Radio Map Interpolation using Graph Signal Processing

Alessandro Enrico Cesare Redondi

Abstract—Interpolating a radio map is a problem of great relevance in many scenarios such as network planning, network optimization and localization. In this work such a problem is tackled by leveraging recent results from the emerging field of signal processing on graphs. A technique for interpolating graph structured data is adapted to the problem at hand by using different graph creation strategies, including ones that explicitly consider NLOS propagation conditions. Extensive experiments in a realistic large-scale urban scenario demonstrate that the proposed technique outperforms other traditional methods such as IDW, RBF and model-based interpolation.

Index Terms—Radio Map Interpolation, Graph Signal Processing, Fingerprinting Localization

I. Introduction

Many problems related to network design and optimization require to rely on a radio map that contains, for each physical location, what is the received signal strength from a particular signal emitter (base stations for outdoor environments or Wi-Fi access point in the case of indoor applications) [1]. However, the process of constructing an accurate radio map is costly and time consuming, as it requires to perform many measurements at different locations. Therefore, radio maps are generally coarsely sampled and contain measurements only in a discrete set of specific locations, typically the most application-critical ones. While such an approach can be effective in some scenarios, other applications may require a finer radio map resolution. Such a case is well represented by fingerprinting localization, a widely popular and cost-effective localization technique. In fingerprinting localization, a set of known *anchors*, such as base stations or access points, transmit/receive radio signals to/from the object to be located. The different signal strengths create a signal fingerprint which is then compared to the ones contained into a pre-stored radio map. Then, the location corresponding to the closest fingerprint in the radio map is returned as the estimated location of the target. It follows that the localization error is inversely proportional to the radio map resolution. To increase the accuracy of fingerprinting localization while at the same time keeping low measurement costs, the radio map is generally interpolated so that measurements for new locations can be created starting from the existing ones. In this paper, we study the problem of interpolating a radio map using a technique named graph-based signal interpolation [2]. Such a technique differs fundamentally from traditional local interpolation techniques generally applied to radio maps (e.g., Inverse Distance Weighting (IDW) [3], Radial Basis Functions

(RBF) [4]) as it treats the radio signal strength measurements as a signal defined over a graph. The graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is composed by nodes $\mathcal{V}$, which represent the different physical locations, and weighted edges $\mathcal{E}$, representing, e.g., the distances between different locations: the radio map can be therefore viewed as a signal defined over such a graph, where each measurement is attached to a different node. Differently from local interpolation methods, where the signal value at an unknown location is obtained as a weighted average of the signals or signal functions at neighbouring locations, the graph-based interpolation methods used in this work captures global information and predicts the value of all unknown nodes at once by selecting a solution that matches the values at known nodes. The solution is obtained leveraging not only the relationships between known and unknown nodes as done in traditional techniques such as IDW and RBF, but also the ones among known nodes yielding to higher interpolation accuracy. The rest of this paper is organized as follows: Section II reviews the basics of graph-based signal processing and describes an interpolation method for graph signals. Section III illustrates how such interpolation methods can be applied to the problem at hand, focusing in particular on different graph creation strategies. Section IV compares the proposed technique with existing methods for the problem of fingerprinting localization in a realistic large-scale outdoor urban scenario. Finally, Section V concludes the paper and describes future work.

II. Notions of Graph Signal Processing (GSP)

A. Background on graph signals

Consider an undirected, connected and weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{W})$ consisting of a set of N vertices $\mathcal{V}$, a set of edges $\mathcal{E}$ and a weighted adjacency matrix $\mathbf{W}$. Entry $w_{i,j}$ in $\mathbf{W}$ contains the weight of the link connecting nodes i and j, where $w_{i,j} = 0$ means absence of the link. A graph signal is any function $f : \mathcal{V} \rightarrow \mathbb{R}$ defined over the vertices of the graph, which can be represented as a vector $\mathbf{f} \in \mathbb{R}^N$.

