Understanding Human Mobility Due to Large-Scale Events

Faber Henrique Z. Xavier
Pontifical Catholic University of Minas Gerais (PUC MINAS)
Brazil – 30.535-901
faber.xavier@sga.pucminas.br

Lucas M. Silveira
Pontifical Catholic University of Minas Gerais (PUC MINAS)
Brazil – 30.535-901
lmsilveira@sga.pucminas.br

Jussara M. Almeida
Universidade Federal de Minas Gerais (UFMG)
Brazil – 31.270-010
jussara@dcc.ufmg.br

Carlos Henrique S. Malab
Oi Telecom
Brazil – 20.230-070
malab@oi.net.br

Artur Ziviani
National Laboratory for Scientific Computing (LNCC/MCTI)
Brazil – 25.651-075
ziviani@lncc.br

Humberto T. Marques-Neto
Pontifical Catholic University of Minas Gerais (PUC MINAS)
Brazil – 30.535-901
humberto@pucminas.br

I. Introduction

Analyzing the mobility patterns of cellphone users is a challenging task, but also a great opportunity to better understand the human dynamics in a covered area. Although recent studies show that human mobility in urban areas can be predictable considering daily routines [1], cellphone carriers still have difficulties for planning the necessary communication infrastructure to support the unusual workload that arises during large-scale events [2]. Such events typically involve a large number of people within an urban area, such as the final match of a soccer championship, a major rock concert (e.g., Rock in Rio), New Year's Eve celebrations, a religious pilgrimage, political manifestations, or the Olympics.

We here consider a large-scale event to be characterized by a huge number of people with similar interests directly related to the event's main subject who move towards/from a specific place in order to participate on a set of collective activities during a period of time. Even though many of these large-scale events are scheduled and planned in advance, and are expected to cause collective changes in the mobile phone workload [3], it remains common to notice the congestion of the carrier's resources during them.

In this paper, we present our on-going work towards understanding the human mobility and the workload dynamics of mobile phone networks due to large-scale events. To that end, we analyze the impact of some types of large-scale events on the workload of a mobile phone network. We use our recently proposed methodology [4], applying it to real anonymized mobile phone datasets provided by a major mobile operator in Brazil. Whereas in [4] the methodology was applied to datasets collected during major Brazilian soccer matches, we here apply it to a different type of large-scale event: New Year's Eve celebrations in three large Brazilian cities. Our results could be used to improve the understanding of human mobility in urban areas due to large-scale events, thus contributing to the network management of the mobile phone operators and also to the development of new applications and devices.

II. Related Work

Better understanding the dynamics of the workload imposed on mobile phone networks is increasingly gaining attention from different research efforts [5]. The prediction of user mobility patterns can help, for instance, detecting routine patterns [6], urban planning efforts (e.g., urban traffic planning [7]), and public health management (e.g., disease spread control [8]). Other authors argue that characteristics of the network workload can be exploited to identify the kind of events users are experiencing (e.g., an emergency or a concert) [3].

Two previous studies are particularly related to our project. Batty et al. [9] analyze human dynamics and social interactions of mobile phone users during large-scale events and Calabrese et al. [10] investigate the relationship between different types of events and the home area of the attendees.

We have recently developed a methodology to analyze the workload dynamics of a mobile phone network during large-scale events, with an initial focus on soccer matches [4]. In that work, we characterized how mobile phone users move during part of the day, around the time and location where major soccer matches took place. We analyzed soccer matches in the same location (one soccer stadium) on different days, aiming at comparing the workloads imposed on the carrier's infrastructure (the same set of antennas near the stadium) on days with and without matches. The analysis from a single point of interest also makes it easier to determine how people move towards the stadium before the match starts and where they go afterwards, following an approach similar to [10].

In this paper, we apply our proposed methodology to analyze another kind of large-scale event, i.e. New Year's Eve celebrations, similar to what has been done by [9]. We also include in our study the analysis of three new metrics: (i) call durations, (ii) call inter-arrival, and (iii) call inter-departure times, extending what was done in our previous work [4].

