TB, DC, US, JPhysA/263312, 12/12/2007

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **41** (2008) 000000 (11pp) UNCORRECTED PROOF

Uncovering individual and collective human dynamics from mobile phone records

Julián Candia[1,2], Marta C González[1,2], Pu Wang[1,2], Timothy Schoenharl[3], Greg Madey[3] and Albert-László Barabási[1,2,4]

[1] Center for Complex Network Research and Department of Physics, Northeastern University, Boston, MA 02115, USA
[2] Department of Physics, University of Notre Dame, Notre Dame, IN 46556, USA
[3] Department of Computer Science and Engineering, University of Notre Dame, Notre Dame, IN 46556, USA
[4] Collegium Budapest, Szentháromság u. 2, H-1014 Budapest, Hungary

Received 16 October 2007, in final form 28 November 2007
Published DD MMM 2007
Online at stacks.iop.org/JPhysA/41/000000

Abstract

Novel aspects of human dynamics and social interactions are investigated by means of mobile phone data. Using extensive phone records resolved in both time and space, we study the mean collective behavior at large scales and focus on the occurrence of anomalous events. We discuss how these spatiotemporal anomalies can be described using standard percolation theory tools. We also investigate patterns of calling activity at the individual level and show that the interevent time of consecutive calls is heavy-tailed. This finding, which has implications for dynamics of spreading phenomena in social networks, agrees with results previously reported on other human activities.

PACS number: **See endnote 1**

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Mobile phones are becoming increasingly ubiquitous throughout large portions of the world, especially in highly populated urban areas and particularly in industrialized countries, where mobile phone penetration is almost 100%. Mobile phone providers regularly collect extensive data about the call volume, calling patterns, and the location of the cellular phones of their subscribers. In order for a mobile phone to place outgoing calls and to receive incoming calls, it must periodically report its presence to nearby cell towers, thus registering its position in the geographical cell covered by one of the towers. Hence, very detailed information on the spatiotemporal localization of millions of users is contained in the extensive call records of any mobile phone carrier. If misused, these records—as well as similar datasets on buying habits, e-mail usage and web-browsing, for instance—certainly pose a serious threat to the privacy of

the users. However, the use of privacy-safe, anonymized datasets represent a huge scientific opportunity to uncover the structure and dynamics of the social network at different levels, from the small-scale individual's perspective to the large-scale, collective behavior of the masses, with an unprecedented degree of reach and accuracy. Besides the inherent scientific interest of these issues, deeper insight into applications of great practical importance could certainly be gained. For instance, urban planning, public transport design, traffic engineering, disease outbreak control and disaster management are some areas that will greatly benefit from a better understanding of the structure and dynamics of social networks [1].

The use of mobile phone data as a proxy for social interaction has already proved successful in several recent investigations. Onnela *et al* [2, 3] have analyzed the structure of weighted call graphs arising from reciprocal calls that serve as signatures of work-, family-, leisure- or service-based relationships. A coupling between interaction strengths and the network's local structure was observed, with the counterintuitive consequence that social networks turn out to be robust to the removal of the strong ties but fall apart following a phase transition if the weak ties are removed. Szabó and Barabási [4] have studied social network effects in the spread of innovations, products and new services. They investigated different mobile phone-based services and found the coexistence on the same social network of two distinct usage classes, with either very strong or very weak community-based segregation effects. In the context of urban studies and planning, Ratti *et al* [5, 6] have considered the potential use of aggregated data from mobile phones and other hand-held devices. Their 'Mobile Landscapes' project aims at the application of location-based services to urban studies in order to gain insight into complex and rapidly changing urban dynamics phenomena. More recently, Palla, Barabási and Vicsek [7, 8] used mobile phone data to study the evolution of social groups. They found that large groups persist for longer times if they are capable of dynamically altering their membership, suggesting that an ability to change the group composition results in better adaptability. In contrast, the behavior of small groups displays the opposite tendency, the condition for long-term persistence being that their composition remains stable.