Let d_i be the degree of node i, that is the sum of link weights connected to node i and define the degree matrix $\mathbf{D} = diag(d_1, d_2, \ldots, d_N)$. The normalized combinatorial graph Laplacian is defined as $\mathbf{L} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}$ with an associated set of orthogonal eigenvectors $\mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N)$ and eigenvalues $(\lambda_1, \lambda_2, \ldots, \lambda_N)$. The Graph Fourier Transform (GFT) of a graph signal $\mathbf{f}$ is defined as $\tilde{\mathbf{f}} = \mathbf{U}^{\mathsf{T}}\mathbf{f}$. Similarly to classical Fourier transform, eigenvectors and eigenvalues of $\mathbf{L}$ give a frequency interpretation of the graph signal.

The author is with the Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Piazza Leonardo da Vinci 32 - 20133, Milano, Italy. e-mail: alessandroenrico.redondi@polimi.it.

B. Interpolation of graph signals

In [5], an interpolation method based on the spectral decomposition of a graph signal is proposed. The method builds on a work by Pesenson [6] which describes how to reconstruct a bandlimited (smooth) graph signal starting only from a subset of nodes called the uniqueness set. The method starts by dividing the set of graph vertices in a known set $\mathcal{S}$, $|\mathcal{S}| = N_\mathcal{S}$ (for which measurement $(\mathbf{f})_\mathcal{S}$ are available) and unknown set $\mathcal{S}^c$, $|\mathcal{S}^c| = N_{\mathcal{S}^c}$ (for which data should be interpolated), $N = N_\mathcal{S} + N_{\mathcal{S}^c}$. The method then computes the optimal cut-off graph frequency $\omega_\mathcal{S}^*$ as the square root of the smallest singular value of the matrix $(\mathbf{L})_{\mathcal{S}^c}^2$, defined as the submatrix of $\mathbf{L}^2$ containing only the rows and columns corresponding to the unknown set $\mathcal{S}^c$. Such choice is proved in [5] to satisfy the sufficient conditions provided by Pesenson for reconstructing a graph signal from the uniqueness set. [6, Theorem 1.2 and 4.1]. Finally, reconstructed values for the unknown set are computed by least-squares projection of the known nodes onto the Paley-Wiener space of $\omega_\mathcal{S}^*$-bandlimited signals defined over the graph. Operatively, this is done according to the following steps:

1) Compute the H^* eigenvectors of $\mathbf{L}$ for which the associated eigenvalues are less than $\omega_\mathcal{S}^*$.
2) Define $\mathbf{U}_H^*$ as the matrix containing such H^* eigenvectors, $(\mathbf{U}_H^*)_\mathcal{S}$ as the submatrix containing only the rows corresponding to the set $\mathcal{S}$ and $(\mathbf{U}_H^*)_{\mathcal{S}^c}$ as the submatrix containing only the rows corresponding to the unknown set $\mathcal{S}^c$.
3) Compute $\hat{\mathbf{g}} = (\hat{\mathbf{g}})_\mathcal{S} \cup (\hat{\mathbf{g}})_{\mathcal{S}^c}$, where $(\hat{\mathbf{g}})_\mathcal{S} = \mathbf{D}^{1/2}(\mathbf{f})_\mathcal{S}$ and $(\hat{\mathbf{g}})_{\mathcal{S}^c} = (\mathbf{U}_H^*)_{\mathcal{S}^c}((\mathbf{U}_H^*)_\mathcal{S}^\mathrm{T}(\mathbf{U}_H^*)_\mathcal{S})^{-1}(\mathbf{U}_H^*)_\mathcal{S}^\mathrm{T}(\mathbf{f})_\mathcal{S}$.
4) Retrieve the interpolated signal $\hat{\mathbf{f}} = \mathbf{D}^{1/2}\hat{\mathbf{g}}$.

As pointed out in [5], the complexity of the method is $O(H^* N^2)$.

