



HAL
open science

Spatio-temporal analysis of mobile service consumption for social signature clustering

Zhaobo Hu, Chuan Li, Vincent Gauthier, Hassine Moun gla, Miguel Nunez del
Prado Cortez

► **To cite this version:**

Zhaobo Hu, Chuan Li, Vincent Gauthier, Hassine Moun gla, Miguel Nunez del Prado Cortez. Spatio-temporal analysis of mobile service consumption for social signature clustering. *Network Mobility (netmob)*, Oct 2023, Madrid, Spain. . hal-04289452

HAL Id: hal-04289452

<https://hal.science/hal-04289452v1>

Submitted on 16 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

SPATIO-TEMPORAL ANALYSIS OF MOBILE SERVICE CONSUMPTION FOR SOCIAL SIGNATURE CLUSTERING

LABS

s@movar

LIPADE
Laboratoire d'Informatique Paris Descartes

Université
Paris Cité

LA BANQUE
MONDIALE

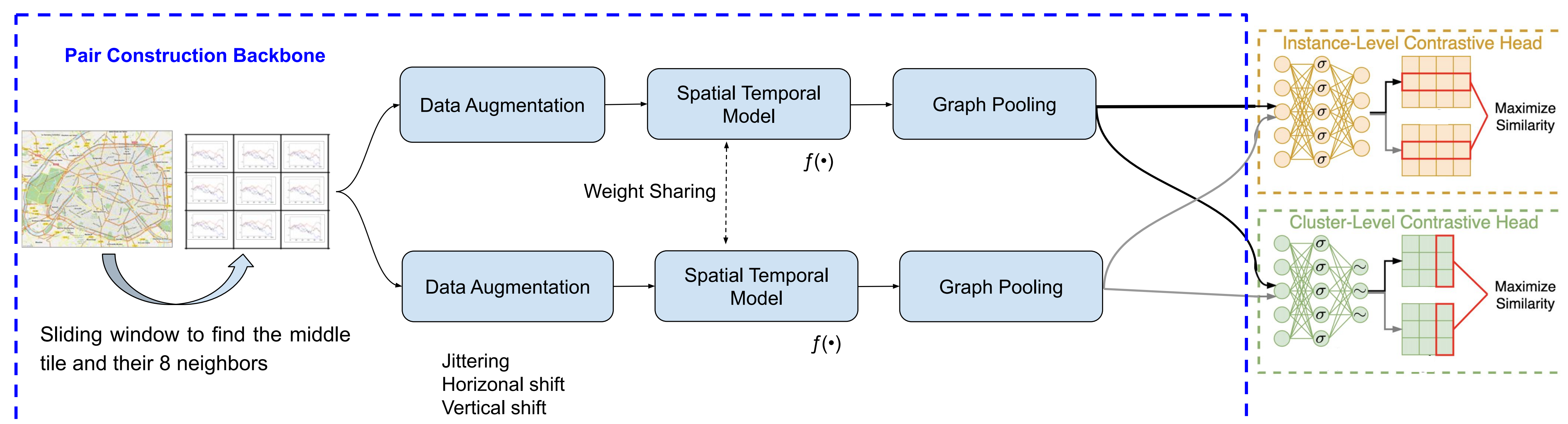
MOTIVATIONS

Mobile phone metadata is now widely used to extract socio-economic activity metrics [1,2] for cities or regions at scale. Properly used, this data can provide unique insights into downstream tasks and business value. We propose a novel scalable method for clustering urban areas based on the spatio-temporal characteristics of mobile traffic data. The development of deep learning techniques makes deep time series clustering feasible. Our approach utilizes deep learning and contrastive learning methods, including contrastive clustering and spatio-temporal model, where the former provides discriminative clusters and the latter provides spatio-temporal correlation between neighboring regions, respectively. Moreover, a novel data augmentation method has been proposed to improve the generalization of the model, the model can be easily transferred to other cities.

METHOD

Clustering, as a fundamental unsupervised learning technique, facilitates the grouping of data points into clusters without the need for explicit labeling. Data points within the same cluster exhibit comparable characteristics, which can lead to new insights and discoveries, thereby enabling downstream tasks to leverage this inherent structure for enhancing the performance of classification or regression tasks. In the context, we employed contrastive clustering [3]. Pair construction backbone (PCB): The PCB constructs data pairs through data augmentations and extracts features from augmented samples. This helps to enrich the training data and improve the robustness of the model:

- **Instance-level contrastive head (ICH):** The ICH applies contrastive learning in the row space of the feature matrix. This helps to learn representations that are discriminative between different instances.
- **Cluster-level contrastive head (CCH):** The CCH applies contrastive learning in the column space of the feature matrix. This helps to learn representations that are discriminative between different clusters.



