Universität Stuttgart

Copyright Notice

Algorithms for hotspot coverage estimation based on field strength measurements

*Stephan Lück, *Christian M. Mueller, *Michael Scharf, and Robert Fetscher
*Institute of Communication Networks and Computer Engineering, Universität Stuttgart, Germany
Email: {stephan.lueck, christian.mueller, michael.scharf}@ikr.uni-stuttgart.de, robert.fetscher@googlemail.com

Abstract— Cellular networks are more and more comple-mented by wireless LAN hotspots. However, when users are mobile, the small coverage area of hotspots is a challenge for the access network discovery and selection. Network selection mechanisms can be improved if, in addition to terminal-based measurements, hotspot information can be retrieved from an information service. In this context, the coverage area is one of the most important characteristics of hotspots. This paper studies algorithms for the precise estimation of hotspot coverage areas based on field strength measurements performed by individual terminals. Different approaches for interpolation, vectorization and complexity reduction of measurement data are presented and evaluated in different scenarios by means of simulation. As a result, guidelines for the optimal parametrization of these algorithms are presented.

I. INTRODUCTION

Today's landscape of wireless access networks is inherently heterogeneous: In addition to cellular access networks, more and more wireless LAN hotspots are deployed. These hotspots offer high data rates, but only have a limited coverage area. In order to be "always best connected" [1], mobile terminals have to discover and select the best access network among the available networks. Access network selection requires an access discovery function, detecting available networks and their properties such as data rates, prices and coverage area. Traditional access discovery based on scanning procedures cannot provide all this information. Hence, it is advanta-geous to complement the scanning procedures with queries to location-based information systems. Since this approach relies on data models, we call it model-based access discovery.

Model-based access discovery requires that information such as location and extent of hotspots is acquired. Some information can be obtained from network operators and third party providers. However, precise information about the coverage area of hotspots requires extensive field strength mea-surements. In [2], we have shown that this can be realized by automatically gathering measurements from many individual users. The accurate prediction of the coverage area requires a rather complex processing of the raw measurement data. Therefore, this paper studies the suitability and parametrization of algorithms to convert field strength measurements into a simple description of the coverage area, such as polygons.

The remainder of this paper is structured as follows: Sec-tion II briefly describes the architecture for model-based access discovery and reviews related work. In section III, different algorithms for hotspot coverage estimation are presented. Sec-tion IV covers our evaluation methodology and in section V, the impact of key parameters on the quality of the estimation is studied. Finally, section VI concludes the paper.

II. MODEL-BASED ACCESS DISCOVERY

An architecture for model-based access discovery consists of four principal functions [2]: measurement acquisition, mea-surement storage, coverage estimation and information provi-sion (see Fig. 1). For measurement acquisition, the mobile terminal records signal strength values measured at the radio interface and adds a timestamp, the cell ID and the geographi-cal position. As soon as a suitable connection is available (high bandwidth or low cost), the terminal transmits the data to the measurement storage function. Both measurement storage and coverage estimation are located within a report database. If a certain amount of new data has been collected, the report database recalculates the coverage information.

The estimated coverage information is represented by mul-tiple polygons, which may contain holes (multi-polygons). This form of information representation is more compact than the raw data. In addition, it is appropriate for network selection decisions because standard geometric operations can easily be used to check whether the terminal is within a particular access network or to determine the distance to it. The coverage information can be complemented by additional information such as the load of the corresponding hotspot, available services, and prices. The report database then inserts the resulting data record into a location-based information system, which is queried by the mobile terminals.

