



**HAL**  
open science

## Mobility-aware estimation of content consumption hotspots for urban cellular networks

Sahar Hoteit, Stefano Secci, Guy Pujolle, Vinh Hoa La, Cezary Ziemlicki, Zbigniew Smoreda

► **To cite this version:**

Sahar Hoteit, Stefano Secci, Guy Pujolle, Vinh Hoa La, Cezary Ziemlicki, et al.. Mobility-aware estimation of content consumption hotspots for urban cellular networks. IEEE/IFIP Network Operations and Management Symposium (NOMS 2014), May 2014, Krakow, Poland. pp.1-7, 10.1109/NOMS.2014.6838312 . hal-01131513

**HAL Id: hal-01131513**

**<https://hal.science/hal-01131513v1>**

Submitted on 16 Mar 2015

**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.

# Mobility-Aware Estimation of Content Consumption Hotspots for Urban Cellular Networks

Sahar Hoteit, Stefano Secci, Guy Pujolle, Vinh Hoa La<sup>†</sup>, Cezary Ziemlicki<sup>†</sup>, Zbigniew Smoreda<sup>†</sup>  
Sorbonne Universities, UPMC Univ. Paris 06, UMR 7606, LIP6, F-75005, Paris, France

Email: [firstname.lastname@upmc.fr](mailto:firstname.lastname@upmc.fr)

<sup>†</sup> Orange Labs, 38 rue du G. Leclerc, 92794 Issy, France

Email: [firstname.lastname@orange.com](mailto:firstname.lastname@orange.com)

**Abstract**—A present issue in the evolution of mobile cellular networks is determining whether, how and where to deploy adaptive content and cloud distribution solutions at the base station and backhauling network level. Intuitively, an adaptive placement of content and computing resources in the most crowded regions can grant important traffic offloading, improve network efficiency and user quality of experience. In this paper we document the content consumption in the Orange cellular network for the Paris metropolitan area, from spatial and application-level extensive analysis of real data from a few million users, reporting the experimental distributions. In this scope, we propose a hotspot cell estimator computed over user’s mobility metrics and based on linear regression. Evaluating our estimator on real data, it appears as an excellent hotspot detection solution of cellular and backhauling network management. We show that its error strictly decreases with the cell load, and it is negligible for reasonable hotspot cell load upper thresholds. We also show that our hotspot estimator is quite scalable against mobility data volume and against time variations.

## I. INTRODUCTION

As a matter of fact, we are witnessing the overtaking of a second decennial milestone in the evolution of telecommunications since the commercial Internet in 1992. After data traffic overcoming voice traffic in the beginning of current century, mobile data traffic is expected to overcome wireline access traffic. As of Institute Mediametrie [1], more than 31% of French mobile users use a smartphone, and about 71% of them daily connect to the Internet. This recent trend seems to accelerate; e.g., the number of mobile Internet users increased by 10% in just one quarter last year.

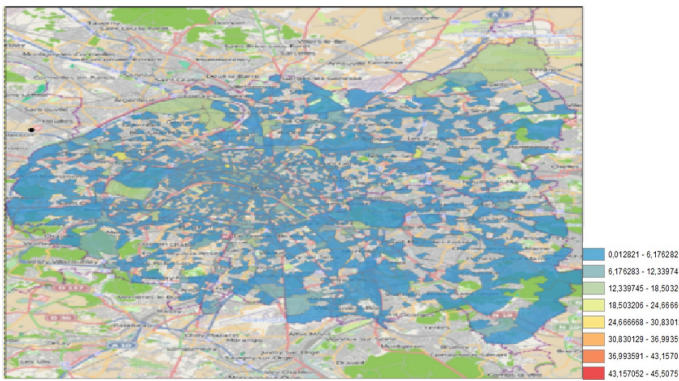
The advent of smartphones, tablets and Internet keys, with generous processing and memory capacity, is shifting the network engineering attention from core networks and wireline local loops to the mobile cellular network. At present, techniques typically implemented in wireline network engineering are being considered for cellular access and backhauling networks. On one hand, content delivery networks and caching protocols could go down close to backhauling gateways and base stations. On the other hand, Cloud servers could approach the user at mini-datacenter connected to backhauling elements. The aim is to offload the transport network when it is congested and to improve user’s quality of experience reducing the access latency.

In order to better characterize these phenomena, in this paper we analyze user content consumption habits from the Orange cellular network in the Paris metropolitan area. We provide experimental distributions of key aspects to allow dataset reconstruction. Moreover, we propose a methodology to estimate hotspot cells based on compact user mobility metrics. In particular, our estimator is based on linear regression of parameters computed over users’ centroids and radius of gyration. We detail the performance of our estimator as a function of cell load and different regressions. We show that as a cell load estimator it has an error on median inferior to 10%, that the error is decreasing with the cell load, and that for hotspot cell identification the load estimation and identification errors are negligible. We also show that the estimator is robust and gives excellent results for different time windows and as a function of the volume of data and the sample regression window.

