Radio map upscaling - adding antennas in an indoor localization scenario

Roland Stenzl, Stefan Wilker, Thilo Sauter
TU Wien
Institute of Computer Technology
Gußhausstraße 27-29/E384, 1040 Vienna, Austria
{firstname.lastname}@tuwien.ac.at

Anetta Nagy, Thomas Bigler, Albert Treytl
Department for Integrated Sensor Systems
Danube University for Continuing Education
Viktor Kaplan Straße 2-E, A-2700 Wr. Neustadt, Austria
{firstname.lastname}@donau-uni.ac.at

***Abstract*— With technological revolutions like the Internet of Things and Industry 4.0 on the doorstep, localization of network nodes is more important than ever. Indoor localization systems mostly work with omnidirectional antennas for simplicity reasons. Using directional antennas has advantages, though, and devices with such antennas will be included in IoT systems someday. In this paper, the focus is on dynamic changes in the configuration of localization systems. Adding an antenna must be handled fast and efficiently. Frequent recalculations of the configuration could be necessary, and limited resources make a reduction of the computational effort desirable. In this work we look at possible methods to cut the computational effort while minimizing the penalty on accuracy. Since accuracy is of utmost importance for indoor localization, there is a fine line which is important to find.**

***Index Terms*—RSS fingerprint, position estimation, directional antenna, clustering,up-scaling**

I. Introduction and previous work

The increasing deployment of mobile, interconnected devices also calls for fast and efficient localization approaches. While outdoor localization is fairly solved today, indoor localization is still a topic for research with many challenges which have to be overcome. There are different methods to design a indoor localization system, but usually omnidirectional antennas are used for this purpose. Using directional antennas in the anchor nodes is less common, but can have advantages such as reduced multipath interference [1]. The trade-off between localization accuracy and computational efficiency is an important topic for indoor localization systems and therefore the focus in this paper.

Our previous work simulates directional antenna patterns to create a radio map with received signal strength (RSS) values [1]. In a two-dimensional localization scenario, a minimum of three anchor antennas is needed. A fingerprinting method [2], [3] then uses the best three reception values. The fingerprint is abstracted into integer values from zero to 20, where the value 20 represents the best reception and zero a loss of connection. A two-dimensional model was preferred for usability and easier development, but the system was planned such that an extension to three dimensions is possible with minimum effort.

The core idea of making localization efficient is to use clustering. To that end, the list of RSS sample points is clustered. This divides the plane into areas of the same RSS value, which are then used for the actual localization. Clustering is done with centroid-based algorithms. This results in an circular area in two dimensions or a sphere in three dimensions. The center of the area is used as starting point and the basis for the localization algorithm. It is therefore important that the clustering is of high quality. This can be evaluated by different scoring methods [4]. K-Means [5] was so far found to be useful, mainly because of the low computation time. The number of clusters per RSS value depends mainly on the geometry of the sample-points, but also on the type of cluster algorithms used. This makes the choice of the algorithm important.

The number of sample points naturally varies for each RSS value. It is known, however, that for good results, a minimum amount of sample points involved in a clustering process is necessary [6]. Therefore an upsampling strategy is reasonable to provide additional, extrapolated data points, if the number of sample points was too low to guarantee a seamless clustering. In our case, a minimum of 50 data points were chosen. If only fewer points of that particular RSS value were measured, every sample point was surrounded by nine additional sample points.

If a target antenna enters the plane, its RSS value is measured and compared to the radio map. The clusters of that RSS value are then used by the localization algorithm to pinpoint the target. A drawback of this numerical approach is that a lot of data accumulate. It must therefore be carefully decided, if certain recalculation effort is worth it. In this work we present the upscaling process. It deals with the changes in configuration when an anchor antenna is added to the scenario. Big sets of data take a long time to cluster which created demand for an alternative. The idea was to regroup sample points which, after changes in the RSS map due to an added antenna occurred, are assigned to wrong clusters.

II. Upscaling a scenario

In an indoor scenario, anchor antennas are fitted to a room and target antennas enter and leave the plane of localization. Moving antennas could help with the localization of the targets. A previously pinpointed target antenna could also be used as moving anchor. Since a moving anchor would mean that the antenna location in the scenario is continuously changed, it should be possible to add anchor antennas with minimal computational effort. The idea of the upscaling module was to reduce the amount of data with reasonable losses in cluster quality. Recalculation of the whole radio map and a re-clustering of the sample points might, given a complex scenario, be a time consuming task. Therefore, ideas to recalculate only parts of the radio map, and doing only simplified changes to the clustering are necessary. In this paper, we focus on adding an antenna to a scenario, however keeping in mind that the underlying principles can be extended to different kinds of changes in the future with low additional effort. Three different methods with different level of processing effort are considered. In the order of decreasing computational effort, they are:

- full recalculation
- cluster change method
- area assessment method

A full recalculation of the scenario is a very complex solution and should be avoided. Per RSS value, one clustering is done, which in our case means that a maximum of 21 clustering processes per scenario is possible. The time for a clustering process to finish fluctuates heavily. Nevertheless, this method was implemented as baseline for the comparison with other methods. It provides an idea of how the RSS map and the clustering should look after adding an anchor antenna. The recalculation consists of two steps: the RSS map is recalculated and the clustering re-done with the new data. The process of recalculating the radio map is a simple iteration in the simulation, and the computational effort is therefore low. The most time consuming and complex process is the clustering, and the most savings can be done here.

The cluster change method recalculates the RSS map in the same way, but instead of re-clustering the new data, the sample points are sorted manually into either an existing cluster of the new RSS value, or an additional clustering is done with only the remaining unclustered sample points. The area assessment method works in the same way, but considers only the area with high reception values of the added antenna, which further decreases complexity, as will be shown in sec. IV.

III. Validation Methods

By changing the cluster data manually, the quality of the clusters changes. To measure the quality of the clustering after these changes, we apply the same cluster scoring methods as in [4]:

- Silhouette analysis
- Calinski-Harabasz score
- Davies-Bouldin score

These three algorithms are popular methods to evaluate cluster quality and are described in more detail in [4]. For the purpose of rating the cluster quality, the most important aspect is how the metrics work: The silhouette score lies between [-1, 1]. This score and the Calinski-Harabasz score use a higher value to indicate better cluster quality, while the Davies-Bouldin score assigns a lower value to a better quality.

IV. Implementation

In Fig. 1, a plane is shown with four anchor antennas, one in each corner directed towards the middle of the plane. This scenario depicts a typical localization scenario. A large plane was chosen and therefore the RSS values max out at the value eleven, in the corners, right in front of every anchor antenna. If the reception in to room should be increased, another anchor

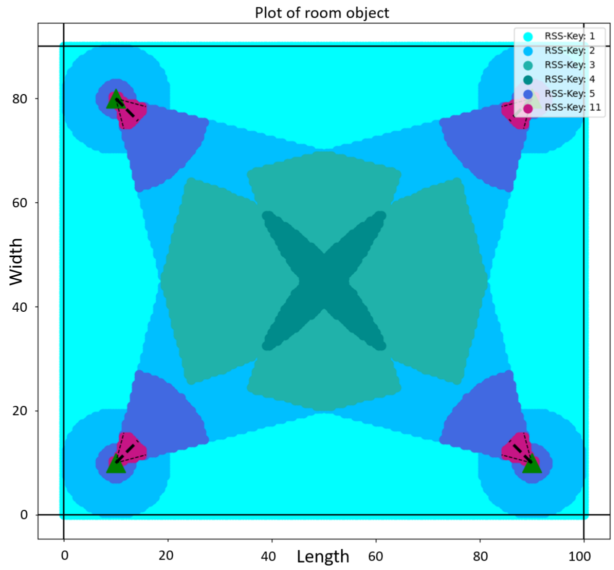

Fig. 1. Reception with four antennas

antenna can be placed on the plane. Here, we add an antenna in the middle of the plane via the upscaling process. In Fig. 2 the changes in the RSS map can be seen. In the lobe of the newly added antenna the RSS values rise and the maximum value increases to 13.

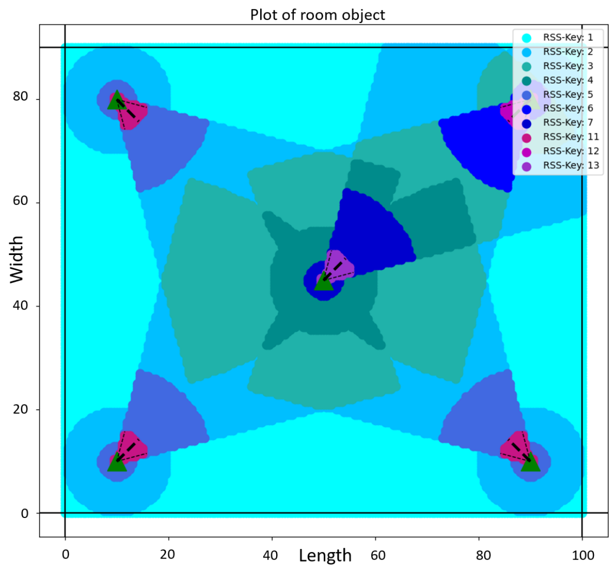

Fig. 2. Reception after adding another antenna

A. Full recalculation

In a full recalculation the RSS map is recalculated and the clustering done again. This method provides most accuracy, since it calculates for every sample point the exact value and re-clusters the whole RSS map. This recalculation has a high computational effort and serves mostly to provide comparison values for the other recalculation methods.