In [2], it is shown how the open context-based information System "Nexus" can be used for this purpose. However, spe-cial information systems dedicated to access discovery can also be applied. For network selection, a terminal centric approach is followed, assuming that the access selection function is performed within the terminal.
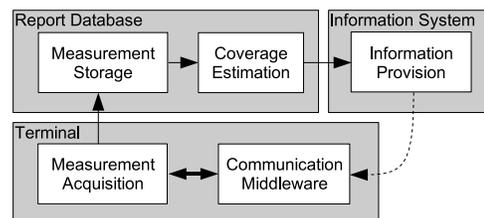

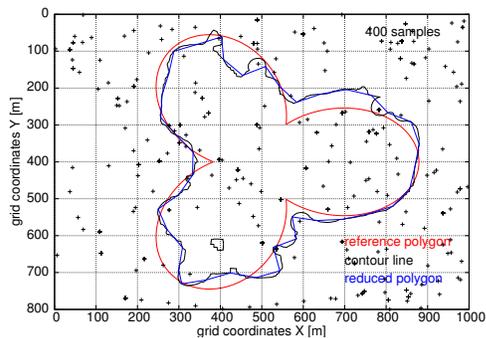Fig. 1.   Architecture for Model-Based Access Discovery

Fig. 2. Coverage estimation of a three sector access network

Some related work also investigated systems that collect signal strength measurement data. For example, the Hybrid Information System (HIS) [3] uses such data to generate handover triggers. In contrast to this work, it mainly operates directly on the measurement values without generating data for context based information systems. Ref. [4] presents a peer-to-peer approach that realizes the data storage for the HIS. The IEEE 802.21 working group is currently specifying the Media Independent Information Service (MIIS) as part of the Media Independent Handover Services (MIH) [5], which defines two alternative approaches of information modeling and representation. However, data acquisition and processing is out of the scope of this draft standard.

III. ESTIMATION OF COVERAGE INFORMATION

In this section, a method to estimate the coverage area of radio cells is described. Some of the algorithms used in this method are well-known from the field of Geographic Information Systems (GIS) and image processing. The overall method consists of three steps: interpolation, vectorization, and complexity reduction. Fig. 2 shows the effect of the algorithms. Besides the true shape of the radio cell (denoted reference polygon), the contour line generated by the vectorization algorithm and the reduced polygon after application of the complexity reduction algorithm are depicted.

A. Interpolation

The interpolation step serves two purposes: First, signal strength values for areas where no measurements are available are calculated. Second, given that the field strength values are randomly distributed, they must be aligned to a regular grid, which is needed for the subsequent vectorization. Several interpolation algorithms [6] such as Voronoi Interpolation, Delaunay Interpolation, and Inverse Distance Weighting (IDW) could be used for this purpose.

The Voronoi interpolation assigns the value of the nearest single data point to a particular grid point. It subdivides the whole area into multiple Voronoi polygons so that each polygon contains one data point. The value of a data point is assigned to all grid points covered by the polygon. As a consequence, gradually varying values are converted into stepwise varying values. The Delaunay Interpolation can avoid this problem. It subdivides the area into triangles whose

vertices are defined by the data points (Delaunay triangulation [6]). It constructs a three-dimensional surface by interpreting the signal strength values as the height (z-coordinate) of the vertices. A drawback is that only three neighboring data points are considered to interpolate a certain raster point.

To avoid the disadvantages of Voronoi and Delaunay interpolation, an algorithm allowing for a higher and better controllable number of data points, should be applied. A quite simple solution is the Inverse Distance Weighting (IDW). To interpolate a point at a given position, IDW calculates the average of all interpolation points weighted by the inverse distance to the data points. If only few interpolation points are available, it averages over a large area. Consequently, good radio conditions might be estimated for areas where no measurements are available although the radio conditions in these areas are actually bad. In order to avoid such an overly optimistic coverage estimation, the interpolation can be restricted to a limited area and consequently includes only a limited set of neighboring data points.

The area taken into account influences how dead spots are estimated. Large interpolation areas tend to eliminate dead spots, but might be too optimistic. Small interpolation areas tend to comprise small dead spots in actually covered areas where no measurements have been taken, leading to complex and inaccurate cell geometries with a high number of holes. Thus, the interpolation area is an important parameter to obtain satisfying results for different numbers of measurements.

Another parameter to be defined is the grid spacing. The smaller the grid spacing, the more accurate is the interpolated data. However, decreasing the grid space augments the number of values to be interpolated such that processing time increases. Thus, a trade-off between grid spacing an computational complexity must be realized.

