Towards Learning an Optimal Metric for Fingerprint-based Localisation

Nicola Saccomanno
University of Udine, Italy
nicola.saccomanno@uniud.it

Andrea Brunello
University of Udine, Italy
andrea.brunello@uniud.it

Angelo Montanari
University of Udine, Italy
angelo.montanari@uniud.it

ABSTRACT

Fingerprinting is a common localisation approach that often estimates a device's position by comparing an observed vector to a set of prior vectors labelled with a ground truth location, typically using methods like k-NN. In Wi-Fi fingerprinting, these vectors represent visible access points and their signal strength. Thus, the choice of metric to compare the fingerprints is crucial. In this work, we discuss our main findings regarding the extent to which metrics in the fingerprint vector space preserve relationships among locations in the 2D/3D geometric/real world. In summary, traditional metrics are not optimal on their own, and while combining them into a learned meta-metric offers slight improvements, deep metric learning, i.e., learning similarities in an end-to-end fashion with deep neural networks, appears much more effective. However, this approach has its challenges given that in the literature the problem has only been formulated for binary similarities rather than continuous ones.

ACM Reference Format:
Nicola Saccomanno, Andrea Brunello, and Angelo Montanari. 2023. Towards Learning an Optimal Metric for Fingerprint-based Localisation. In *The 29th Annual International Conference on Mobile Computing and Networking (ACM MobiCom '23), October 2–6, 2023, Madrid, Spain.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3570361.3615739

1 INTRODUCTION

Wi-Fi fingerprinting is the prevalent method for indoor positioning due to the ubiquity of Wi-Fi access points (APs) and the fact that specialized equipment is not needed. Only the Received Signal Strength (RSS) from APs, not their exact locations, is necessary for accurate position information.

In fingerprinting, a training set of n examples is definable as a collection of (fingerprint, location) pairs $\mathcal{P} = \{(\mathbf{x}, \mathbf{y})_i | \, 1 \leq i \leq n, \mathbf{x} \in \mathbb{R}^m, \mathbf{y} \in \mathbb{R}^p\}$, where m is the number of available APs and p is the number of dimensions used to represent the locations (e.g., $p = 3$ in a 3D-based modelling). Each vector in $\mathbb{R}^m$ contains the RSSs of all the APs that are visible at the specific location. We can define the localisation task as a (learnable) function $l_\theta : \mathbb{R}^m \rightarrow \mathbb{R}^p$. The goal is to find the parameters θ that allow to obtain the best possible solution of the problem, also generalisable to newly observed fingerprints. Classical deterministic algorithms, like (k-)Nearest Neighbour, match a new fingerprint to previously collected ones, estimating its position based on the most similar instances' coordinates, according to a given metric.

Existing literature has largely emphasized the importance of metric choice in achieving accurate positioning (e.g., [7]). However, the focus has almost entirely been on the relationships between the metrics and the resulting position estimation error. Moreover, in practice, the Euclidean distance is frequently used as fingerprinting metric, although empirically recognised as sub-optimal in several scenarios.

2 FINGERPRINT SPACE PROPERTIES

The perspective we followed in our work is radically different, as we aim to investigate the degree to which classical metrics (working in $\mathbb{R}^m$) capture the spatial relationships (in $\mathbb{R}^p$) among the locations associated with the fingerprints. Thus, we treat the problem separately from position estimation.

Indeed, path loss models allow to estimate the distance from a single AP in a Line-Of-Sight (LOS) scenario. However, when dealing with fingerprinting the situation is radically different because: (i) it is unrealistic to assume LOS in indoor settings; (ii) the fingerprint is composed of multiple APs; and, (iii) the exact position of each AP is not known. All these elements contribute to the impossibility of deriving a precise mathematical formulation of a metric for fingerprinting. Also, the high dimensionality and the sparsity of the problem (scenarios with 500 or 1000 detectable APs are common, i.e., $m = 500$ or $m = 1000$) intrinsically makes classical metric less meaningful due to the curse of dimensionality.