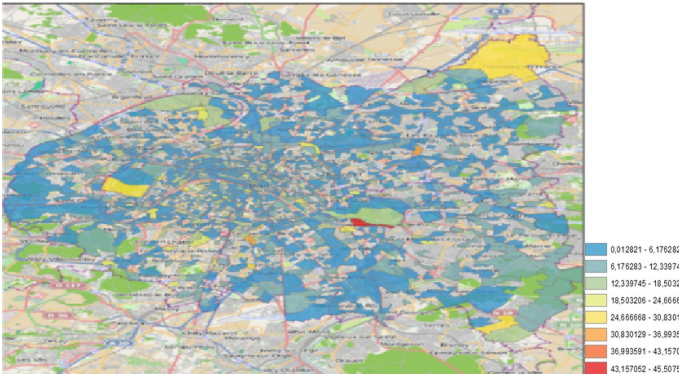
The paper is organized as follows. Section II synthetically presents the dataset. Section III presents content consumption maps and analyzes them from a spatial distribution and application perspective. Section IV presents the hotspot estimation issue and the estimator we propose. Section V presents estimation results for the Orange’s Paris metropolitan area network. Section VI summarizes our work.

## II. DATASET

The dataset we consider in this work comes from network management tickets, generated each time a mobile device uses wireless mobile network for Internet data exchange (i.e., what is commonly referred to as “mobile Internet” service). The network probe data exploited in the study provide information on the protocol used for the communication, so it is possible to categorize the traffic by application (Web, VoIP, P2P, streaming etc.). Data are individual, so all user identifiers were irreversibly anonymized before the analysis to protect user’s privacy. The probe collects 6-minute spaced data session information, assigning the session to the cell identifier of the last used antenna. Data are recorded on a per-user basis and cover about 1,5 million mobile phone users from the “Ile-de-France” French department, giving about 100 millions of records per day. We limit our study to two single days and to the Inner Ring “La Petite Couronne” of the Paris metropolitan area,



(a) June 25, 2012 (whole day)



(b) July 1, 2012 - during the football game

Fig. 1. User density (nb. of users per minute)

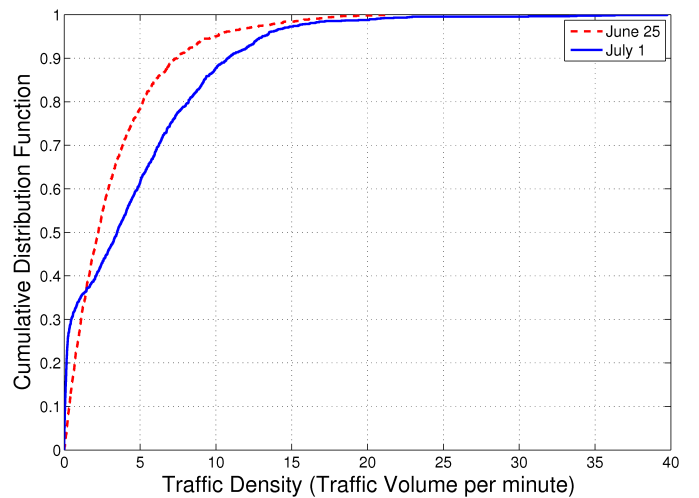
formed by three Ile-de-France peri-urban departments (92, 93, 94) bordering with the Paris urban department (75) and forming a geographical crown around it. We decompose the region according to the IRIS2000 system, developed by INSEE (National Institute of Statistics and Economics Studies), which divides a territory into units of equal size; a French acronym, ‘IRIS’ stands for “aggregated units for statistical information”, and “2000” refers to the target size of 2000 residents per IRIS unit (using IRIS-localized data, an IRIS zone may not contain base stations). For hotspot cell estimations, we further divide an IRIS into cells, each cell being composed of a few antennas able to host up to roughly 1000 users.

### III. CONTENT CONSUMPTION DISTRIBUTIONS

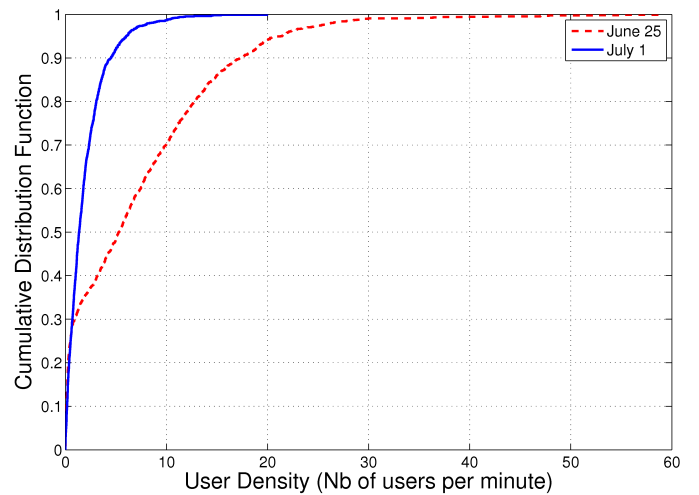
In this section, we synthetically characterize content consumptions, as a function of time, place and applications as done in [2]. We consider datasets of two days, one normal day with standard content consumption activity (June 25<sup>th</sup>, 2012), and one special day where a particular content consumption is expected (July 1<sup>st</sup>, 2012), the day of the Italy-Spain Eurocup football final; for this special day, only the 2h period during the match is taken into account.