III. GRAPH CONSTRUCTION FOR RADIO MAP INTERPOLATION

The problem of radio map interpolation can be defined as it follows: assume K anchors (base stations, access points, concentrators) are deployed in the area under analysis and let $s_k(\mathbf{x}_i)$ be the measured signal strength received from the k-th anchor at location $\mathbf{x}_i = \{x_i, y_i\}$, $k = 1 \ldots K$. The radio fingerprint at location $\mathbf{x}_i$ is defined as $\mathbf{s}(\mathbf{x}_i) = \{s_1(\mathbf{x}_i), \ldots, s_K(\mathbf{x}_i)\}$. Let $\mathcal{M}$ be the set of location indexes where measurements have been taken and a signal fingerprint is available. Given $\mathbf{s}(\mathbf{x}_i)$, $i \in \mathcal{M}$, radio map interpolation estimates the unknown signal fingerprints $\hat{\mathbf{s}}(\mathbf{x}_j)$ at *new* locations $\mathbf{x}_j$, $j \in \mathcal{U}$.

In order to cast the radio map interpolation problem as a graph signal interpolation problem, one needs to construct a graph first. That means transforming the problem at hand into a set of nodes and a set of weighted links between them.

For what concerns nodes in the graph, and referring to Section II, let $\mathcal{S} = \mathcal{M}$ and $\mathcal{S}^c = \mathcal{U}$. With such an assignment nodes in the graph correspond to physical locations in the radio map, which may or may not have an associated fingerprint measurement.

The implicit assumption behind the graph interpolation method explained in Section II is that the signal to be interpolated is "smooth" over the graph, i.e., nodes close to each other (in terms of the weights of the links connecting them) should have similar signal values. In our case it is reasonable to assign the link weights $w_{i,j}$ to an inversely proportional function of the physical distance between locations i and j: two locations close in space are expected to have similar fingerprints, therefore the weight of the link connecting them should be high. Following this rationale, several options for the link weights are possible:

- *Inverse distance:* a natural choice is the inverse of the distance between the real-world locations $\mathbf{x}_i$ and $\mathbf{x}_j$.

$$w_{i,j} = \begin{cases} \frac{1}{d_{i,j}} & \text{if } d_{i,j} < d_{\text{thr}}, \, d_{i,j} \neq 0 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

The distance threshold d_{thr} controls the level of connectivity of the graph.

- *Gaussian kernel weighting function:* a common choice widely used in the field of graph signal processing is the following:

$$w_{i,j} = \begin{cases} \exp\left(-\frac{d_{i,j}^2}{2\sigma_d^2}\right) & \text{if } d_{i,j} < d_{\text{thr}} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

- *Physical distance and correlation coefficients:* in this case the edge weights are determined not only by the physical distance between nodes, but also by the correlation coefficients $\rho_{i,j}$ between the signal fingerprints $\mathbf{s}(\mathbf{x}_i)$ and $\mathbf{s}(\mathbf{x}_j)$ at those nodes. Note that $\rho_{i,j}$ can be computed only for those couples of nodes where a fingerprint is available, that is:

$$w_{i,j} = \begin{cases} \exp\left(-\frac{\tilde{d}_{i,j}^2}{2\sigma_d^2}\right) \cdot \exp\left(-\frac{(1-|\rho_{i,j}|)^2}{2\sigma_c^2}\right) & \text{if } i,j \in \mathcal{M}, \\ \exp\left(-\frac{\tilde{d}_{i,j}^2}{2\sigma_d^2}\right)^2 & \text{otherwise} \end{cases} \quad (3)$$

where $\tilde{d}_{i,j}$ is the normalized distance between i and j, and σ_d and σ_c are the parameters controlling the exponential decay rate. As in (1) and (2), $w_{i,j} = 0$ if $\tilde{d}_{i,j}$ is smaller than a normalized distance threshold $\tilde{d}_{\text{thr}}$.