Thus, a systematic analysis was performed in [4] to assess the ability of individual metrics, when applied in the fingerprint vector space, to characterise real-world geometric distances. The study considered 16 datasets, multiple metrics, fingerprint normalisation strategies, and problem granularities (i.e., restricting to individual floors/buildings or considering the entire premises). Precisely, the aim was

ACM MobiCom '23, October 2–6, 2023, Madrid, Spain

ACM ISBN 978-1-4503-9990-6/23/10.
https://doi.org/10.1145/3570361.3615739

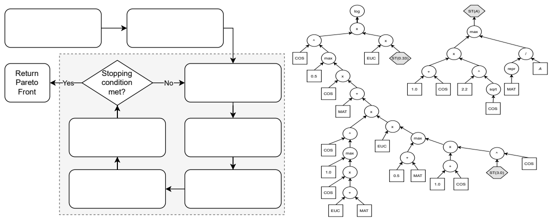


Figure 1: Meta-metric learning via GP ([1], ©Elsevier).

to establish the extent of the correlation between fingerprint and real-world location distances. It emerged that all the metrics had a performance far from the optimal case in all the considered scenarios. Nevertheless, it was also observed that some (e.g., cosine distance) possess better properties than others, and that the metric behaviours are heterogeneous and seemingly linked with the environment.

3 META-METRIC LEARNING

Inspired by the heterogeneous behaviours of metrics, a key question emerged: is it possible to combine classical metrics to leverage their unique strengths and characteristics?

To such an extent, in [1], an approach to learn a meta-metric, i.e., a function whose inputs are the classical metrics' outputs for a given pair of fingerprints, is proposed. The problem is framed as a symbolic regression one, where both the model structure as well as the parameters are learned, so to have maximum flexibility at the expense of a larger search space. The solution is achieved through genetic programming (GP), which is very often employed to solve symbolic regression tasks. The objective of the algorithm is to maximise the correlation at the centre of the previous study. Figure 1 reports the learning schema and the meta-metric.

The outcome is that the meta-metric, when learned on a single dataset and tested on all the other 15 ones, proved to be extremely generalisable. The analyses were done following the same protocol as the previous study, and resulted in a statistical improvement, although marginal and still far from the optimal scenario. Still, the meta-metric, when used in conjunction with k-NN to solve a positioning estimation task, resulted to be in par with or superior to all the single metrics taken in isolation, also on the datasets not considered at training time. Such an extremely positive and unexpected result led to two considerations. First, it appears that trying to maximise the correlation between distances measured in the fingerprint space and those in the real-world one is a good proxy task to obtain, in the end, a function that works well also for position estimation (note that the two objectives are radically different). Second, the meta-metric could represent an off-the-shelf solution for position estimation. For example, it could act as a universal baseline for indoor positioning, removing the need to determine the most suitable metric in each considered setting.

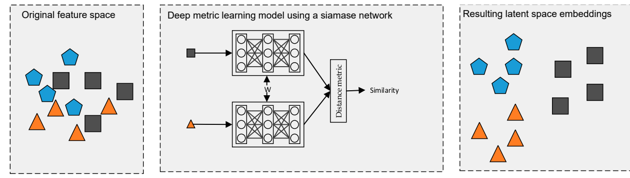


Figure 2: DML general workflow (adapted from [2]).