#### A. Content Consumption Spatial Distributions

Looking for differences between urban and peri-urban region is a first naturally arising research question. In the



(a) Traffic Density CDF



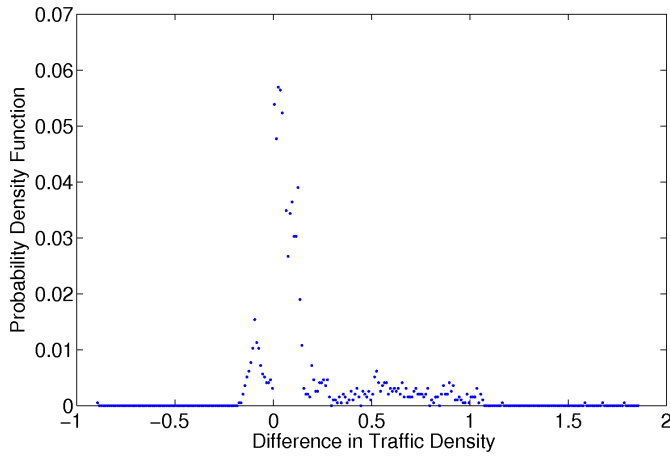
(b) User Density CDF

Fig. 2. Traffic and User Densities - CDF

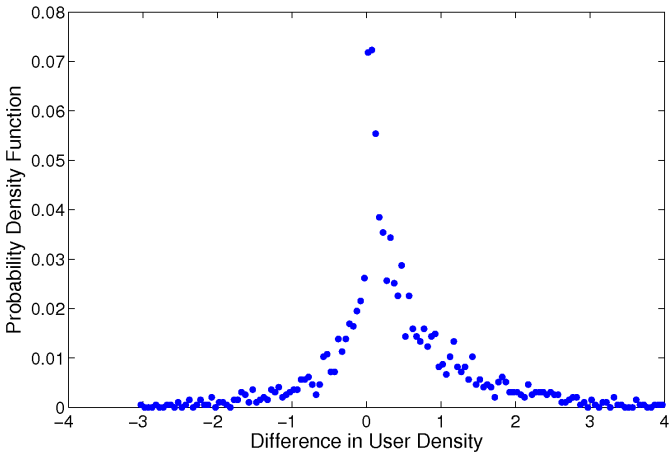
following, we define the user density as the total number of users in a given IRIS, to the time of observation in the two periods, and the traffic density as the exchanged byte volume by IRIS’s base stations to the time of observation.

Fig. 1 represents the user density over seven-quantile intervals. We can remark that a few IRISs are clearly more overloaded than others, presenting a large number of users and a large traffic volume; the reason is that they cover clearly identifiable content consumption hotspots such as the two Paris international airport terminals (in the north-east and in the center-south) with travelers actively surfing using their mobile phone while waiting, and the La Defense business center. During the football game, hotspots appear in some IRIS inside the capital as well as outside it. These crowded places are very likely some public spaces where people use to gather together.

In order to further explore this latter aspect and the density distributions, Fig. 2(a) and Fig. 2(b) plot the experimental cumulative distribution functions (CDFs) of traffic and user



(a) PDF of the differential traffic density (normal day - special day)



(b) PDF of the differential user density (normal day - special day)

Fig. 3. PDF of the Differential Traffic and User Densities

density in the two considered days. We can easily notice that:

- In terms of traffic density, the game period shows a higher ratio of low density IRISs (e.g., 30% with less than density = 5) than the normal day, and a double ratio of high density IRISs (e.g., with a density > 10 are 10% during the game, and 5% in the normal day);
- In terms of user density, during the normal day dense zones show a density from 5 to 15 times higher than during the game (e.g., the zones with a density > 10 are 30% in the normal day, and only 2% during the game).