AUTHORS

Zhaobo Hu

zhaobo.hu@telecom-sudparis.eu

Chuan Li

chuanlifelix@gmail.com

Vincent Gauthier

vincent.gauthier@telecom-sudparis.eu

Hassine Mougla

hassine.mougla@parisdescartes.fr

Miguel Nunez del Prado Cortez

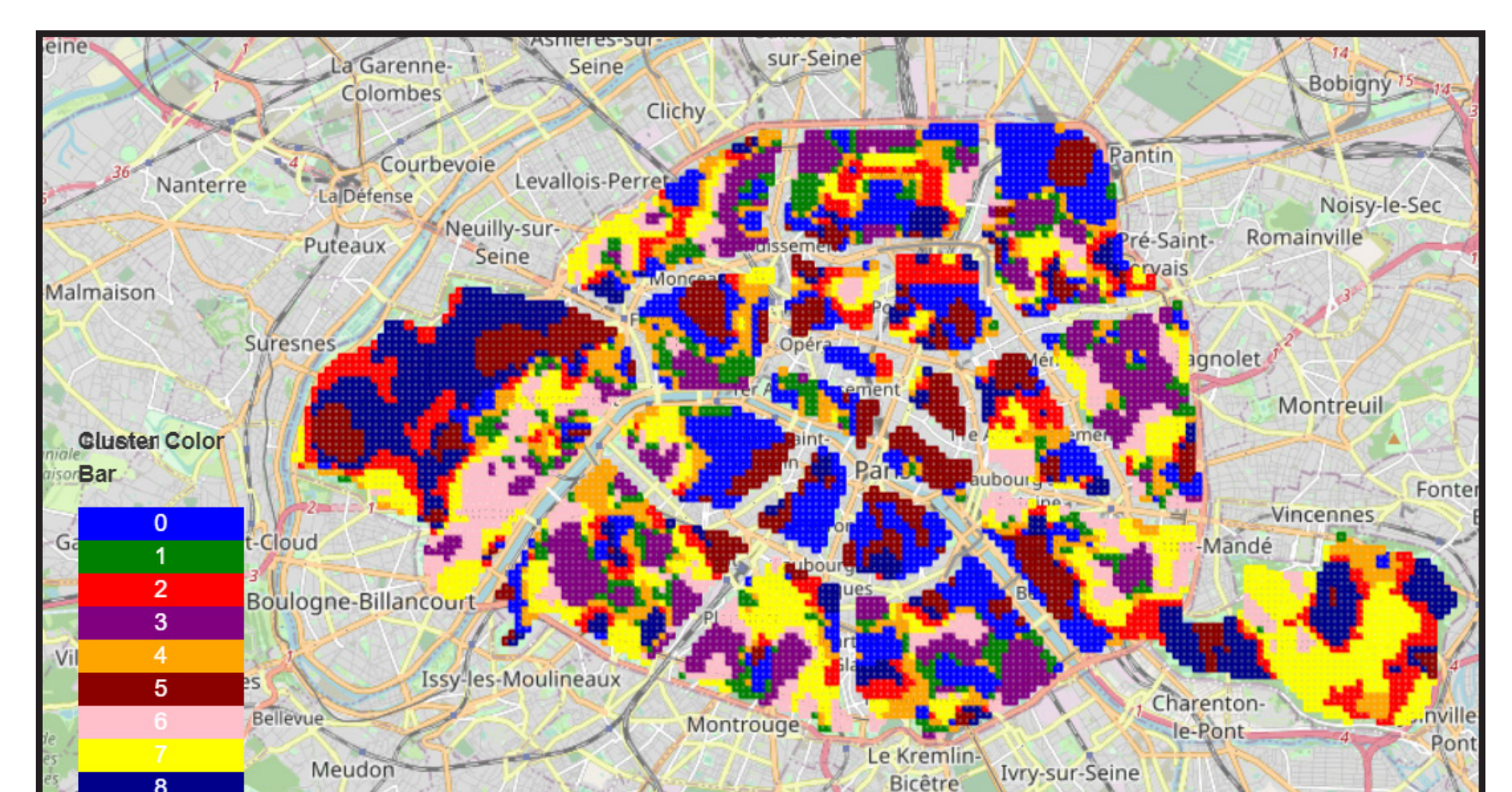
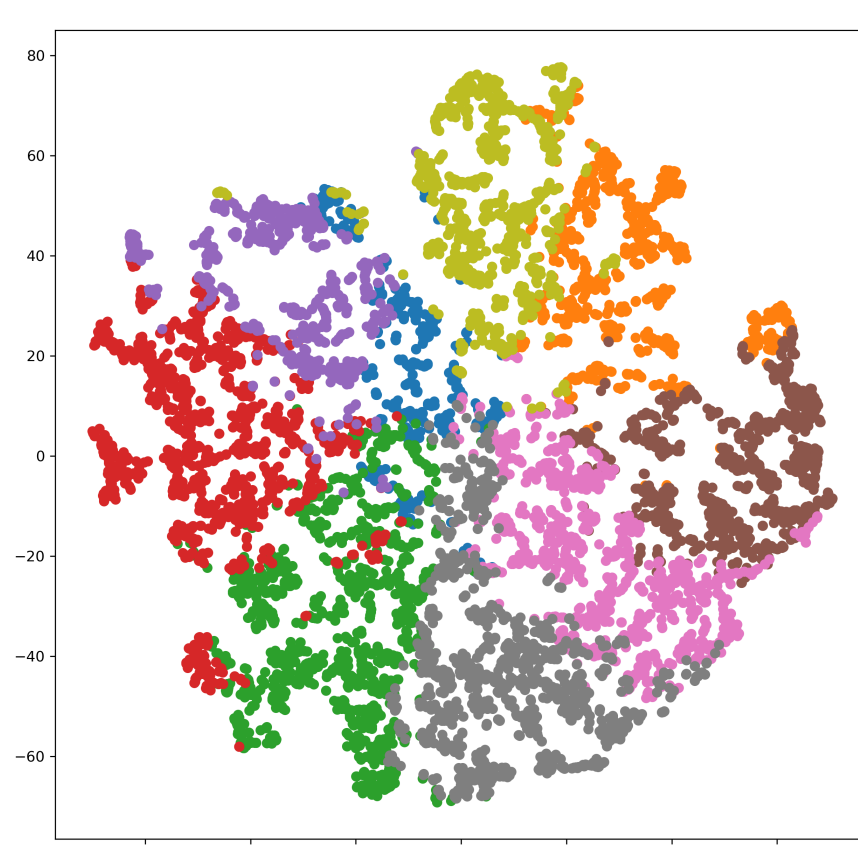
mnunezdelpradoco@worldbank.org