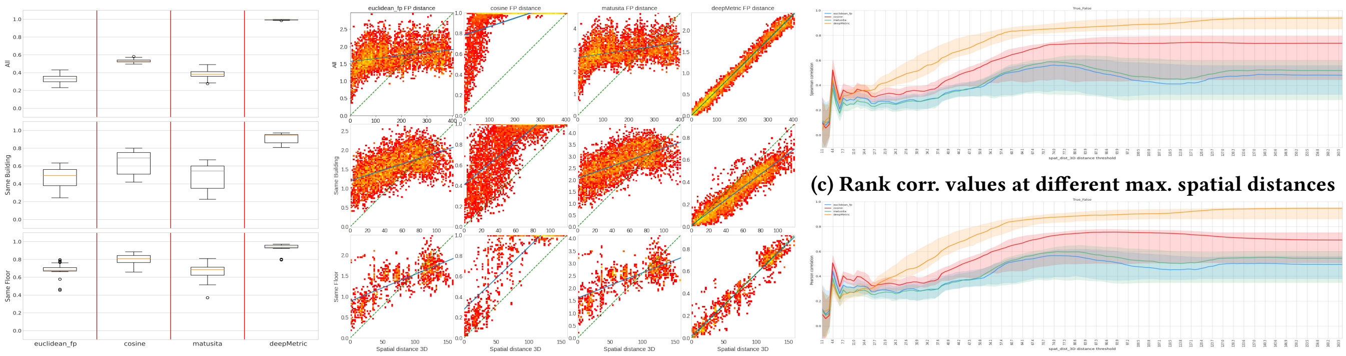
4 DEEP METRIC LEARNING

A limitation of the previously described study is that the ability to solve the target task is likely to be constrained by the considered representation, i.e., by working with the output of individual metrics instead of on the fingerprints directly. Richer representations could be obtained relying on deep metric learning (DML), which produces similarity metrics in an end-to-end fashion with deep neural networks.

In its easiest, binary formulation, the idea behind DML (intuition in Figure 2) is to learn an embedding function $\phi_\theta : \mathcal{X} \to \mathcal{Z}$ from the feature space ($\mathcal{X}$) to a new latent one ($\mathcal{Z}$) in such a way that elements deemed similar according to a given similarity function (often evaluated over $\mathcal{X}$'s labels $\mathcal{Y}$), i.e., belonging to the same class (positive pair), are mapped closer in $\mathcal{Z}$ than those considered to be dissimilar, i.e., belonging to two different classes (negative pairs). Note that, although there might be multiple classes, the semantic of the similarity over the labels is binary, as the aim is to distinguish whether two (or more) objects are the same or not. A well-known and widely adopted loss function for DML is the triplet loss [5].

In principle, such a technique can be employed to capture any kind of semantic similarity between labels, delivering great results in many different applications, such as face recognition, person re-identification, zero-shot and self-supervised learning. However, most of the literature has been focusing on computer vision and binary labels over pairs of elements, simply assessing whether they are similar or not [3]. The main reason is that, if continuous labels are instead considered, the semantics associated with their similarities is far more complex, as it involves both the concepts of ranking and proportionality. In such a scenario, when comparing elements in the latent space, a good approach should *(i)* preserve their relative ordering, and *(ii)* produce distance values that resemble some properties observed among distances computed in the label space. Moreover, the problem is different when compared to the binary case, as, e.g., the notions of negative and positive elements for tuples construction do not apply anymore. Defining a threshold for binarisation is, in principle, an approach to reduce the continuous case to the binary one, but such a definition is domain and application-dependent, thus complex and often rather ineffective.

Focusing on fingerprinting, the goal is to obtain distances in the latent space that are proportional to those calculated among the labels, i.e., on the real-world locations. A solution



(a) Correlation performance **(b) Distance functions behav. at 3 granularities** **(c) Rank corr. values at different max. spatial distances** **(d) Corr. values at different max. spatial distances**

Figure 3: Preliminary results of proportional DML: canonical metrics vs learned one.

to achieve this is to rely on distance ratios, requiring that:

$$\frac{s(\mathbf{y}_i, \mathbf{y}_j)}{s(\mathbf{y}_k, \mathbf{y}_h)} = c \cdot \frac{\| \phi_\theta(\mathbf{x}_i) - \phi_\theta(\mathbf{x}_j) \|_2^2}{\| \phi_\theta(\mathbf{x}_k) - \phi_\theta(\mathbf{x}_h) \|_2^2},$$

where $s : \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}$ calculates the distance (or similarity) between pairs of elements i, j by looking at their labels $\mathbf{y}_i$ and $\mathbf{y}_j$ (e.g., for fingerprinting, the Euclidean distance), ϕ_θ is a parameterized neural network $\phi_\theta : \mathbb{R}^m \to \mathbb{R}^z$ ($\mathbb{R}^z$ is the latent space), c is a scaling factor, and $i, j, k, h \in \{1, \ldots, n\}$.

