



HAL
open science

Human Presence Hotspot Extraction: A Mobile Network Data-Oriented Analysis

Gwladys Ornella Djuikom Foka, Solohaja Rabenjamina, Razvan Stanica,
Diala Naboulsi

► **To cite this version:**

Gwladys Ornella Djuikom Foka, Solohaja Rabenjamina, Razvan Stanica, Diala Naboulsi. Human Presence Hotspot Extraction: A Mobile Network Data-Oriented Analysis. NetMob 2024 - International Conference on the Analysis of Mobile Phone Datasets, Oct 2024, Washington DC, United States. hal-04741769

HAL Id: hal-04741769

<https://inria.hal.science/hal-04741769v1>

Submitted on 17 Oct 2024

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

Human Presence Hotspot Extraction: A Mobile Network Data-Oriented Analysis

Gwladys Ornella Djuikom Foka*, Solohaja Rabenjamina†, Razvan Stanica†, Diala Naboulsi*

*Software and IT Engineering Department, École de Technologie Supérieure (ÉTS), Montreal, Canada

{gwladys-ornella.djuikom-foka.1, diala.naboulsi}@ens.etsmtl.ca

†INSA Lyon, Inria, CITI, Villeurbanne, France

{solohaja.rabenjamina, razvan.stanica}@insa-lyon.fr

Index Terms—Hotspot, Data analysis, Cellular Network Management

I. INTRODUCTION

Mobile network traffic is not distributed evenly in time and space: there are moments and places when the network needs to accommodate more users and an increased demand. This is a consequence of two main factors. The first is that users move in the network, according to their daily schedules (e.g., work areas are obviously more popular during the day, while residential areas are more visited during evenings and at night [1]). The second reason is that user demand is not uniform in time: less network activity happens at night and during weekends. The result is the existence of *hotspots*, i.e. areas presenting a high network activity at certain times.

Hotspots have been defined from a human mobility point of view [2] as manifestations of concentrated human activities in an environment. Thus, they were also characterized from a network traffic standpoint [3]–[5], considering the traffic or the occupancy of the network itself. Analyzing and understanding hotspots behavior in detail is important for network optimization purposes, as well as for guaranteeing seamless connectivity and enhancing user experience. However, we are confronted with an unexpected problem: while the term *hotspot* is a common one and easily understandable, widely accepted formal definitions are lacking in the community.

II. HOTSPOT DEFINITIONS

In this section, we present a non-exhaustive overview of the definitions proposed in the literature while proposing a series of original definitions for the hotspot concept.

Definition II.1. A hotspot is a base station i which fulfills the following inequality [6]:

$$I_i \geq \frac{1}{N} \cdot \sum_{j=1}^N I_j + \Delta,$$

This work was supported by the National Natural Sciences and Engineering Research Council of Canada (NSERC) through research grant RGPIN-2020-06050 and by the Fonds de Recherche du Québec – Nature et Technologies (FRQNT) through research grant 2022-NC-297403. This work was partially funded by the ANR Mob Sci-Dat Factory (ANR-23-PEMO-0004) project.

where I_i is the user density of base station i , N is the number of base stations, and the value of Δ is calculated as follows :

$$\Delta = \left(MaxDensity - \frac{1}{N} \cdot \sum_{j=1}^N I_j \right) \cdot p,$$

where $MaxDensity$ is the maximum density in all the regions (base stations) and p is a parameter to determine the cutoff threshold.

Definition II.2. A hotspot is constituted of a set of cell sites and time periods where the aggregate user density encompasses over 50% of the total user density in the dataset [3].

From our perspective, the definitions provided above lack an acknowledgment of the relative strength of a potential hotspot compared to its closely surrounding areas. Additionally, they do not consider any historical element in the definition of the hotspot. Therefore, in the following, we propose a number of other possible definitions:

Definition II.3. A base station is considered to be a hotspot at time t if it presents a user density (number of users/ km^2) at least n times larger than the mean of the BS densities during the chosen time slot (date and time).

Definition II.4. A hotspot is a BS that satisfies the following properties during a given time slot:

- it has a higher user density than its k closest neighbours (geographically) during the time slot;
- it has a higher user density than its own mean historic density at that specific time of the day.

Definition II.5. A BS is defined as a hotspot during a time slot if it meets the two following conditions:

- it has a user density n times higher than the average density over the studied area during the specific time slot.
- it has a higher user density than its own mean historic density at that specific time of the day.

For the sake of simplicity, we refer to the two definitions from the state of the art as **Baseline 1** and **Baseline 2**, respectively, whereas our proposed definitions are referred to as **Definition 1**, **Definition 2** and **Definition 3**, respectively.