B. Vectorization

After the measurements have been transformed to raster data, a vectorization algorithm is applied to compute geometric shapes representing the radio coverage of a cell. The choice of vectorization algorithms depends on the characteristics of the input data, e.g., whether the raster data is binary or has a continuous range, on the structure of the shapes, e.g., large line segments or curved shapes, and on the ratio of pixels that are part of the shape. It further depends on the characteristics of the desired output data, whether it consists of independent line segments, contiguous segments, or polygons. Finally, the computational complexity should be kept in mind. For coverage estimation, input data are continuous range values and the coverage area can be best described by a curved shape. To simplify processing in subsequent operations, polygons are expected as output data. In consideration of these characteristics, we chose a contouring algorithm by Bourke [7] to perform the vectorization.

The Bourke algorithm calculates contour lines for one or more contour levels. The algorithm subdivides the radio field into non-overlapping rectangles, each defined by four neighboring values of the grid. For each rectangle it calculates
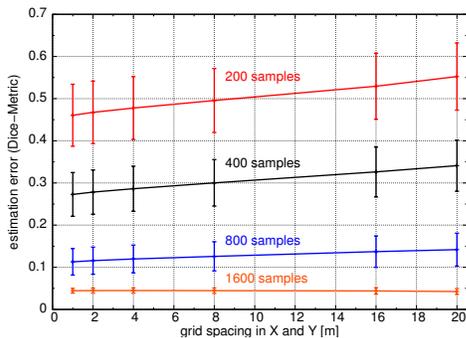
Fig. 3. Sensitivity of polygon accuracy to grid spacing

Fig. 4. Correlation of interpolation radius and polygon accuracy

a point in the center whose value is the mean value of the four vertices. By joining each vertex with the point in the center, it constructs four triangles per rectangle. The intersection of these triangles with the horizontal plane, whose height is given by the contour level, defines the segments of the contour lines.

In order to use the result of the contouring algorithm for network selection, the contour level is determined by the minimal signal strength value for which communication using the considered radio technology is possible. Although, only one contour level is used here, it is possible to generate contour lines for multiple levels. For instance, one contour level can be defined at which communication is possible, but the signal strength is low, such that a handover to this cell is not recommended. A higher contour level can be defined to generate shapes for areas having good radio conditions, in which case a handover is recommended.

C. Complexity Reduction

Vectorization yields complex shapes with a large number of vertices. The aim of complexity reduction is to transform the complex shapes into less complex shapes that are as similar as possible to the complex ones. The reduced shapes are more simple to process and consume less resources when being transmitted or stored. Complexity reduction can formally be defined as an optimization problem, which minimizes the number of vertices for a given maximal deviation of the reduced geometry from the original geometry (min-#-problem), or which minimizes the deviation for a given number of vertices (min-ϵ-problem).

In practice, complexity reduction can be achieved by heuristic algorithms. Distance-based algorithms show comparatively good results [2]. These algorithms traverse the geometries along their vertices and remove vertices that are located within a tolerance region around each visited vertex. In addition, the distance-based algorithms can be complemented or replaced by algorithms that operate on slope differences between individual line segments. The *euclidean-distance* algorithm, which is applied here, uses circles as tolerance regions.

IV. EVALUATION METHODOLOGY

A. Simulation Experiments

The proposed radio coverage estimation algorithms have been evaluated by means of simulation. Radio characteristics
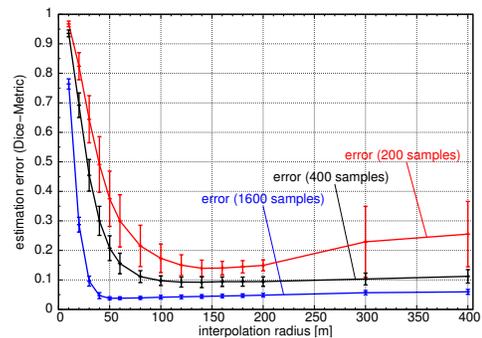
of a cell or hotspot have been modeled by the Walfish-Ikegami radio propagation model [8]. All results presented here have been computed for a model of three sector antennas with beam width $65°$, and radiation patterns according to [9]. Mobile users move around in a rectangular plane of 1000x800 m, following a random direction mobility model. Field strength measurements are collected by each terminal every 50 s. Evaluations of further scenarios as well as details of the parameterization can be found in [10].