B. Cluster change

For the cluster change method the RSS map recalculation is done the same way as before. After that, the sample points that are affected by the lobe of the additional antenna might have a different measured RSS value. The problem is that these sample points are now assigned to wrong clusters, since their RSS value is not the same as the RSS values of the other sample points in the cluster. This is solved in the full recalculation method by re-clustering every sample point. This behaviour is very inefficient, since many sample points are not affected by the new antenna. The cluster change method assigns these sample points to clusters of their new RSS value manually if possible. If there is no available cluster, they are saved for a separate clustering process. If a sample-point with a different RSS value is found in a cluster, it is deleted from that cluster and we search for available clusters of that new RSS value. The following rules ensure that every sample point gets assigned to a new cluster:

- If no clusters of the new RSS value exist, the sample-point is saved.
- If another cluster is found which already inherits the sample point (which is possible because the clusters can overlap), the sample point is added to the new cluster.
- If no such cluster is found, the next option is a cluster which size would increase only by a small value if the sample point was added. The sample point is added to the new cluster.
- If no such cluster is found, the sample point is saved.

The sample point is then deleted from its old cluster and all changes in cluster size considered. After this process is done for all sample points, the saved sample points get clustered separately and added to the scenario. This strategy has two benefits:

- Unnecessary clustering runs are avoided. If a sample point is near a valid cluster, it is just added instead of re-clustering the whole cluster.
- Only the sample points which are affected by the additional antenna are re-clustered, if it was not possible to assign a cluster to them manually.

C. Area assessment

The area assessment method works similar to the cluster change method, but instead of considering all RSS changes in the map, changes are only considered in a certain area. This area inherits the best reception values of the new antenna. The RSS fingerprint method used takes the best three reception values of the anchor antennas in the scenario. Therefore the loss of information is reduced, if lower reception values are neglected. The area is defined as rectangular and inherits the most significant parts of the antennas front and back lobe. Outside of a simulation, this area has to be measured for every antenna beforehand. This saves further recalculation effort and respects the most significant changes in RSS. In Fig. 3 the result of changes in RSS value based on the area assessment can be seen.

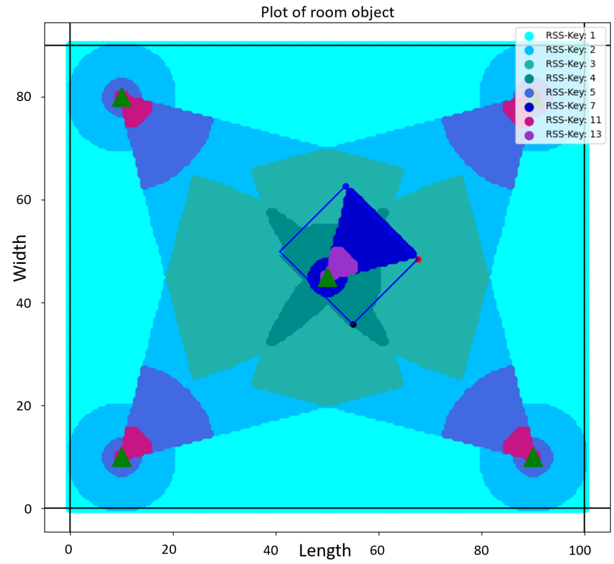

Fig. 3. Area assessment method

If an antenna is added to increase reception, the antenna should be added where the reception is low. In the case that area assessment is used, only the RSS value changes where the antenna is located, are considered.

V. Preliminary Results

Like in previous work, a scenario generator was used to generate 50 different scenarios. The parameters of these scenarios were randomly chosen. The size of the plane could vary from two meters to 20 meters, with a number of anchors from three to 10. All anchor antennas were directional and their direction was chosen randomly in 15° steps. An antenna with randomized parameters was added to the map. The three different methods were used on all 50 scenarios, processing times and the cluster scores were measured. In Tab. I the results are listed and the respective best values are underlined. The full recalculation method provides benchmark values.

TABLE I
UPSCALING - RESULTS

Method	Time[s]	SC	CH	DB
Full	26.0	<u>0.658</u>	<u>5755</u>	<u>0.475</u>
Change	23.7	0.591	5581	0.554
Area	<u>8.0</u>	0.607	5674	0.542

All measurements in Tab. I are mean values across all scenarios. The measured processing time shows that recalculating the scenarios took a long time in general, which supports the suspected need for less complex calculations. The cluster change method took 91.2% the time of the full recalculation, while the area assessment method needed only 30.8%. Cluster change therefore provides not enough reduction of the data to be processed, and area assessment is preferable. The processing times are in general too high to consider a continuous flow of changes in a scenario. From these times, the share needed for clustering varies:

- full recalculation: Average clustering time was 19.77 seconds (76.0% of the complete run time)
- cluster change: Average clustering time was 1.94 seconds (8.2%)
- area assessment: Average clustering time was 1.20 seconds (15.0%)

In a full recalculation, clustering takes most of the time. This shrinks dramatically if cluster change is used. Area assessment saves further time in re-clustering and in calculating RSS changes.