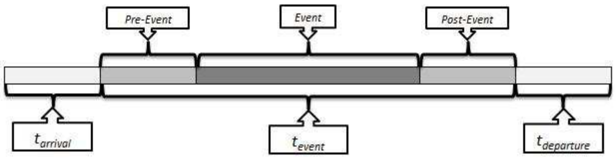

Fig. 1. Timeline adopted in our methodology.

III. Proposed Methodology

Our proposed methodology [4] aims at providing insights for better management and capacity planning of the carrier's infrastructure to support the demands of cellphone users due to their mobility during large-scale events. Basically, our methodology is designed with the purpose of answering three main questions: (i) *who moved towards the surroundings of the large-scale event when it took place?* (ii) *where did they come from?* and (iii) *where did they go after the event?*

Towards answering such questions, we restrict our analysis to cellphone calls made during a period of time around the event time. Specifically, we adopt the *timeline* notation shown in Figure 1, which is discussed next.

The first step of our methodology is to identify the antennas that cover the region where the event was held. We then select the users who made at least one cellphone call in one of the selected antennas before and after the event. To that end, we define the *Pre-Event* and *Post-Event* periods as the time intervals that cover the k minutes preceding the beginning of the event and following its end, respectively. We refer to the period from the beginning of *Pre-Event* until the end of *Post-Event*, including the event's total duration, as t_{event}. Users who made at least one call during t_{event} in one of the selected antennas are considered *attendees*. The durations of these time intervals, i.e., determining k, somewhat depends on local aspects and on the event characteristics.

In the second step, we identify the subset of event attendees who also made calls before the beginning (i.e., before *Pre-Event*) or after the end (i.e., after *Post-Event*) of the event. These calls allow us to track the movements of these users before or after the event duration. Thus, our analysis of mobility patterns focuses on these selected users. To identify these users, we define two other periods ($t_{arrival}$ and $t_{departure}$) corresponding to the time intervals before and after the t_{event} period, respectively, during which people are moving towards and arriving at the event place as well as departing from the location.

At this point, the third and last step of our methodology, we use the geolocation of the antennas to locate where each call started and ended, thus determining from where the event attendees (identified in the previous step) came and to where they moved after the event. Ultimately, this enables us to analyze the workload dynamics of the antennas located along the main routes to and from where the large-scale event takes place. To visualize such dynamics, we use *heat maps*[1] to represent the intensity of activity in the mobile phone network (the darker the color in the heat map, the larger the number of calls received by the antenna in this area).

[1]Heat maps are generated using the Google Maps Javascript API V3.

TABLE I. Overview of the Datasets (event days in bold).

City	Day	# Calls done by Attendees	# Attendees	Average Calls per Attendees
BH	**Dec 31, 2011**	**5187**	**1938**	**2.7**
BH	Jan 03, 2012	779	365	2.1
Recife	**Dec 31, 2011**	**9951**	**3566**	**2.8**
Recife	Jan 03, 2012	924	444	2.1
Salvador	**Dec 31, 2011**	**12826**	**7458**	**1.7**
Salvador	Jan 03, 2012	1019	689	1.5
Rio	**Dec 04, 2011**	**4284**	**1754**	**2.4**
Rio	Oct 30, 2011	1270	691	1.8

IV. Results

Our methodology was applied to datasets containing mobile phone calls made during the 2012 New Year's Eve celebrations in three large Brazilian cities: Belo Horizonte (BH), Recife, and Salvador. The datasets contain for each call a unique user identifier[2], the geographical locations (latitude and longitude) of the antennas where the call started and ended, as well as the time instants when it started and ended.

To define the event location and associated antennas, we considered large-scale New Year's Eve celebrations organized in each city, such as a celebration on a beach in Salvador which received, reportedly, one million attendees. Six antennas covered this area. The celebrations in BH and Recife, hosted in an area covered by 3 antennas each, received around 100,000 and 10,000 people, respectively. Regarding the proposed timeline, we considered the total period from 9:45PM to 2:30AM, with the event starting at 11:15PM and lasting for 105 minutes. We set the durations of $t_{arrival}$, $t_{departure}$, *Pre-Event* and *Post-Event* to 45 minutes each. For comparison purposes, we also analyzed other datasets from the same cities, collected on January 3rd, 2012 (a day without event) using the same antennas and timeline. Similarly, we compared these results with those obtained for another type of large-scale event - a soccer match in Rio on December 4th 2011 - reported in [4]. The antennas covering the location (stadium) of the match and the timeline were defined using the same methodology, taking the time and soccer match stadium into account. As basis for comparison, we also analyzed data collected on October 30th corresponding to the same antennas and timeline of the match.