The simulation model represents an instationary system. Therefore, all simulation experiments have been repeated multiple times with different random generator initializations to obtain statistically valid results. All plots in the following sections are given with 98 % confidence intervals.

B. Metrics

To quantify the performance of the algorithms, two metrics have been defined: The complexity of geometries is quantified by the number of their vertices. Since information systems handle geometries by indicating the vertices, this metric allows to estimate the processing, storage, and transmission costs of the corresponding data records. In order to evaluate the accuracy of the coverage estimation, the estimated geometry is compared to a reference one. To quantify the difference between the estimated and the reference geometry, the Dice coefficient [11] is used. Given a set of points A and B of the geometries, the Dice coefficient is defined by

$$C_D = \frac{2|A \cap B|}{|A| + |B|}$$

It corresponds to the ratio of the area of the intersection of the geometries and the sum of their areas. Therefore, the values of this metric can be interpreted intuitively. To indicate the estimation error, we use the value $1 - C_D$. A value of zero means that no estimation error has occurred. If the value is one, there is no similarity between the estimated geometry and the reference geometry.

V. SIMULATION RESULTS

As it has already been mentioned in section III, Fig. 2 shows the rectangular plane and the geometry of the reference geometry. The position of field strength measurements are also indicated in this plot. After interpolation and application of the
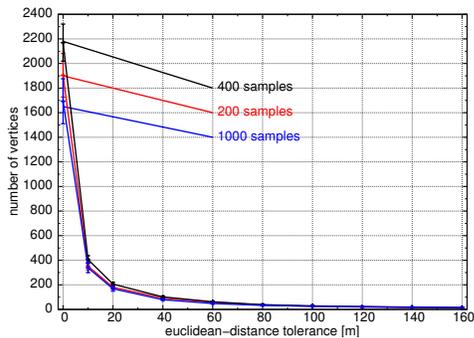
Fig. 5.   Effectivity of complexity reduction with Euclidean distance algorithm

Fig. 6.   Influence of complexity reduction on polygon accuracy

contouring algorithm, a polygon is computed, which needs to be further simplified to allow for efficient data transmission and storage. In this process, several parameters influence accuracy and size of the final data record.

A. Influence of grid spacing

For the interpolation of the measured field strength values, the spacing of the interpolation grid has an impact on processing time and accuracy. Fig. 3 shows the influence of the grid spacing parameter on the deviation of the computed polygon to the reference shape, which is directly derived from radio propagation characteristics. The error depicted on the y-axis here refers to the polygon before any complexity reduction steps have been applied. The interpolation radius of the inverse distance weighting algorithm has been set to 40 m.

In general, very fine grids tend to increase processing time, even though they do not result into significant improvement of the polygon's accuracy, as it can be observed in Fig. 3. The error in terms of the dice metric remains fairly constant over the evaluated parameter range, although the influence is slightly higher for coverage estimation with a small number of field strength measurements. For the following simulation experiments, the grid spacing has been set to 8 m.

B. Optimal choice of IDW interpolation radius

The impact of the interpolation radius is much more significant than the influence of grid spacing. Fig. 4 shows the accuracy of the computed contour line for different radii and simulation experiments with 200 to 1600 samples. The error here again refers to the polygon before complexity reduction.

For a small radius, the error is very large, given that the interpolation is not able to compensate for irregularities introduced by small regions without any field strength measurement. The resulting polygon resembles a patchwork of small coverage isles without interconnection.