In the following sections, we present new results that address novel aspects of human dynamics and social interactions obtained from extensive mobile phone data. In section 2, we show how large-scale collective behavior can be described using aggregated data resolved in both time and space. We stress the importance of investigating large departures from the average and develop the basic framework to quantify anomalous fluctuations by means of standard percolation theory tools. In section 3, we focus on the individual level and study patterns of calling activity. We show that the interevent time of consecutive calls is heavy-tailed, a finding that has implications for the dynamics of spreading on social networks [9–17]. Furthermore, by fixing the time of observation between consecutive calls it is possible to use the phone call data to characterize some aspects of human mobility.

2. Fluctuations in aggregated spatiotemporal call activity patterns

The spatial dependence of the call activity at any given time can be conveniently displayed by means of maps divided in Voronoi cells, which delimit the area of influence of each transceiver tower or antenna. The Voronoi tessellation partitions the plane into polygonal regions, associating each region with one transceiver tower. The partition is such that all points within a given Voronoi cell are closer to its corresponding tower than to any other tower in the map.

Figure 1 shows activity maps for aggregated data corresponding to a 1 h interval. The upper panel shows the activity pattern (in $\log_{10}$ scale) for a peak hour (Monday noon), while the lower panel shows the same urban neighborhood during an off-peak hour (Sunday at 9 am).

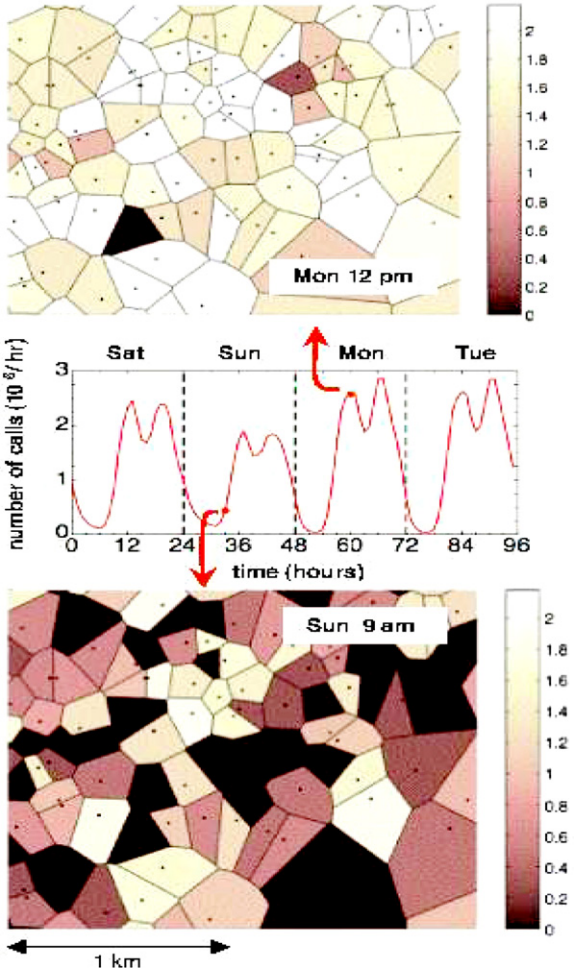

Figure 1. Call activity maps in an urban neighborhood, showing the number of calls per hour managed by each transceiver tower or antenna (dots). The division in terms of Voronoi cells defines the area of reach of each tower. Call traffic patterns depend on time and day of the week, as shown by comparing the map on a Monday at noon (upper panel) with that on a Sunday at 9 am (lower panel). The bars on the right side of each panel correspond to the number of calls per hour and tower in $\log_{10}$ scale.

The differences between both panels reflect the intrinsic rhythm and pulse of the city: we can expect call patterns during peak hours to be dominated by the hectic activity around business and office areas, whereas other, presumably residential and leisure areas can show increased activity during off-peak times, thus leading to different, spatially distinct activity patterns. Besides different spatial patterns, each particular time of the day, as well as each day of the week, is characterized by a different overall level of activity. This phenomenon is shown by the plot at the center of figure 1, in which aggregated data for a country are shown as a function of time (data were binned in time intervals of 1 h). As expected, the overall normalization of the aggregated pattern is lower during weekends than during weekdays, except around weekend midnights and early mornings, when many people go out.