These two aspects clearly show important differences. Users are much less concentrated during the special event, which appears as counterintuitive at a first view, but could be explained with the presence of very-high-density residential zones in many peri-urban departments (e.g., Montrouge, Malakoff are among the highest density communities in Europe): during the event, people gathering elsewhere might decrease the ratio of very-high density residential IRISs. Moreover, during the event opposite behaviors are exalted, i.e., users usually not surfing much, surf even less, and those usually surfing more, surf even more. This latter aspect seems being confirmed by

the differential density distribution in Fig. 3(a) and Fig. 3(b). We can observe that traffic density difference distribution roughly follow a normal distribution, and that user density difference distribution is dominated by positive differences (higher densities in normal days).

There are therefore important differences in the geographical localization of content consumption hotspots. As evidenced, their location can change at different periods of the week and of the day. Moreover, a large majority of the traffic volume is related to bulk transfer and web-related services whose content could be either partially cached or whose Cloud server could be located close to the hotspots to well manage the cloud and the content offloading. Users could indeed be better served if delivery facility could be located closely to the hotspots. The backhauling network performance could also benefit from traffic offloading and traffic engineering techniques aware of user mobility and content consumption point deflections (e.g., adjusting link auto-bandwidth, path reservations and packet tunneling).

#### IV. ESTIMATION OF HOTSPOT CELLS

The type of dataset we use is easily available in most of mobile cellular network carriers. These data are derived from network management tickets and generated in real-time as the user consumes cellular network services, it appears technically feasible to implement a hotspot estimation technique indicating where the traffic load is going to increase closely.

##### A. Related Work

A limited amount of work exists in the literature on the estimation of hotspots and rendez-vous points in wireless access networks. E.g., in [3] vehicular data is exploited to determine accident-risk points. Authors in [4] propose a framework that discovers regions of different functions in a city using both human mobility among regions and points of interests located in a region. Many other works, such as [5], [6], [7], and [8], while assuming the availability of mobility information, focus in user-profile aware QoS provisioning, load balancing and network signaling improvement techniques.

Traffic load forecasting has also been investigated from an analytical and mathematical modeling perspective. For example, authors in [9] show how under certain conditions periodic sinusoidal functions can be used as cellular traffic profile. Unfortunately the simplicity and the too theoretical properties of these approaches fail from precisely matching with the actual real traffic load, which is a strict requirement of our investigation. In this direction, relevant works targeting mobility pattern detection from real cellular network data have been studied, e.g., [10] [11] [12]. In particular, authors in [12] categorize users with respect to their radius of gyration, allowing to define a circle, centered at user's centroid, around which a user is expected to geographically move.

##### B. Mobility-Aware Hotspot Estimator

Following [12], given a user  $u \in U$  who has been located at  $n^u(t)$  locations until time  $t$ , its centroid  $\bar{r}_{cm}^u(t)$  can be

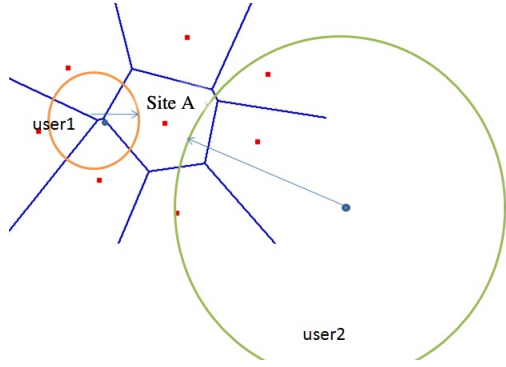


Fig. 4. Illustration of two user's radius of gyration around their centroids.

computed as the location pointed by the sum of all vectors related to all previous sampled positions  $\vec{r}_i^u(t)$ , i.e.:

$$\vec{r}_{cm}^u(t) = \frac{1}{n^u(t)} \sum_{i=1}^{n^u(t)} \vec{r}_i^u(t) \quad (1)$$

The radius of gyration can then be computed for each user as:

$$r_g^u(t) = \sqrt{\frac{1}{n^u(t)} \sum_{i=1}^{n^u(t)} |\vec{r}_i^u(t) - \vec{r}_{cm}^u(t)|^2} \quad (2)$$

Aiming at defining an adequate and implementable hotspot cell estimator for cellular network management, we adopt as estimation parameters the centroid and the radius of gyration to account for user's spatial coverage. The 'territory' of a user can therefore be defined as the area covered by a circle centered at its centroid with a radius equal to the radius of gyration. Our proposed estimator takes into account the non negligible intersections of different users' territories to estimate the cellular user density. For example, in Fig. 4, based on the centroid and radius of gyration of the two users, it is reasonable to account for the possibility that the two users pass by site A. The efficiency of the estimator should be evaluated toward its capability of estimating hotspot cells rather than lightly loaded cells.

Let  $m[r_g^u(t), \vec{r}_{cm}^u, c, u, t]$  be a spatial mapping counter equal to 1 if the circle of user  $u$ , with radius  $r_g^u(t)$  around the centroid  $\vec{r}_{cm}^u$  at time  $t$ , covers at least 10% of cell  $c$ , 0 otherwise (i.e., the 10% is actually 10% of the area of each cell presented in Fig. 1); other thresholds than 10% could certainly be considered, depending also on the way the environment is architected, its aim being to avoid the small overlaps since not all the cells covered by one user's radius will reasonably be visited.