- *Non-line-of-sight (NLOS) correction:* In NLOS scenarios, such as outdoor urban environments, the presence of obstacles greatly impacts on the propagation of radio signals. In such cases, the smoothness assumption may be invalid, as two close locations separated by an obstacle may receive a very different signal strength. Therefore, we propose the following correction to the edge weights in (1)–(3):

$$w_{i,j}^* = \mathcal{I}_{i,j} w_{i,j}, \quad (4)$$

where $\mathcal{I}_{i,j}$ is an indicator function equal to 1 if locations i and j are in line-of-sight and 0 otherwise. The effect of the correction is to de-activate the graph links connecting two locations in NLOS conditions. Section IV gives details on how such a correction can be operatively computed in urban scenarios.

Having defined nodes, edges and edge weights, the graph signal $\mathbf{f} \in \mathbb{R}^{K \times N_\mathcal{S}}$ is obtained with the fingerprints $\mathbf{s}(\mathbf{x}_i)$ measured over the set of known nodes $\mathcal{S}$. The method described
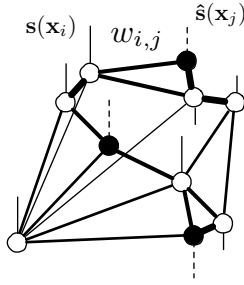
Fig. 1. Graphical representation of graph-based radio map interpolation. White nodes represent location where signal fingerprints $s(x_i)$ are available (solid lines stemming from the nodes). Black nodes represents locations where the radio map should be interpolated computing the unknown signal $\hat{s}(x_j)$ (dashed lines). The width of the links represent weights $w_{i,j}$ and is inversely proportional to the distance between the nodes.

in Section II can be used to retrieve the interpolated signal $\hat{f} \in \mathbb{R}^{K \times N_{S^c}}$, defined over the set of unknown nodes $\mathcal{S}^c$. Figure 1 provides a graphical representation of the graph construction process.

IV. EXPERIMENTAL RESULTS

We test the performance of graph-based radio map interpolation in a real-life large-scale outdoor urban scenario. Specifically, we leverage the availability of an already deployed city-wide low-power wireless sensor network composed of more than 2500 sensor nodes scattered in a rectangular area of about 4 Km². The area under interest is located in the historical center of a city in northern Italy and therefore characterized by severe NLOS propagation conditions. Sensor nodes operate the Wireless M-Bus standard (EN 13757-4:2013) at the physical and data link layer, using the N mode (long range, narrow-band) at 169 MHz. The nodes are deployed statically, geo-tagged with GPS and periodically broadcast packets containing application specific sensor measurements. Such messages are received by K data concentrators deployed in the area. The coordinates x_k of each concentrator are also known. For each received message, the k-th concentrator estimates the Received Signal Strength (RSS) $s_k(x_i)$ from the i-th node. This allows to construct, for each node location, a fingerprint signal $s(x_i)$ which contains the RSSs computed by the different concentrators. Note that a concentrator can be located outside the communication range of a particular node, causing the measurement for that node to be missing. We rely here on a simple heuristic that replaces missing values in the fingerprints with a small default value below the sensitivity threshold for the technology under consideration, equal to -148 dBm.

A. Computing the NLOS indicator

To estimate the NLOS indicator function in (4) we leverage the OpenStreetMap database, which provides free worldwide geospatial data including street maps. In particular, we rely on the Open Source Routing Machine (OSRM) project [7], a service that exploits the OpenStreetMap database to compute routing directions for several types of transportation modes (cars, bikes, walking, etc.). Given two real-world nodes locations x_i and x_j, the OSRM service is queried asking for

walking directions and the corresponding output is processed through a parsing script. The indicator $\mathcal{I}_{i,j}$ is set to 1 if a straight walking path exists between the two locations. Otherwise, if any obstacle blocks the path between x_i and x_j, $\mathcal{I}_{i,j}$ is set to 0. Note that the complexity of the method can be reduced focusing only on those location pairs (i,j) for which $d_{i,j} < d_{\text{thr}}$.