The minimum spatial resolution is determined by either the typical distance between towers or, in rural regions with sparse tower density, by the reach of the radio-frequency signals exchanged between the mobile handset and the antenna (typically ranging from a few hundred meters to several kilometers). To explore activity differences at larger scales, the data of neighboring cells can be aggregated. At the expense of some loss of spatial resolution, aggregating data into larger spatial bins (taking, e.g., a regular spatial grid covering the entire country) allows for better statistics and for a more stable activity pattern. That is, the number of calls made from a group of nearby cells at a certain time and day of the week is expected to be fairly constant, except for small statistical fluctuations.

Usually, activity patterns are strongly correlated with the daily pulse of populated areas (such as those shown in figure 1) and, at a larger scale, to variations in population density between different regions within the country. In contrast, departures from the mean expected activity are in general not trivially correlated with population density and describe instead interesting dynamical features.

The measurement of fluctuations around the mean expected activity is of paramount importance, since it allows a quantitative measurement of anomalous behavior and, ultimately, of possible emergency situations. This indeed constitutes the base of proposed real-time monitoring tools such as the *Wireless Phone-based Emergency Response* (WIPER) system [18]. Anomalous patterns indicative of a crisis (such as the occurrence of natural catastrophes and terrorist attacks) could be detected in real time, plotted on satellite and GIS-based maps of the area, and used in the immediate evaluation of mitigation strategies, such as potential evacuation routes or barricade placement, by means of computer simulations [18, 19].

The call volume shows strong variations with time and day of the week, as shown in figure 1, but differences across subsequent weeks are generally mild (provided one considers call traffic in the same place, time and day of the week). To capture the weekly periodicity of the observed patterns, we define $n_i(\mathbf{r}, t, T)$ as the number of calls recorded at location $\mathbf{r}$ (which can either denote a single Voronoi cell or a group of neighboring cells) during the ith week between times t and $t + T$, where time is defined modulo 1 week. Assuming we have access to continuous data for N weeks, the mean call activity is given by

$$\langle n(\mathbf{r}, t, T)\rangle = \frac{1}{N}\sum_{i=1}^{N} n_i(\mathbf{r}, t, T). \tag{1}$$

Note that in the same way as one can trade off spatial resolution for increased statistics by summing over a group of Voronoi cells, varying T one can regulate time accuracy versus statistics. This certainly depends on the extent to which aggregated data show a regular, stable behavior. The results presented here correspond to $T = 1$ h.

The scale to measure departures from the average behavior is set by the *standard deviation*, defined as

$$\sigma(\mathbf{r}, t, T) = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(n_i(\mathbf{r}, t, T) - \langle n(\mathbf{r}, t, T)\rangle)^2}. \tag{2}$$

Hence, using recorded data for an extended period of time, one can determine the expected call traffic levels and corresponding deviations for all times and locations. Once this *normal* behavior is established, *anomalous* fluctuations above or below a given threshold can be obtained using the condition

$$|n_i(\mathbf{r}, t, T) - \langle n(\mathbf{r}, t, T)\rangle| > A_{\text{thr}} \times \sigma(\mathbf{r}, t, T), \tag{3}$$

where $A_{\text{thr}} > 0$ is a constant that sets the fluctuation level.

