This is the author's accepted version of the article. The final version published by IEEE is O. E. Martínez-Durive, S. Bakirtzis‡, C. Ziemlicki and M. Fiore, "Demonstrating Deep Learning-based Spatial Diffusion," IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), 2025, pp. TBP, doi: TBP.

Demonstrating Deep Learning-based Spatial Diffusion

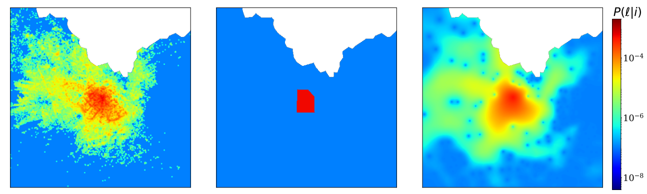
Orlando E. Martínez-Durive*†, Stefanos Bakirtzis‡, Cezary Ziemlicki¶ and Marco Fiore*
*IMDEA Networks Institute, Spain, †Universidad Carlos III de Madrid, Spain, ,‡University of Cambridge, United Kingdom,
¶SENSE / Orange Innovation, France
{orlando.martinez, marco.fiore}@imdea.org, ssb45@cam.ac.uk, cezary.ziemlicki@orange.com

***Abstract*—Metadata geolocation, *i.e.*, mapping information collected at a cellular Base Station (BS) to the geographical area it covers, is a central operation in producing statistics from mobile network measurements. This task requires modeling the probability that a device attached to a BS is at a specific location, and it is currently accomplished via simplistic approximations based on Voronoi tessellations. However, Voronoi cells exhibit poor accuracy compared to real-world geolocation data, which can reduce the reliability of downstream research pipelines. To overcome this limitation, DEEPMEND proposes a new data-driven approach relying on a teacher-student paradigm that combines probabilistic inference and deep learning. Similarly to other benchmarks, DEEPMEND can produce geolocation maps using only the BS positions, yielding a 56% accuracy gain compared to Voronoi tessellations. Our demonstrator will show visual and qualitative comparisons between DEEPMEND and several competitor approaches, allowing users to explore BS deployments from different geographical regions and operators.**

I. Background and Motivation

The metadata collected in operational mobile networks provides rich information about the movement, communication patterns, activities, and interests of large populations of subscribers, with high spatiotemporal resolution and at broad geographical scales. Therefore, such data has been used in various scientific fields, including networking, demography, geography, sociology, and epidemiology, enabling dependable quantitative analyses of phenomena that can only be studied qualitatively otherwise. In all these studies, the metadata is usually geo-referenced at the level of the individual Base Station (BS). In other words, all metadata records (*e.g.*, the traffic generated by a device or the total demand of all devices for a given service) are associated with the geographical site of BS that serves the device at each time instant. Yet, network analyses building on such metadata need to be carried out over the continuous space of the target region and not at the discrete BS locations only. This makes the mapping of BS-referenced data to the underlying geographical space an essential step in any mobile network metadata processing pipeline and calls for *spatial diffusion* models that describe *the probability that a device attached to a BS is at a specific geographical location served by that BS*.

Accurate spatial diffusion models can be directly obtained from client-side measurements or estimated by post-processing the results of radio propagation solvers. However, these approaches are not viable in academic research, as they require massive measurement campaigns or access to confidential information about the network infrastructure that is typically not available. As a result, the vast majority of the literature has relied on simplistic approximations of spatial diffusion based on Voronoi tessellations of the space, assuming that the



(a) Ground truth (b) Voronoi (c) DEEPMEND

Fig. 1: Spatial diffusion of one real-world BS, obtained from (a) elaborate post-processing of radio propagation simulations by a network operator based on complete RAN information, (b) Voronoi tessellation, (c) our proposed DEEPMEND. Colors denote diffusion probabilities over space. From [1].

metadata associated with one BS is uniformly diffused over the surface of its Voronoi cell, as illustrated in Figure 1b. Evidently, this approximation poses strong performance limitations and yields substantial discrepancies when compared to ground truth data, as shown in Fig. 1a; yet, many of the works in the literature rely on this model due to its simplicity. Indeed, computing a Voronoi tessellation is a straightforward operation that only requires the geographical coordinates of the BSs, which are typically provided to academic researchers along with the metadata.

An alternative is a probabilistic density model that estimates spatial diffusion via probability density inference [2]. That allows partially overcoming the limitations of Voronoi cells by generating a statistical representation of the location of UE given solely its distance from the BS of attachment, computing an approximate probabilistic spatial diffusion of the BS. This model results in a smoothed Voronoi cell, controlled by a power attenuation parameter that needs to be properly configured through field measurements. An in-depth discussion on alternative approaches can be found in [1].

II. Proposed Solution

DEEPMEND [1] aspires to change the status quo and democratize dependable spatial diffusion by proposing a model that can combine the minimal input requirement of a Voronoi tessellation (*i.e.*, the sole locations of BSs, making the model accessible to any academic researcher) with the high accuracy granted by a full-fledged modeling process based on operator-proprietary information.