An interesting property of Fig. 4 is the sharp decrease of the overall error within a small range of interpolation radii. A lower bound for the estimation error is reached for radii around 50 m to 150 m, where the corresponding polygon is characterized by a closed region which already encloses much of the reference polygon area. For a large number of measurements, the generated polygon evidently converges towards this structure much faster than in case of scarce data.
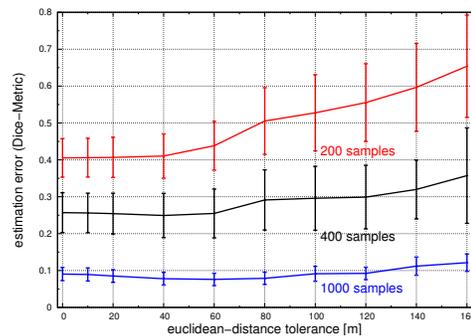
The estimation error increases again for large interpolation radii, particularly if only few field strength measurements are available. In such a configuration, dead spots and edges in the shape of the reference polygon are erroneously leveled out, as it has been mentioned in section III.

Fig. 4 indicates that the ideal interpolation radius of the inverse distance weighting algorithm depends on the number of available field strength values. It can also be concluded that in case the amount of available field strength measurements is large, the coverage estimation is less sensitive to the choice of the interpolation radius. Consequently, the interpolation radius for a given scenario can be chosen such that satisfactory results are achieved even for polygon calculation with scarce field strength data. In practice, the choice of the interpolation radius might be subject to technical constraints of the respective radio access technology and might be chosen differently for urban and rural areas. In this case, the number of required field strength measurements for a given radius has to be determined, which is further evaluated in subsection D. The absolute values of the interpolation radius presented here obviously depend on the cell size and the radio technology under investigation.

C. Trade-off between data record volume and accuracy

The polygon computed by the Bourke algorithm consists of a large number of vertices. For efficient storage and transmission, a simplification of these shapes is essential. However, an excessive reduction might result into drastic degradation of accuracy. Thus, a feasible trade-off has to be identified.

The heuristic approaches applied here lead to a substantial reduction in the number of vertices required to describe the polygon shape. Fig. 5 depicts the extent of complexity reduction by application of the euclidean distance algorithm. The x-axis indicates the maximum distance for which vertices in the perimeter of a given vertex will be discarded. For the simulation experiments shown here, an IDW interpolation radius of 40 m has been used. An increased euclidean distance results into a higher degree of simplification, respectively a higher tolerance towards the final shape of the polygon. A distance of zero in the plot denotes "no simplification", i.e., the number of vertices on the y-axis here represents the contour line computed by the Bourke algorithm. Due to the constant interpolation grid spacing for all curves in Fig. 5, there is only a marginal difference for varying numbers of
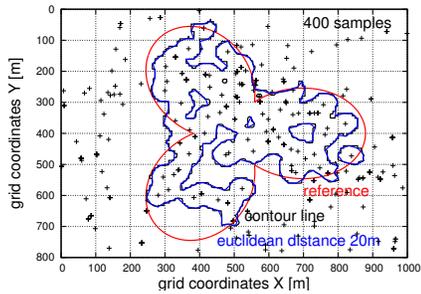
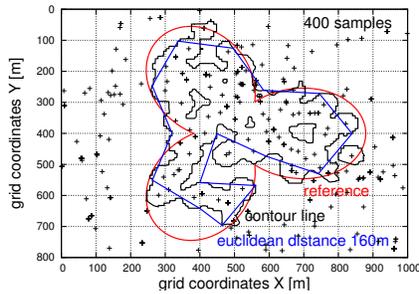Fig. 7.   Appropriate polygon simplification

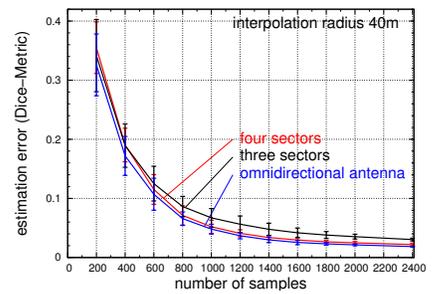Fig. 8.   Oversimplification of contour line

Fig. 9.   Different antenna characteristics

collected field strength values. Fig. 6 depicts the corresponding accuracy degradation of the estimated radio coverage for an increasing degree of simplification.