Hence such a spatial mapping counter is the core metric of our estimator. However, alone it is not precise enough. Simply counting the number of intersections significantly covering a given cell would certainly lead to an over counting that needs to be appropriately scaled. A simple scaling could be, e.g., to divide it by the average number of users per cell during the past measurements, yet this does not prevent from high deviations.

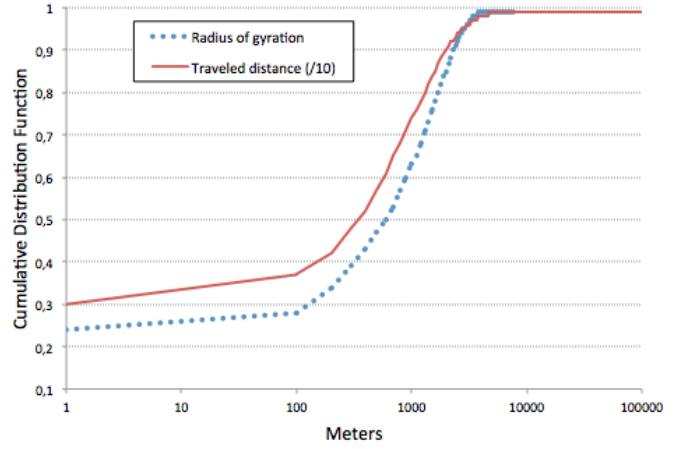


Fig. 5. Radius of gyration and traveled distance CDFs

We propose to scale it by the scale factor that would generate null estimation error in an arbitrary instant  $t - T$  in the recent past, with a regression time  $T$  adequately set.

Then, as an estimator of the number of users per cell  $c$ , we propose:

$$\hat{e}(c, t) = \frac{\sum_{u \in U} m[r_g^u(t), \vec{r}_{cm}^u, c, u, t]}{\sum_{u \in U} m[r_g^u(t), \vec{r}_{cm}^u, c, u, t - T]} * n(c, t - T) \quad (3)$$

It uses therefore a linear regression over past measurements to adequately weight the spatial mapping counter. Intuitively, the smaller  $T$  is, the closer  $\hat{e}(c, t)$  is to the unknown  $n(c, t)$ , i.e., the more accurate the estimation is. However,  $T$  should be big enough to allow network management system to retrieve the real number of users in cell  $c$  at time  $t - T$ , i.e.,  $n(c, t - T)$ . Depending on network management tickets, session duration, network latency and network size, this parameter  $T$  could range from a few dozens of minutes to a few hours.

## V. ESTIMATOR EVALUATION

We evaluate our estimator using an available network management dataset of the Orange France cellular network, in an arbitrary working day, computing centroid and radius of gyration of all users passing by the considered region, discarding those sampled positions falling outside the Paris metropolitan area network.

In order to qualify the dataset, Fig. 5 reports the CDFs of the users' radius of gyration and of the overall traveled distance, and Fig. 6 shows the CDF of the real number of users per cell at a given instant (the end of the sampled day). We can appreciate that the radius of gyration adequately represents the user mobility habits as it increases in parallel with the traveled distance. Moreover, 40% of the cells have more than 100 users, and 10% of the cells have more than 400 users. It is worth noting that the distance between two neighboring cells is typical of a few hundreds of meters.

### A. Hotspot cell estimation accuracy

We intend to estimate hotspot cells positions using our proposed estimation technique (3). In this evaluation, to in-

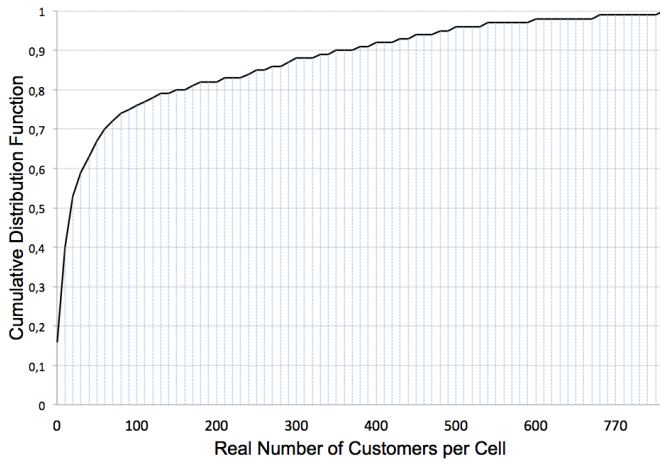


Fig. 6. CDF of the real number of customers per cell

crease the estimation pertinence, we further segment each IRIS into smaller cells, each cell grouping few co-located antennas. Moreover, we set a regression time  $T$  to 1 hour, which is a quite pessimistic value (in practice, in carrier grade networks, it could even be set to a few minutes, hence allowing a higher accuracy). We consider the user position samples from the whole day (more precisely, the sample temporal window is  $24 - (t - T) = 23$  hours) in the computation of users' centroids and radius of gyration.