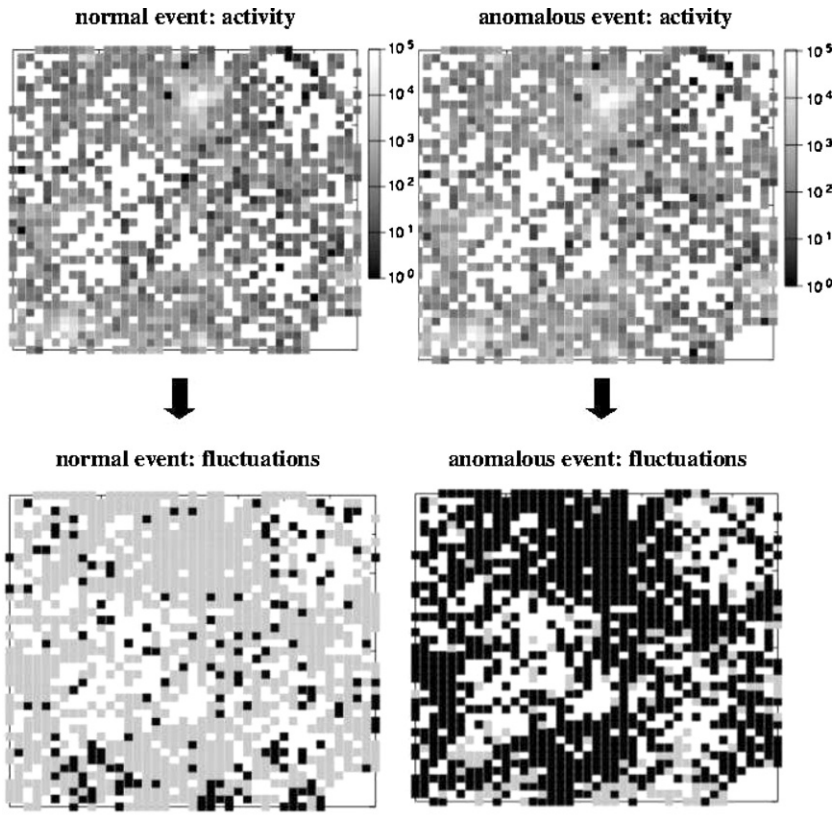

Figure 2. Activity and fluctuations in a regular 2D grid showing a normal event (left panels) and an anomalous one (right panels). The activity is displayed in terms of the number of calls per hour inside each square bin in $\log_{10}$ scale (upper panels). High-activity bins above the fluctuation threshold $A_{\rm thr} = 0.25$ are shown in black, while bins with normal activity are shown in gray (bottom panels). Bins in white correspond to areas not covered by the mobile phone carrier.

We grouped Voronoi cells together generating a regular 2D grid made of square bins of about 12 km of linear size. Considering a fixed time slice, we study the spatial clustering of bins showing anomalous activity at different fluctuation levels. In order to illustrate our procedure, figure 2 shows the activity and fluctuations in a grid of size 40×40 bins (i.e. 480×480 km^2 area). We compare the activity in the same region for 2 different weeks (corresponding to the same time and day of the week). The left panels show a *normal event*, in which fluctuations around the local mean activity are typically small, with just a few scattered bins having somewhat larger deviations. The right panels, however, show an *anomalous event*, characterized by extended, spatially correlated fluctuations that indicate the emergence of a large-scale, coordinated activity pattern. As pointed out above, the existence of anomalous activity patterns could be indicative of possible emergency situations. Similarly to the Voronoi maps already discussed, the upper panels in figure 2 show the activity (number of calls per hour inside each square bin) in $\log_{10}$ scale. White bins correspond to areas not covered by the mobile phone provider. Taking a fixed threshold value $A_{\rm thr} = 0.25$, the bottom panels show the high-activity bins above the fluctuation threshold (in black) and the bins with normal activity (in gray). Note that although the activity maps have a similar appearance to the

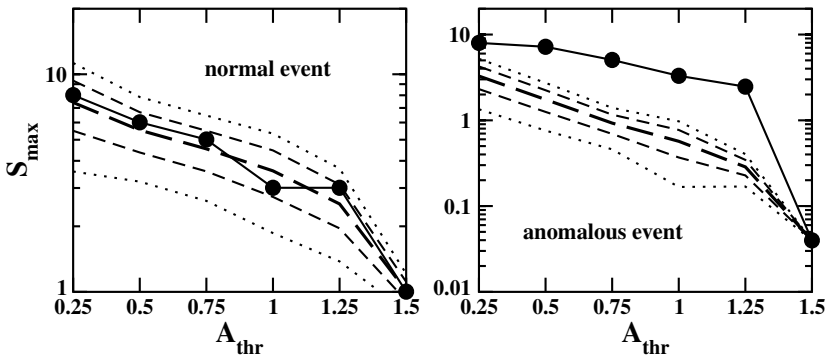

Figure 3. Size of the largest cluster as a function of the fluctuation threshold for the normal case (left) and the anomalous one (right). Measurements on the call data (solid line with circles) are compared to those of randomized distributions, of which we show the mean (long-dashed line) and confidence bounds at $\pm\sigma_{\rm rdm}$ (short-dashed lines) and $\pm 2\sigma_{\rm rdm}$ (dotted lines).

degree that they seem at first look indistinguishable, the fluctuation maps display striking differences.