In Figure 3, we report some preliminary results applied to the UJIIndoorLoc dataset's test set [6], following the same evaluation protocol as the previous two experiments. We can see that there is an overall improvement in the quantitative outcomes, with an average Pearson correlation higher than the classical metrics (Figure 3a). This result is also confirmed qualitatively by Figure 3b, where the points lie in a line following the optimal case scenario. However, there is still a major challenge to be addressed. Looking closely at both Figure 3c and Figure 3d (resp., rank and linear correlation), it can be seen that the performances for short distances, i.e., the left-hand side of the figures, are lower than those of the contenders. Other than making Figure 3a and Figure 3b overoptimistic, this shortcoming must be solved to learn an optimal metric, as short distances are the most informative from the positioning point of view. The causes of this phenomenon are probably diverse and still under investigation. Explanations include: an uneven distribution of the training samples used to train the deep learning model; and, an inherent difficulty of capturing distance relationships among fingerprints that are very close, both with respect to their position labels as well as their original feature representations, thus proportionally more affected by signal perturbations.

5 USEFULNESS OF AN OPTIMAL METRIC

We conclude by pointing out why being able to learn an optimal metric for fingerprinting matters. It is well-known that the performance of fingerprinting is linked with the sampled points for which the ground truth is known (composing the so called radio-map). The higher the number of points, the finer the sampling granularity, the lower the chance of committing a large positioning error. However, it is well recognised that collecting a radio-map is a time and money-consuming task. Also, radio-maps need to be periodically maintained, sampling and labelling new points over time, even after the deployment, to cope with environmental changes. An optimal metric would allow to reduce the number of fingerprints needed to build the radio-map, thanks to the better distance estimation. Moreover, additional cost savings may come by framing the meta-metric learning task as a weakly/semi-supervised one: indeed, learning a metric in our setting requires knowing just the spatial distance between two fingerprints. Thus, instead of relying on a professionally built radio-map, a crowdsourcing approach could be used to assemble the required training dataset: each user simply walks around a building and, meanwhile, its device periodically collects RSS data together with the inertial measurement unit recordings between them, so to obtain coarse spatial distances directly. Finally, we highlight that the approach for learning on continuous similarities is rather general and thus may be interesting per se for the machine learning community, beyond the fingerprinting domain.

REFERENCES

[1] Andrea Brunello et al. 2022. A genetic programming approach to WiFi fingerprint meta-distance learning. *Pervasive Mob. Comput.* 85 (2022), 101681.
[2] Mahmut Kaya et al. 2019. Deep Metric Learning: A Survey. *Symmetry* 11, 9 (2019).
[3] Sungyeon Kim et al. 2019. Deep Metric Learning Beyond Binary Supervision. In *IEEE CVPR*. 2288–2297.
[4] Nicola Saccomanno et al. 2021. What you sense is not where you are: On the relationships between fingerprints and spatial knowledge in indoor positioning. *IEEE Sens. J.* 22, 6 (2021), 4951–4961.
[5] Florian Schroff et al. 2015. FaceNet: A unified embedding for face recognition and clustering. In *IEEE CVPR*. 815–823.
[6] Joaquín Torres-Sospedra et al. 2014. UJIIndoorLoc: A new multi-building and multi-floor database for WLAN fingerprint-based indoor localization problems. In *IPIN 2014*. IEEE, 261–270.
[7] Joaquín Torres-Sospedra et al. 2015. Comprehensive analysis of distance and similarity measures for Wi-Fi fingerprinting indoor positioning systems. *Expert Syst. Appl.* 42, 23 (2015), 9263–9278.