To qualify the accuracy of the proposed hotspot detection method, we analyze the estimation error (i.e., the difference between the real and the estimated number of users, to the real number of users) for all the cells in the Paris metropolitan area network.

Fig. 7 depicts the CDF of the absolute value of the estimation error defined previously. Fig. 8 is a log-log scatter representation of the estimation error as a function of the real number of users per cell. We can observe that the estimation error decreases more than linearly as the cell load increases. In particular, the CDF shows that including all cells, the estimation error is always lower than 50%, and the median error is lower than 10%. Including only cells with a load higher than the load of 60% of the cells, i.e., 50 users, the error is inferior to 5%. With stricter definitions of the load threshold beyond which a cell is considered as a hotspot, 100 (70% percentile) and 200 (80% percentile) users, the error becomes negligible.

To further check the accuracy of the proposed method, we evaluate type I and II errors in hotspot identifications, i.e., false positive (a cell is identified as a hotspot when it is not) and false negative (a cell is not identified as a hotspot when it is a hotspot) errors, respectively. We choose five different thresholds beyond which we can say that a given cell is a hotspot one: at the 50<sup>th</sup> percentile ( $\approx 35$  users), the 60<sup>th</sup> percentile ( $\approx 50$  users), the 70<sup>th</sup> percentile ( $\approx 100$  users), the 80<sup>th</sup> percentile ( $\approx 200$  users), the 90<sup>th</sup> percentile ( $\approx 400$  users). We plot for each threshold case the false positive and false negative ratios evaluated by directly counting

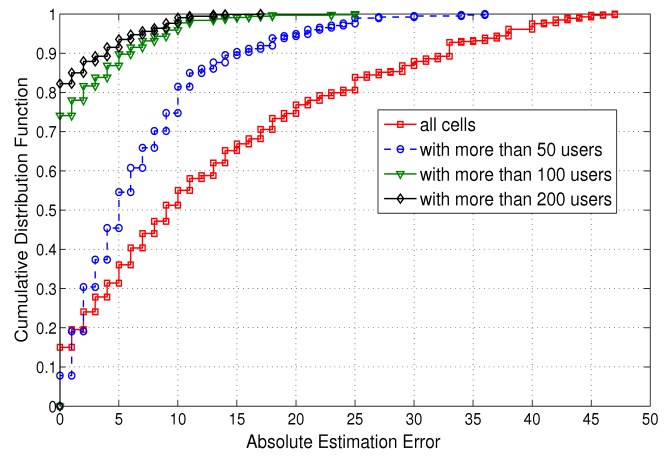


Fig. 7. Absolute Estimation Error Distribution

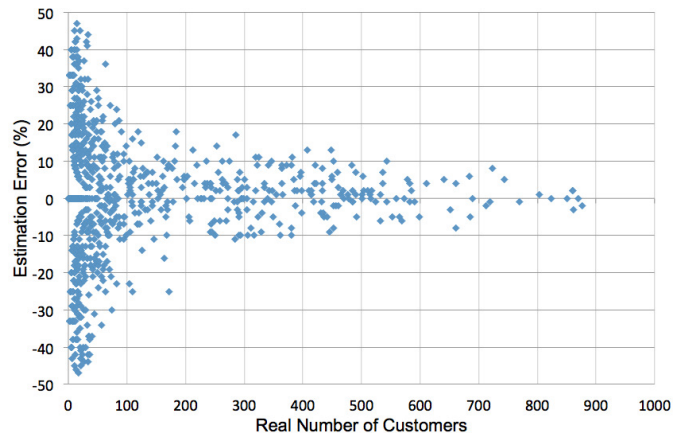


Fig. 8. Scatter of the estimation error as a function of real cell load

from the complete dataset against these obtained through our proposed method. Fig. 9 shows encouraging results with more and more negligible and decreasing erroneous identifications with stricter definitions of hotspots. Especially, for the 80<sup>th</sup> percentile case the error is inferior than 1%, and for the 90<sup>th</sup> percentile case, no false negative occurred and only 0.2% false positive.

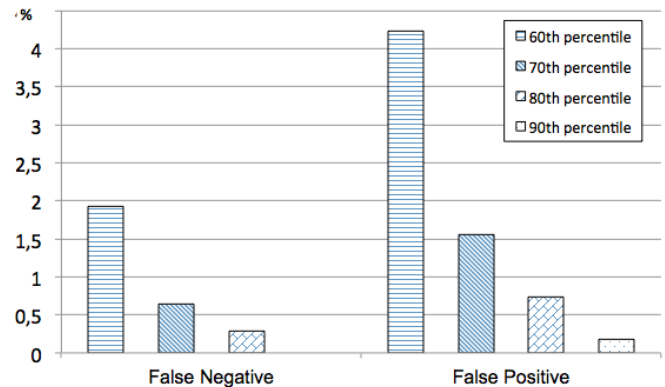


Fig. 9. Type I and Type II errors as a function of hotspot thresholds.