In order to quantify the clustering of anomalous bins, we will use the standard tools of percolation theory and determine the size of the largest cluster, the number of different clusters and the size distribution of all clusters. The statistical significance of the measured clustering is evaluated by comparing it to results from randomized distributions, in which many different configurations are randomly generated, keeping fixed the total number of high-activity bins above the fluctuation threshold. The substrate, which is formed by all bins with non-zero activity, remains always the same (in figure 2, for instance, the substrate is the set of all gray and black bins). Clusters are defined by first- and second-order nearest neighbors in the square 2D grid. In the remainder of this section, we will focus on a specific large-scale anomalous event and compare it to the normal behavior observed in data of a different week (but corresponding to the same time and day of the week). The comparison between normal and anomalous events will illustrate the use of percolation observables as diagnostic tools for anomaly detection.

Figure 3 shows the size of the largest cluster, $S_{\rm max}$, as a function of the fluctuation threshold $A_{\rm thr}$, for the normal case (left) and the anomalous one (right). Each measured plot (solid line with circles) is compared to results from randomized distributions. The latter correspond to the mean (long-dashed line) and confidence bounds at $\pm\sigma_{\rm rdm}$ (short-dashed lines) and $\pm 2\sigma_{\rm rdm}$ (dotted lines), as obtained from generating 100 random configurations in each case. As expected, the plots show that the size of the largest cluster monotonically decreases with the fluctuation threshold. However, while the clustering in the normal case lacks any significance, the anomalous event shows large departures from the clustering expected in a random configuration.

In the same vein, figure 4 shows the number of different clusters, $N_{\rm cl}$, as a function of the fluctuation threshold $A_{\rm thr}$, where measurements on the call data for the same normal (left) and anomalous (right) events are compared to results from randomized configurations. As before, in the normal case the number of clusters agrees well with the expectations for random configurations, while significant departures are observed in the anomalous case.

Figure 5 shows the cumulative size distribution of all clusters, $N_{\rm cl}(s_{\rm cl} > S)$, as a function of the cluster size S, compared to random configurations. The upper panels display results

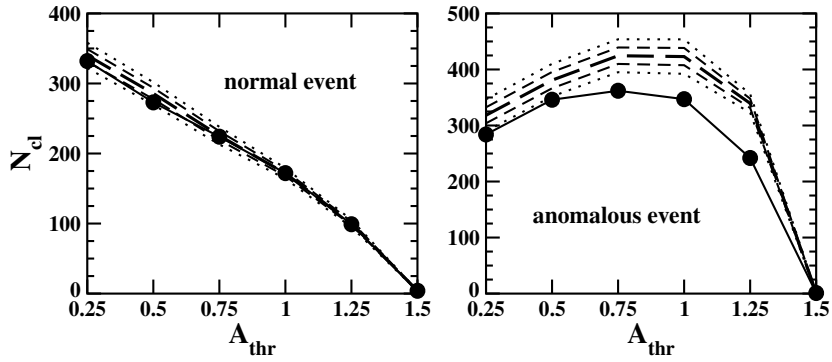

Figure 4. Number of different clusters as a function of the fluctuation threshold for the normal case (left) and the anomalous one (right). Measurements on the call data (solid line with circles) are compared to results on random configurations (dashed and dotted lines).