As the scores show, the full recalculation provides the best results as suspected. Interestingly, the scores of cluster change are worse than those of the area assessment method. In contrast to area assessment, the cluster change method considers all changes in RSS. Therefore more clusters are exposed to bigger changes and the cluster scores get worse. Nevertheless, the loss of cluster quality in the example is over all not drastic. This supports the idea that further simplifications are possible, without sacrificing much cluster quality.

Another important metric is the change in RSS value for the area assessment method. Since not every sample point changes to the correct RSS value, the mean RSS value was compared with the same value of the full recalculation. The mean value of RSS for the full recalculations was calculated to be 3.178, while the mean value of RSS measured in a area assessment upscaling was 3.057. This loss of accuracy can be very well neglected for smaller and simpler scenarios. The difference will most likely increase, especially if a third dimension is added, which is the case in real IoT applications.

VI. Suitability of methods

From the results, it can be seen that further streamlining is necessary. The area assessment method is a good way to trade-off wrongly calculated RSS values to gain better cluster quality. The loss in RSS accuracy of the area assessment method is small enough to be neglected in these simple scenarios. These results make future changes to the system, with reduction of complexity in mind, feasible. The limitations of computing resources in simple IoT devices mean that the whole process must be redesigned to further save complexity. The best way seems to further decrease the impact of an added antenna. The lobe could be further limited to areas were bigger changes occur.

In previous work Meanshift and Affinity Propagation were the cluster algorithms that were predominantly used. These algorithms provided a simple way to cluster data, but a higher computational complexity as they are more complicate algorithms. In the course of this work, k-Means was used. K-Means needs additional calculations to define the best number of clusters. This was done with silhouette analysis. This change of the preferred algorithm was done because in preliminary tests, k-Means was more than three times faster than the other algorithms. The change to k-Means was crucial to save time in these tests. The low computational complexity is its main advantage and the most important argument for k-Means. This means that k-Means will have a bigger role in the future research. In an upscaling run, many clustering processes are done. Most clustering processes are done in a few milliseconds, while sometimes such a process can take seconds. This opens a significant possibility to further speed the process up.

VII. Conclusion and Outlook

The recalculations of the RSS map are very complex and take a lot of time. A serious effort must be taken to reduce complexity of the system if the system should support dynamic reconfigurations. This is even more important with future restrictions of real hardware. In the future, the focus will be shifted on methods with low computational effort like k-Means. This will ensure that the system is as efficient as possible. Analyzing the defining parameters for increased re-clustering times will be the next effort in the project. Further steps include tests with data from real antennas. Testing the software on this kind of data will bring more information about the actual requirements. Especially the increase of data sets will show which algorithms will provide enough processing power to bring good results. The simulation will be expanded to use transmission and reflection models.

Acknowledgments

This work is partially co-funded by the NÖ Forschungs- und Bildungsges.m.b.H. (NFB) within the FTI-Call under the project PinPoint IoT (FTI17-006). The authors are responsible for the contents of this publication.

References

[1] A. Nagy, T. Bigler, A. Treytl, R. Stenzl, S. Wilker, T. Sauter, and T. Wien, “Rss-based localization for directional antennas,” in *2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, vol. 1, 2020, pp. 774–781.

[2] S. Tomic, M. Beko, and R. Dinis, “3-d target localization in wireless sensor networks using rss and aoa measurements,” *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 3197–3210, 2017.

[3] W. K. Zegeye, S. B. Amsalu, Y. Astatke, and F. Moazzami, “Wifi rss fingerprinting indoor localization for mobile devices,” in *2016 IEEE 7th Annual Ubiquitous Computing, Electronics Mobile Communication Conference (UEMCON)*, 2016, pp. 1–6.

[4] R. Stenzl, S. Wilker, T. Sauter, A. Nagy, T. Bigler, and A. Treytl, “Work-in-progress: Usage of clustering algorithms for analysis of radio maps for localization using directed antennas,” in *2020 16th IEEE International Conference on Factory Communication Systems (WFCS)*. IEEE, 2020, pp. 1–4.

[5] D. Arthur and S. Vassilvitskii, “k-means++: The advantages of careful seeding,” Stanford, Tech. Rep., 2006.

[6] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.