Using an Euclidean distance of 20–40 m, a reduction of the data record size by a factor of 10 to 20 can be achieved without accuracy degradation. With increasing tolerance, the number of vertices required to describe the polygon decreases at a much slower rate, such that further benefits of simplification are virtually negligible. On the other hand, the resulting error scales up for distance values larger than 60 m. Fig. 7 and 8 provide an example of an adequate, respectively excessive degree of simplification. Similar to what has been observed for the influence of grid spacing and IDW radius, increasing the degree of simplification has a larger impact the smaller the amount of available field strength measurements is.

D. Number of required field strength measurements

The results presented so far constitute an evaluation of the radio coverage estimation algorithms for a scenario of three sector antennas. In [10], several other scenarios using the same radio propagation and mobility models have been analyzed. Notably, a number of simulation experiments have been conducted for an antenna with omni-directional shape, as well as for a setup of four sector antennas in each corner of the observation area, pointing towards the area center. They exhibit the same principal behavior (see Fig. 9). Considering a Dice metric between 0.1 and 0.2 as an sufficiently accurate description of the radio coverage area, a rule of thumb for the number of required field strength measurements can be formulated. For an observation area of size $A = a \cdot b$ and interpolation radius r, the required number of measurement points N can be approximated as

$$N = m \cdot \frac{A}{\pi r^2}.$$

The multiplication factor m has been found to be in a range of 4 to 6 for reasonable interpolation radii. For an UMTS picocell or an IEEE 802.16e hotspot with a diameter of roughly 500 m and an interpolation radius of 30 m, the required number of samples will be in the order of 400. In an indoor scenario with hotspot diameters around 100 m, an interpolation radius of 10 m might be better suited. As a rough estimate, around 150 samples will be needed to map the coverage area.

The previous considerations are valid under the assumption that measurements are distributed in the plane according to a 2-dimensional Poisson process. This is a strong limitation and consequently only allows to give the order of magnitude of the required number of field strength values. The behavior of the algorithms in case of an inhomogeneous samples distribution is left for further studies.

VI. CONCLUSION

This paper studies algorithms for the conversion of field strength measurements into a description of radio coverage areas in form of polygons. The proposed method consists of three different steps: Interpolation, vectorization, and complexity reduction. The evaluation of these algorithms for different scenarios reveals that the algorithms are insensitive to the grid spacing used in the interpolation step. The radius of the inverse distance weighting interpolation depends on the respective radio technology. A larger radius performs better than a small value, given that a smaller radius can have a heavier negative impact, particularly if measurement data is scarce. Regarding the degree of simplification, there is a trade-off between accuracy and simplicity: Satisfactory results can be achieved for relatively small values of the Euclidean distance, while at the same time the accuracy of the simplified polygon shape compared to the raw data is maintained.

REFERENCES

[1] E. Gustafsson and A. Jonsson, "Always best connected," *IEEE Wireless Communications*, vol. 10, no. 1, 2003.
[2] S. Lück, M. Scharf, and G. Gil, "An architecture for acquisition and provision of hotspot coverage information," in *European Wireless*, 2005.
[3] M. Siebert, M. Schinnenburg, M. Lott, and S. Göbbels, "Location aided handover support for next generation system integration," in *European Wireless*, 2004.
[4] S. Oechsner, T. Hossfeld, K. Tutschku, and F.-U. Andersen, "Supporting vertical handover by a self-organizing multidimensional p2p overlay," in *VTC Spring*, 2006.
[5] *Draft IEEE Standard for Local and Metropolitan Area Networks: Media Independent Handover Services*, IEEE Std. 802.21, October 2006.
[6] P. A. Burrough and R. A. McDonnell, *Principles of geographical information systems*.   Oxford: Oxford University Press, 1998.
[7] P. D. Bourke, "A contouring subroutine," *BYTE journal*, June 1987.
[8] J. D. Parsons, *The Mobile Radio Propagation Channel, 2nd edition*. Wiley, 2000.
[9] B. Stockbroeckx and A. V. Vorst, "Microwave antennas at ucl: A design point of view," *Revue HF*, 2002.
[10] R. Fetscher, "Entwurf und Implementierung von Software zur Verarbeitung von Signalstärkemesswerten in Funkzellen," diploma thesis, Universität Stuttgart, IKR, 2006.
[11] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, 1945.