, February 2003.
14. P. Prasithsangaree, P. Krishnamurthy, and P. Chrysanthis. On indoor position location with wireless lans. In *The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, volume 2, pages 720–724, sept. 2002.
15. Sensite Solutions. Logisphere rfid system. <http://www.sensite-solutions.com/>.
16. Hugues Smeets, Chia-Yen Shih, Marco Zuniga, Tobias Hagemeyer, and Pedro José Marrón. Trainsense: A novel infrastructure to support mobility in wireless sensor networks. In *EWSN*, pages 18–33, 2013.
17. Ubisense. Ubisense webpage. <http://www.ubisense.net>, April 2011.
18. S. Wagner, M. Handte, M. Zuniga, and P.J. Marron. On optimal tag placement for indoor localization. In *IEEE International Conference on Pervasive Computing and Communications*, pages 162–170, march 2012.