To this end, we use a rich set of spatial diffusion maps from $5,947$ outdoor BSs located in urban, suburban, and rural areas of France, making up the radio access network of Orange, a major local operator. The spatial diffusion dataset is provided directly by the network operator, and it is derived from realistic distributions of the *association probability*, $P(\ell|i)$,

©2025 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

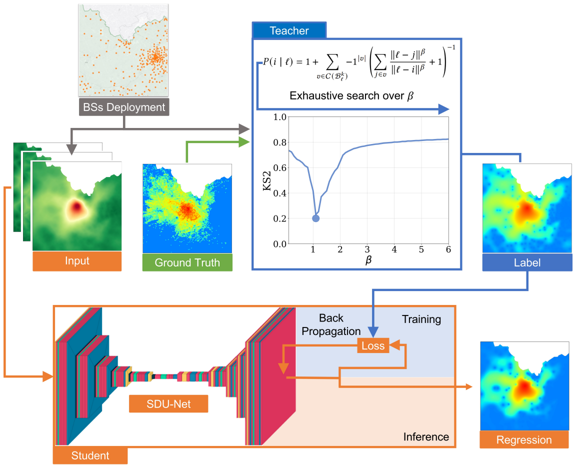


Fig. 2: Diagram of the teacher-student workflow of DEEPMEND, with inputs, outputs, and interactions. From [1].

which describe the probability that a UE attached to the i-th BS is located at position ℓ. For each BS, the probability of association over a $3,600$ km^2 area is in the form of a 600×600 regular grid, with the BS located at the center of the grid and each grid element ℓ matching a 100×100 m^2 physical area.

These voluminous data are used to develop DEEPMEND, a deep learning-based spatial diffusion model [1] that exploits the concept of *knowledge distillation* [3]. Specifically, knowledge distillation involves a complex and computationally expensive model, the *teacher*, to derive soft labels that are then used to train a simpler model, the *student*. In our case, we consider as a teacher the model of [2], with its hyperparameter optimally selected via an exhaustive search, whilst the student is a deep convolutional encoder-decoder. The general concept of DEEPMEND is shown in Figure 2. As can be seen, the teacher model processes the BS locations and uses an exhaustive search to determine the hyperparameter of [2] which minimizes the Kolmogorov–Smirnov test [4] error between the ground truth and the emerging model-based [2] spatial diffusion distribution. That corresponds to a *soft label*, *i.e.*, a high-fidelity representation of the ground truth as shown in Fig. 1a, which is though easier to be learned by the student than the ground truth itself. Finally, the student model uses these soft labels to understand how to transform simple BS location information into realistic spatial diffusion maps for each BS. More details on the performance and scalability of DEEPMEND can be found in [1].

III. DEMONSTRATOR

Our demonstrator is a web app that allows the user to visually and quantitatively assess the performance of DEEPMEND among other geolocation map approaches. The dashboard of our demonstrator is shown in Fig. 3. The web app builds on *React.js* for the User Interface, *Mapbox* for map visualizations, and *Python* for data load and inference tasks.

The user can select the country, city, and mobile operator of interest in the menu at the top. Currently, four countries are available: United Kingdom, Ireland, Belgium, and France. In the latter case, we offer more than 10 cities to visualize, including Paris. Operators include Orange, Proximus, Vodafone, Three, Free, and SFR. In the same menu, it is possible to select the model to infer geolocation maps: this list includes our proposed solution DEEPMEND together with other benchmarks such as Voronoi and VoronoiBoost.

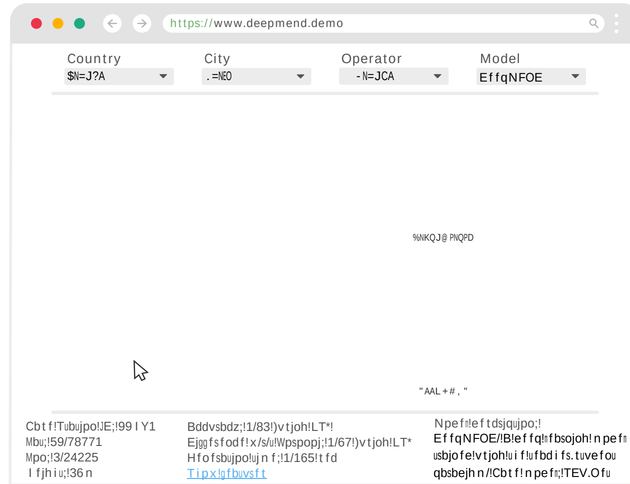


Fig. 3: Web app demonstrator of DEEPMEND.

After selecting the country, city, and operator, the app shows the corresponding base station deployment in the left panel. Then, the user can zoom in and out or move over the map; details will be displayed in a popup while the user hovers over a base station. When a base station is selected, the ground-truth geolocation map (if available) is shown in the top right panel. Based on the chosen model, the app requests the inferred geolocation map using the Python API, which is displayed in the bottom right panel.

The bottom panel displays relevant information about the selected base station, such as lat, long, and height. Furthermore, it also shows information about the selected model, such as inference time, accuracy with respect to the ground truth (if available), and error metrics compared to the corresponding Voronoi cell and the other benchmarks. There is also an option to show the features fed to the model.

ACKNOWLEDGMENT

This research was supported by NetSense Talent Attraction grants 2019-T1/TIC-16037 and 2023-5A/TIC-28944, funded with public funds from Comunidad de Madrid, as well as by ORIGAMI project (GA 101139270) funded by SNS JU and the European Union. Stefanos Bakirtzis' work is supported by the Foundation for Education and European Culture.

REFERENCES

[1] O. E. Martínez-Durive *et al.*, "Deepmend: Reliable and scalable network metadata geolocation from base station positions," in *IEEE Conference on Sensing, Communication, and Networking (SECON)*, 2024.

[2] V. A. Traag *et al.*, "Social event detection in massive mobile phone data using probabilistic location inference," in *Third International Conference on Privacy, Security, Risk and Trust*, 2011.

[3] G. Hinton *et al.*, "Distilling the knowledge in a neural network," in *NIPS Deep Learning and Representation Learning Workshop*, 2015.

[4] G. Fasano and A. Franceschini, "A multidimensional version of the kolmogorov–smirnov test," *Monthly Notices of the Royal Astronomical Society*, 1987.