REFERENCES

- [1] G KHODABANDELOU, V GAUTHIER, M FIORE, AND M. EL-YACOUBI. "ESTIMATION OF STATIC AND DYNAMIC URBAN POPULATIONS WITH MOBILE NETWORK METADATA". IN: IEEE TRANSACTIONS ON MOBILE COMPUTING 2019.
- [2] A Furno, M Fiore, R Stanica, et al. "A Tale of Ten Cities: Characterizing Signatures of Mobile Traffic in Urban Areas". In: IEEE Transactions on Mobile Computing 2016.
- [3] Y Li, P Hu, Z Liu, ET AL. "CONTRASTIVE CLUSTERING". IN: IN PROCEEDINGS OF THE AAAI CONFERENCE ON ARTIFICIAL INTELLIGENCE. 2020.

PRELIMINARY RESULTS

In TSNE plot we can observe the evolution of the latent space with 9 clusters through different training steps. Based on the T-SNE plot, initially the latent data points are clustered into 2 clusters, at the end most of the data points are well separated, providing meaningful information for future analysis. We project the different clusters on Paris maps, we can observe that cluster we found are spatially coherent. This is important since our model didn't take into account explicitly the spatial dependency between map titles. Moreover, some specific area of the Paris map such as parc «parc de Boulogne» and «parc de Vincennes» and some well-known residential area in the outskirts of Paris city (in the 14th, 15th districts for instance) are clustered together. However, more inquiries are needed to provide a deeper analysis of the found cluster.



LIMITATION AND FUTURE WORK

Although we achieve good clustering on the hidden space, the demarcation between clusters is not obvious, which we attribute to the fact that the transition between certain map regions is not obvious. Meanwhile, the clustering numbers are arbitrarily defined. Future work will focus on improving the interpretability of the model, e.g. by analyzing the properties of the data points within the clusters. Experiments will also be conducted to define appropriate clustering numbers. Regarding the transferability of the model from one city to others, data augmentation of contrastive learning would of course be feasible, but needs experiments to prove it, since city-wide properties should be shared, and data augmentation provides generalization.