In our opinion, these are extremely encouraging results. Our estimator permits reaching an excellent accuracy. It can be used to feed a network management policy aiming at taking a decision when detecting a hotspot (e.g., link auto-bandwidth bounds setting in mobile backhauling, adaptive content-delivery-network redirections, adaptive virtual machine migrations, etc.). For operational consideration, we however need to qualify its scalability.

### B. Implementation complexity and scalability performance

Our estimator (3) is of a rather simple and intuitive nature. Certainly, a more complex estimator, e.g., weighting by a user-cell crossing probability computed using other mobility information, might be defined to reach lower errors and can be left for future work, yet the estimator computation complexity has to be as low as possible, and the utility of having a more precise estimator is a matter of discussion since the accuracy of (3) is already extremely high.

With our estimator, for each new user position, (1) and (2) could be updated handling only a limited arbitrary number  $N$  of last positions, then (3) can be updated. Hence, all these operations have a  $O(N*|U_m(t)|)$  time complexity,  $U_m(t) \subset U$  being the subset of users moving and sampled in the sampling interval  $(t-T, t)$ . As already mentioned, the regression time  $T$  and the estimator's recomputation interval, should be chosen so that it is sufficiently higher than the (3) computation time.

However, from a more practical perspective, in very high mobility environments the above complexity could become quadratic, which may rise scalability concerns. Scalability concerns could also rise from the volume of the data to mine in order to extract estimator's metrics. The larger the sample temporal window is, the larger the data volume is and the higher the computation complexity is. Hence the last questions we want to answer are: is the hotspot detection accuracy decreasing with the sample temporal window? What happens at different times of the day?

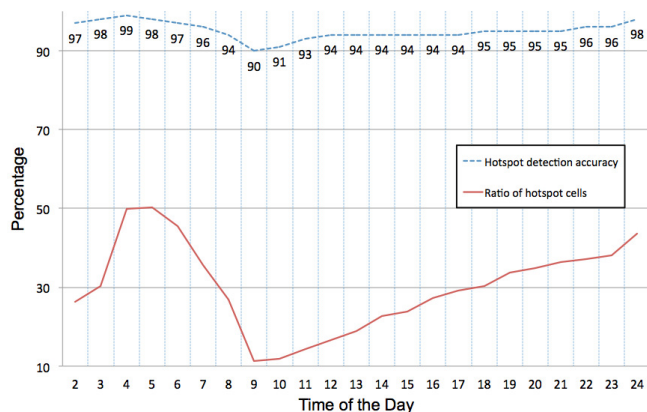


Fig. 10. Hotspot detection accuracy and ratio of hotspots as a function of time and sample regression window

In order to answer these two questions jointly, we test our estimator in the same day, with a sample temporal regression ranging from 1 to 23 hours (it is worth remembering that the

sample regression window of the previous simulations is 23 hours), and the sample time ranging from 1:00 in the morning to 23:00 in the evening.

Fig. 10 indicates our evaluation results in different instants of the day, with a step of 1 hour starting by 2:00. For example, when performing the hotspot estimation at 10:00, we exploit the data collected from 0:00 to 9:00 (as before,  $T$  is set to 1 hour). These results highlight two important aspects. Firstly, the accuracy is influenced by the time of the day, and hence the user propensity to move; it is extremely high when most users maintain fixed positions (i.e., around 4:00 am when people rather sleep than move around), and decreases when most users move out (i.e., around 9:00 am). The decrease is however relatively low (9% gap). Secondly, whenever users are on the move, they generate a higher number of sample positions. Hence we see that the more mobility data is taken into account, the more accurate estimations become (i.e., from 9:00 to 24:00, the accuracy is monotonously increasing). When instead most users are sedentary (i.e., from 4:00 to 9:00), even if the sample temporal window increases, the accuracy slightly decreases. An important aspect therefore is that the number of detected hotspots, also depicted in Fig. 10, increases when the overall user mobility decreases, and viceversa.

## VI. CONCLUSION

A present matter in mobile access network management is to find scalable and effective ways to offload the backhauling and core network as a function of user mobility and consumption patterns. The motivation is the arising weight mobile Internet traffic taked over legacy wireline access traffic in todays Internet Service Provider networks. Different offloading techniques are currently studied which could be classified as traffic offloading (e.g, Wifi offloading [14]), content offloading (e.g, ICN [13]) and Cloud (computation) offloading (e.g., CloudLet [15] [16]). In this scope, the contribution of this paper is twofold.