Table I summarizes the analyzed datasets, presenting the total numbers of calls and attendees. Note that the numbers of calls are increased by a factor between 6.7 and 12.6 during the New Year's Eve celebrations (BH, Recife, and Salvador), comparing with the numbers of a day without event at each city. This could be expected if we consider the time (late at night), the event nature, and the huge number of attendees. Note also that the growth factor of the number of calls during the soccer match, comparing with a day with no events in Rio (two last rows of Table I), is lower (factor of 3.4). This indicates the importance of the event nature for better understanding the human mobility and the workload dynamics of the mobile phone network.

To further understand the characteristics of the workload imposed on the selected antennas during each event, Figure 2 shows the numbers of calls made by attendees of the

[2]This id is generated by the cellphone carrier. It is completely anonymized, and thus, cannot be used for identifying the user; although, it allows us to identify multiple calls made by the same user in different points in time.

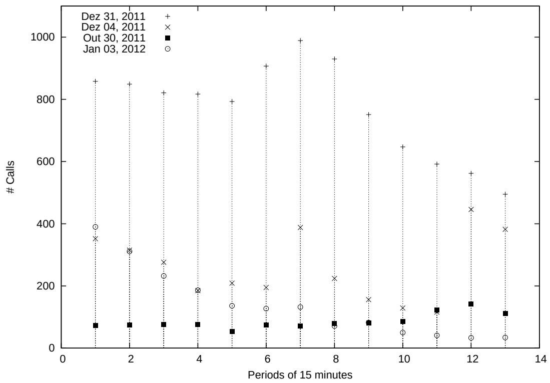

Fig. 2. Calls made by attendees of Recife's New Year's Eve (Dec 31, 2011), at one soccer match in Rio (Dec 04, 2011), and on two days with no large-scale event in Rio (Oct 30, 2011) or in Recife (Jan 03, 2012).

celebration in Recife, in successive 15-minute time bins during t_{event}. For comparison purposes, corresponding results for a day with a soccer match (Dec 04, 2011) and two days without event (Oct 30, 2011 in Rio and Jan 03, 2012 in Recife) are also shown in Figure 2. This figure shows that, for the day of Recife's New Year's Eve, the number of calls peaks at around midnight (bins 6–8), dropping quickly after the celebration finishes. For a soccer match and regular days, instead, the number of calls shows different patterns, decreasing sharply in the ending of t_{event}. Results for the other cities are very similar, being thus omitted.

Call durations, inter-arrival times (IAT), and inter-departure times (IDTs) are key metrics for capacity management and planning as they allow us to assess whether an antenna throughput is meeting its imposed load. We computed all these metrics over 15-minute intervals for each New Year's Eve celebration and soccer matches. In general, all events are typically composed of a large volume of short calls. We also observed that the average of IATs and IDTs are similar for each celebration and slightly different of days with soccer matches. This also points out the difference of the workload imposed on a mobile phone network during events of distinct nature.

Finally, in our methodology, we use heat maps to analyze the mobility of event attendees before, during and after the analyzed events, which allows us to infer the most used access routes to/from the event's location. Figures 3 and 4 show the heat maps produced for the $t_{arrival}$ and t_{event} periods of the New Year's Eve celebration in Salvador. The observation of these heat maps in sequence illustrates human mobility towards the location of the large-scale event.

V. Conclusions and Future Work

In this paper, we extended our recently proposed methodology [4], applying it to analyze human mobility and the workload dynamics of a mobile phone network during large-scale events. Our results show that user behavior patterns that arise during different kinds of events can be used to analyze phenomena related to human mobility due to such events. As future work, we intend to expand our analysis to include other types of large-scale events, like major concerts. We believe that identifying mobility patterns based on cellphone calls during large-scale events can drive the development of target applications and gadgets tailored for such events.