B. Radio Map Interpolation

To evaluate the quality of the proposed graph-based radio map interpolation, the original set of fingerprints $s(x_i)$ is divided in a training set (80% of the measurements) and a test set (20% of the measurements). The training set is used as the set of known measurements $\mathcal{M}$, while the test set is used to assess the performance of the method. Additionally, we also estimate the fingerprints of the locations in the test set using several traditional radio map interpolation techniques such as Inverse Distance Weighting (IDW) [3], Radial Basis Functions (RBF) [4] and Model-based Interpolation (MBI) [8]. IDW estimates the signal at an unknown location as a weighted average of the signals at known locations. RBF approximates the fingerprints as sum of weighted radial basis functions centered at known locations. MBI estimates the signal fingerprint using the log-distance path-loss model, whose parameters are learnt on the training data. For the graph-based interpolation, we distinguish between the three different options of graph construction, namely Inverse Distance (GBI_d), Gaussian kernel (GBI_e) and distance with correlation coefficients (GBI_ρ) with or without NLOS correction. All tested methods are sensible to the input parameters: IDW requires to set the power to the distance weights p, RBF requires to set the standard deviation σ_r of the Gaussian basis functions and graph-based techniques require to set the distance threshold d_{thr} and the parameters of the Gaussian functions σ_d and σ_c. To perform a fair comparison we run extensive experiments and we retrieve the parameters for which the best results were obtained through linear search. For IDW this corresponds to setting $p = 2$. For RBF, we set $\sigma_r = 10^3$ and for GBI techniques we set $d_{\text{thr}} = 100$m, $\sigma_d = 35$ and $\sigma_c = 1$.

For each interpolation technique, we compute the root mean squared error (RMSE) between the estimated and actual fingerprints, in dBm. The process is repeated 5 times, each time changing the training and the test set, according to 5-fold cross validation, and results are averaged.

Table I reports the RMSE obtained by the different methods. As one can see, the model-based interpolation is the worst-performing among all tested techniques. This is not surprising considering that the simple log-distance path-loss model does not capture the complex NLOS characteristics of the urban propagation environment, including reflection, diffraction and scattering. IDW and RBF perform almost at par, while graph-based interpolation is the best-performing method. Among the different tested graph construction techniques, the one using both pyhsical distances between locations and correlation co-efficients between the fingerprints at known locations performs the best. Moreover, the proposed NLOS correction allows to further improve the performance.

TABLE I
RMSE FOR THE DIFFERENT INTERPOLATION TECHNIQUES [dBm]

IDW	RBF	MBI	GBI_d	GBI_d-NLOS	GBI_e	GBI_e-NLOS	GBI-ρ	GBI_ρ-NLOS
9.82	9.87	11.34	9.56	9.41	9.53	9.38	9.48	**9.31**

TABLE II
SUMMARY OF LOCALIZATION ERRORS [M]

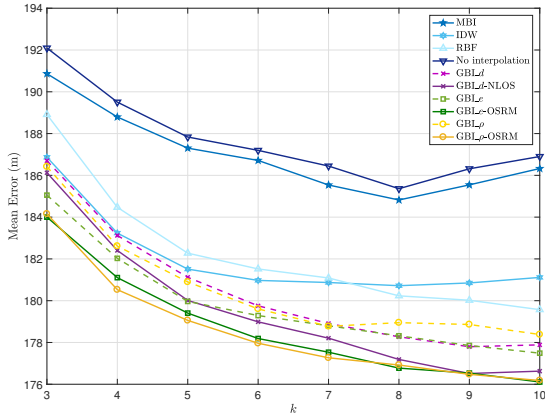|	None	IDW	RBF	MBI	GBI_d	GBI_d-NLOS	GBI_e	GBI_e-NLOS	GBI_ρ	GBI_ρ-NLOS
min	54.5	47.8	45.3	53.8	50.8	50.2	46.7	46.1	47.8	**42.7**
max	478.0	472.3	465.7	467.1	469.9	466.6	471.8	468.0	472.5	**464.2**
avg	187.9	182.8	181.2	187.1	179.8	178.3	179.7	177.8	179.9	**177.7**
std	75.3	71.1	70.7	72.6	71.1	70.7	70.6	70.7	70.1	**69.2**