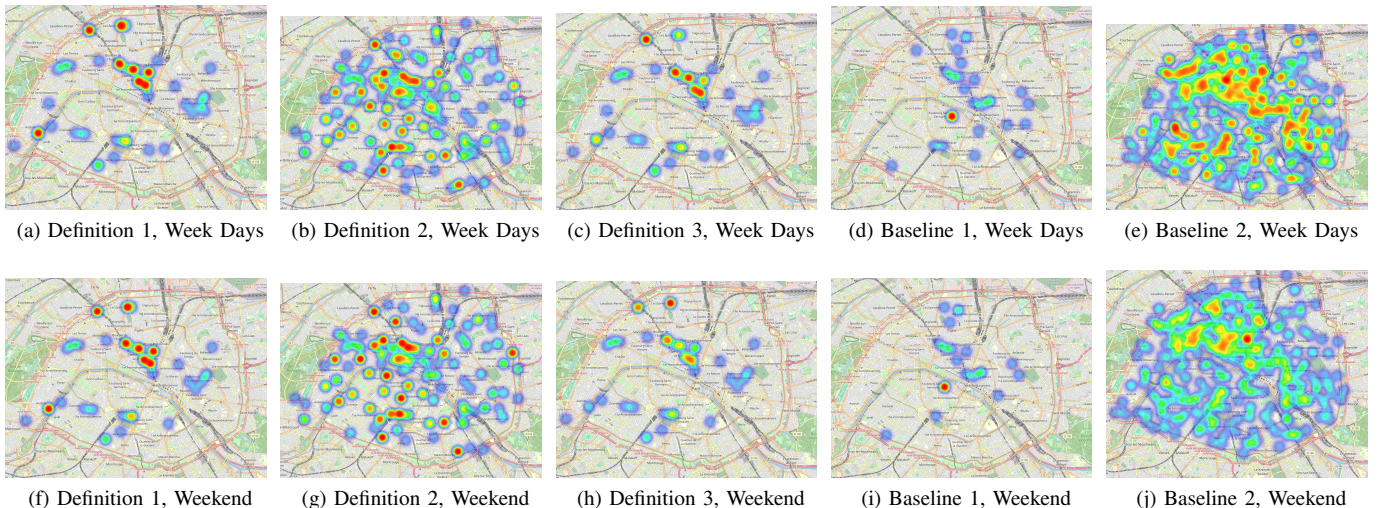


Fig. 1: Spatial distribution of hotspots occurrences for weekdays and weekends.

III. RESULTS AND DISCUSSION

In this section, we conduct an in-depth spatio-temporal analysis of the aforementioned hotspot definitions. Our study is based on a dataset collected by Orange, the largest French mobile operator, on the 4G network in the city of Paris. We consider 8 days of data, from 9th April 2019 to 16th April 2019. We are focused on the time series of the number of users with an active radio connection to the cellular network during one-hour intervals. This metric is representative of human presence in a given area, and it also gives a direct view of the utilization of the mobile network.

A. Hotspot Spatial Distribution

For each base station, we calculate the percentage of time slots for which the BS is detected as a hotspot, according to the five definitions. We plot these results as heat maps in Figure 1, distinguishing between weekdays (Figure 1a-e) and weekends (Figure 1f-j). In these figures, warmer colors indicate areas that are more frequently detected as hotspots.

There are two important observations we make based on Figure 1. First, at a close inspection, we notice some differences between weekends and weekdays. This is probably more visible in the case of Baseline 2, where the weekdays heatmap is more illuminated. However, differences are present for the other definitions as well. For example, for Definition 3, locations such as Gare de l’Est and Gare du Nord, or the Sorbonne campus, are frequently detected as hotspots during weekdays, while they never appear as hotspots during weekends.

B. Landmarks Analysis

To further analyze the spatial distribution of hotspots, we select several locations generally considered as important Paris landmarks, and compute the percentage of time when they are detected as hotspots during the 8 days in our dataset. The results are shown in Table I, demonstrating an important

Places	Definition 1	Definition 2	Definition 3	Baseline 1	Baseline 2
Chatelet les Halles	0.52	20.42	0.52	0.52	15.18
Louvre Museum	–	–	–	–	31.41
Gare de Lyon	–	2.62	–	–	3.66
Gare du Nord	–	–	–	–	95.29
Opera Garnier	100	67.54	67.54	34.55	57.07
Gare Saint-Lazare	–	–	–	–	45.55

TABLE I: Hotspots occurrences per definition (%)

heterogeneity among definitions. As expected, since Baseline 2 and Definition 2 detect many more hotspots than the other definitions, they are also much more present in this table.

IV. CONCLUSION

Hotspots are a common phenomenon in mobile networks, a result of non-uniform user presence and user demand, and network operators need to carefully manage them. Using a real 4G network dataset from Paris, France, we show significant discrepancies in the hotspots detected using different definitions, underlying the importance of having accurate and agreed-upon terminology.

REFERENCES

- [1] A. Furno, M. Fiore, R. Stanica, C. Ziemlicki, and Z. Smoreda, “A tale of ten cities: Characterizing signatures of mobile traffic in urban areas,” *IEEE Transactions on Mobile Computing*, vol. 16, no. 10, pp. 2682–2696, 2017.
- [2] F. Zhai and B. Li, “Point of interest recommendation based on graph convolutional neural network,” in *Journal of Physics: Conference Series*, vol. 1883, no. 1. IOP Publishing, 2021, p. 012132.
- [3] A. Nika, A. Ismail, B. Y. Zhao, S. Gaito, G. P. Rossi, and H. Zheng, “Understanding and predicting data hotspots in cellular networks,” *Mobile Networks and Applications*, vol. 21, pp. 402–413, 2016.
- [4] H. Klessig, V. Suryaprakash, O. Blume, A. Fehske, and G. Fettweis, “A framework enabling spatial analysis of mobile traffic hot spots,” *IEEE wireless communications letters*, vol. 3, no. 5, pp. 537–540, 2014.
- [5] A. Cao, Y. Qiao, K. Sun, H. Zhang, and J. Yang, “Network traffic analysis and prediction of hotspot in cellular network,” in *2018 International Conference on Network Infrastructure and Digital Content (IC-NIDC)*, 2018, pp. 452–456.
- [6] E. Mededovic, V. G. Douros, and P. Mähönen, “Node centrality metrics for hotspots analysis in telecom big data,” in *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2019, pp. 417–422.