Firstly, working on real network probe data from the Orange cellular network of the Paris metropolitan area, we show how content consumption hotspots spatially move as a function of the occurrence of special events. We provide useful user, traffic density and traffic type experimental distributions to the research community.

Secondly, motivated by the experimental findings, we propose a hotspot cell estimaton technique based on two compact and easy-to-compute user mobility metrics, i.e. the user's centroid and radius of gyration, and relying on linear regression. The technique basically consists of measuring the cell intersections between users' geographical 'territories' defined by users' centroids and radius of gyration, weighting it by a regression factor while accounting for the gap between estimated and real numbers in the recent past. Results on real data show that the error of our estimator exponentially decreases with the cell load, and that the hotspot identification presents a quite small estimation error for the most loaded cells, becoming negligible ( $< 1\%$ ) when defining the hotspots as those 20% more loaded cells. We also stressed our estimator for more

challenging conditions with smaller regression window and at different time of the day. We show that the estimation accuracy remains very high, and only marginally decreases in high mobility temporal slots.

#### ACKNOWLEDGMENT

This work was partially supported by the ANR ABCD project (Grant No: ANR-13-INFR-005), and by the EU FP7 IRSES MobileCloud Project (Grant No. 612212).

#### REFERENCES

- [1] "L'audience de l'Internet mobile en France", Mediametrie, <http://www.mediametrie.fr/internet/communiqués/telecharger.php?f=e8c0653fea13f91bf3c48159f7c24f78>, 2011.
- [2] S. Hoteit, S. Secci, G. Pujolle, Z. He, C. Ziemlicki, Z. Smoreda and C. Ratti "Content Consumption Cartography of the Paris Urban Region using Cellular Probe Data", in *Proc. of ACM URBANE 2012, CoNext 2012 Workshop*, 2012, Nice, France.
- [3] TK. Anderson, "Kernel density estimation and K-means clustering to profile road accident hotspots", *Accident Analysis and Prevention*, Vol. 41, No. 3, 2009
- [4] J. Yuan, Y. Zheng, and X. Xie, "Discovering regions of different functions in a city using human mobility and POIs.", In *Proc. of the 18th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, New York 2012.
- [5] K. Seada, "Rendezvous regions: a scalable architecture for service location and data-centric storage in large-scale wireless networks", in *Proc. of 2004 Parallel and Distributed Processing Symposium*.
- [6] S.K. Das, S.K.S. Jayaram, "A novel load balancing scheme for the tele-traffic hot spot problem in cellular networks", *Wireless Networks*, Vol. 4, No. 4, 2004.
- [7] D. Ghosal, B. Mukherjee, "Exploiting user profiles to support differentiated services in next-generation wireless networks", *IEEE Networks*, Vol. 18, No. 5, 2004.
- [8] H. Zang, J. Bolot, "Mining call and mobility data to improve paging efficiency in cellular networks", in *Proc. of ACM MOBICOM 2007*.
- [9] E. Oh and B. Krishnamachari, "Energy Savings through Dynamic Base Station Switching in Cellular Wireless Access Networks", In *Proc. of IEEE Globecom 2010*.
- [10] F. Calabrese et al., "The Geography of Taste: Analyzing Cell-Phone Mobility and Social Events", In *Proc. of IEEE PerComp 2010*.
- [11] H. Hohwald, E. Frias-Martinez, and N. Oliver, "User modeling for telecommunication applications: Experiences and practical implications", in *Proc. UMAP*, pp. 327-338, 2010.
- [12] M. Gonzalez, CA . Hidalgo, Al. Barabasi, "Understanding individual human mobility patterns", *Nature* 458, pp. 238-238, 2008.
- [13] A. Detti, M. Pomposini, N. Blefari-Melazzi, S. Salsano and A. Bragagnini, "Offloading cellular networks with Information-Centric Networking: The case of video streaming", In *Proc. of the Int. Symp. on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, IEEE, San Francisco, CA, USA, 1-3, 2012.
- [14] K. Lee , J. Lee, Y. Yi, I. Rhee, S. Chong "Mobile Data Offloading: How Much Can Wi-Fi Deliver?", *IEEE/ACM Transactions on Networking*, Vol. 21, No. 2, April 2013.
- [15] M. Satyanarayanan, P. Bahl, R. Caceres, and N. Davies. "The Case for VM-based Cloudlets in Mobile Computing", *IEEE Pervasive Computing*, 8(4), 2009.
- [16] Lei Jiao, Roy Friedman, Xiaoming Fu, Stefano Secci, Zbigniew Smoreda, Hannes Tschofenig "Challenges and Opportunities for Cloud-based Computation Offloading for Mobile Devices", in *Proc. of Future Network and Mobile Summit 2013, 3-5 July, 2013, Lisbon, Portugal*.