Fig. 2. Average WKNN localization error for the different techniques

C. Fingerprinting Localization

In a second experiment, we compare the performance of radio map interpolation techniques when applied to fingerprint localization. The baseline technique used for the task at hand is the weighted k nearest neighbor (WKNN) algorithm: given a fingerprint captured at an unknown location and a database of pre-stored fingerprints in known locations, the WKNN computes the estimated location as a weighted average of the locations corresponding to the k nearest fingerprints in signal space. Weights are chosen to be inversely proportional to the squared euclidean distance between fingerprints. For distance computation, only the elements of the fingerprints containing valid values (different from -148 dBm) are considered.

The original set of available fingerprints is divided in a training set and test set in the same way as done in Section IV-B. As a first test, fingerprint localization is performed without radio map interpolation: the training set is used as fingerprints database, and localization is performed for the samples in the test set. Then, interpolation is performed using the aforementioned interpolation techniques. The 4 Km^2 area is spatially sampled with a step of 50 m, producing a grid of 40×40 new locations. For each location $\mathbf{x}_j$ in the grid, a fingerprint is interpolated and added in the database. Finally, localization is performed for each sample in the test set and for each technique, the average location error is computed. Again, 5-fold cross validation is performed and the corresponding

errors are averaged. Figure 2 shows the average error of the different techniques when varying the number of neighbors k. Fingerprint localization without radio map interpolation is the worst-performing technique, as expected, and the model-based interpolation does not improve much its performance. Better results can be obtained with RBF and IDW interpolation. The best performance is obtained by the proposed graph-based interpolation approach. The three tested graph-construction techniques perform almost at par when used without the NLOS correction. When the correction is applied, an increase in the localization performance is observed, with GBI_e and GBI_ρ performing slightly better than GBI_d. The details of the localization errors are reported in Table II for $k = 10$.

V. CONCLUSION

In this letter, a graph signal processing based technique for interpolating radio maps is proposed. Different graph construction strategies are tested and experimental evaluations with realistic data in a urban NLOS environment show that they provide higher interpolation accuracy compared to existing methods. Future works will focus on other radio technologies and on indoor environments.

REFERENCES

[1] H. B. Yilmaz, T. Tugcu, F. Alagoz, and S. Bayhan, "Radio environment map as enabler for practical cognitive radio networks," *IEEE Communications Magazine*, vol. 51, no. 12, pp. 162–169, 2013.

[2] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 83–98, 2013.

[3] S.-P. Kuo and Y.-C. Tseng, "Discriminant minimization search for large-scale rf-based localization systems," *IEEE Trans. on mobile computing*, vol. 10, no. 2, pp. 291–304, 2011.

[4] J. Krumm and J. Platt, "Minimizing calibration effort for an indoor 802.11 device location measurement system," *Microsoft Research, Nov.*, 2003.

[5] S. K. Narang, A. Gadde, and A. Ortega, "Signal processing techniques for interpolation in graph structured data," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on.* IEEE, 2013, pp. 5445–5449.

[6] I. Pesenson, "Sampling in paley-wiener spaces on combinatorial graphs," *Trans. of the American Mathematical Society*, vol. 360, no. 10, pp. 5603–5627, 2008.

[7] D. Luxen and C. Vetter, "Real-time routing with openstreetmap data," in *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, ser. GIS '11. New York, NY, USA: ACM, 2011, pp. 513–516. [Online]. Available: http://doi.acm.org/10.1145/2093973.2094062

[8] M. Lee and D. Han, "Voronoi tessellation based interpolation method for wi-fi radio map construction," *IEEE Communications Letters*, vol. 16, no. 3, pp. 404–407, 2012.