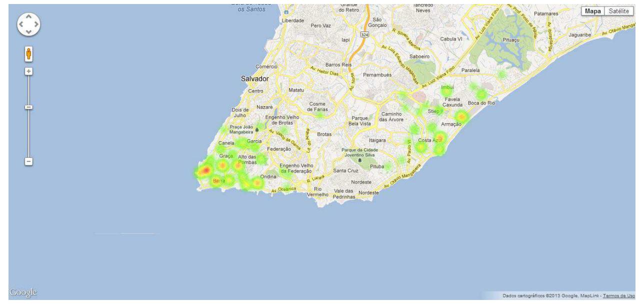

Fig. 3. Heat map of New Year's Eve at Salvador during its $t_{arrival}$.

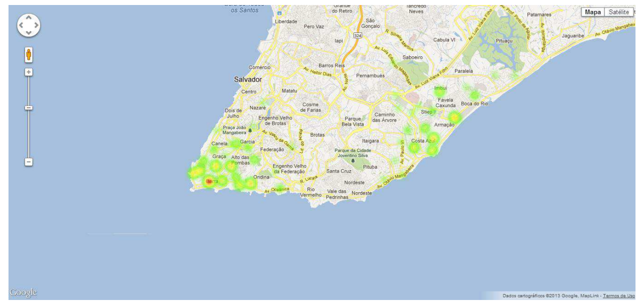

Fig. 4. Heat map of New Year's Eve at Salvador during its t_{event}.

Acknowledgment

This work is supported by FIP-PUC MINAS, InWeb (MCT/CNPq 573871/2008-6), CNPq, CAPES, FAPEMIG, and FAPERJ.

References

[1] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási, "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, 2010.

[2] A. Bleicher, "The on-demand olympics," *IEEE Spectrum*, vol. 49, no. 7, pp. 9–10, jul 2012.

[3] J. P. Bagrow, D. Wang, and A.-L. Barabási, "Collective response of human populations to large-scale emergencies," *PLoS ONE*, vol. 6, no. 3, p. e17680, 03 2011.

[4] F. H. Z. Xavier, L. M. Silveira, J. M. Almeida, A. Ziviani, C. H. S. Malab, and H. Marques-Neto, "Analyzing the workload dynamics of a mobile phone network in large scale events," in *Proceedings of the UrbaNe Workshop – ACM CoNEXT 2012*, 2012.

[5] N. Eagle, A. Pentland, and D. Lazer, "Inferring social network structure using mobile phone data," *Proceedings of the National Academy of Sciences (PNAS)*, vol. 106, no. 36, pp. 15 274–15 278, 2009.

[6] M. C. González, C. A. Hidalgo, and A.-L. Barabási, "Understanding individual human mobility patterns," *Nature*, vol. 453, pp. 779–782, 2008.

[7] J. Yuan, Y. Zheng, and X. Xie, "Discovering regions of different functions in a city using human mobility and POIs," in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, ser. KDD '12. New York, NY, USA: ACM, 2012, pp. 186–194.

[8] D. Balcan, V. Colizza, B. Gonçalves, H. Hu, J. J. Ramasco, and A. Vespignani, "Multiscale mobility networks and the spatial spreading of infectious diseases," *Proceedings of the National Academy of Sciences*, vol. 106, no. 51, pp. 21 484–21 489, 2009.

[9] M. Batty, J. Desyllas, and E. Duxbury, "The discrete dynamics of small-scale spatial events: agent-based models of mobility in carnivals and street parades," *International Journal of Geographical Information Science*, vol. 17, no. 7, pp. 673–697, 2003.

[10] F. Calabrese, F. Pereira, G. Di Lorenzo, L. Liu, and C. Ratti, "The geography of taste: Analyzing cell-phone mobility and social events," in *Pervasive Computing*, ser. Lecture Notes in Computer Science, P. Floréen, A. Krüger, and M. Spasojevic, Eds. Springer Berlin / Heidelberg, 2010, vol. 6030, pp. 22–37.