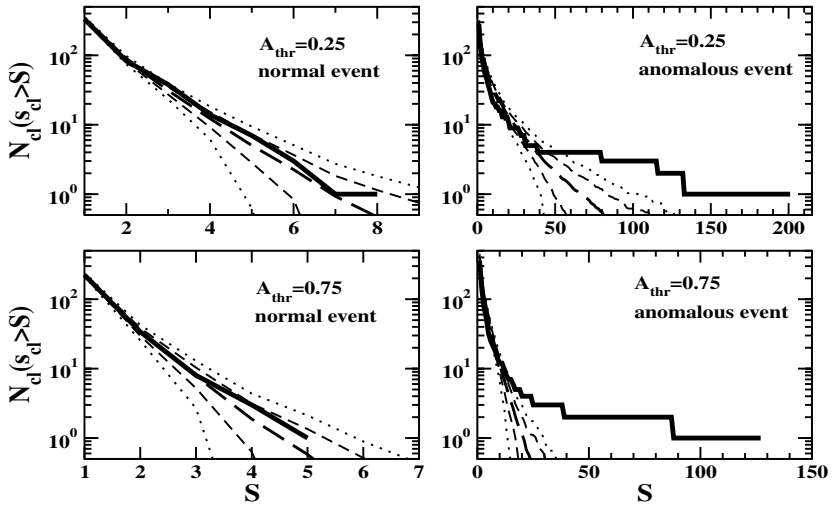

Figure 5. Cumulative size distribution of all clusters as a function of cluster size, for $A_{\rm thr} = 0.25$ (upper panels), $A_{\rm thr} = 0.75$ (bottom panels), normal case (left panels), and anomalous case (right panels). Thick solid lines are measurements on the call data, while dashed and dotted lines are results from random configurations.

for $A_{\rm thr} = 0.25$, while the bottom ones show results for $A_{\rm thr} = 0.75$, as indicated. Moreover, the left panels correspond to the normal event, while the right panels to the anomalous event. Again, the measured cluster size distribution in the normal case is in good agreement with the expected one for a random configuration. In contrast, the anomalous event shows the occurrence of a few very large clusters formed by many highly active bins. These unusually large structures cannot be explained as arising just from random configurations, but instead are the result of the spatiotemporal correlation of large, highly active regions.

As a summary, in this section we showed how large-scale collective behavior can be described using aggregated data resolved in both time and space. Moreover, we developed the basic framework for detecting and characterizing spatiotemporal fluctuation patterns, which is

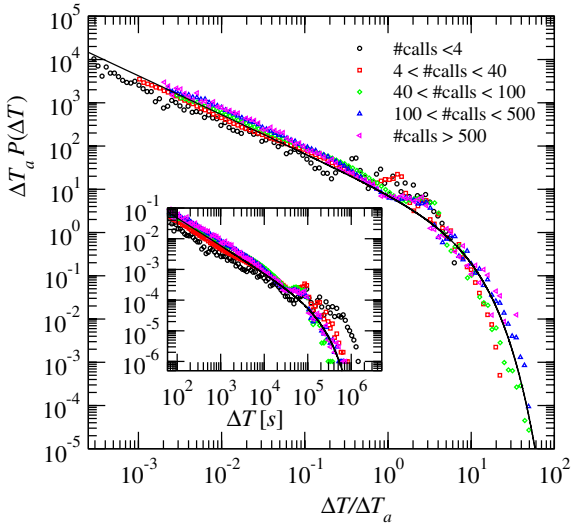

Figure 6. Interevent time distribution $P(\Delta T)$ for calling activity. ΔT corresponds to the time interval between two mobile phone calls sent by the same user. Different symbols indicate the measurements done over groups of users with different activity levels (# calls). The inset shows the unscaled interevent time distribution and the solid line corresponds to equation (4).

based on standard procedures of statistics and percolation theory. These tools are particularly effective in detecting extended anomalous events, as those expected to occur in emergency scenarios due to e.g. natural catastrophes and terrorist attacks.

3. Individual calling activity patterns

In order to use the huge amount of data recorded by mobile phone carriers to investigate various aspects of human dynamics [1, 20–23], a necessary starting point is to characterize the dynamics of the individual calling activity *per se*. Previous studies have measured the time between consecutive individual-driven events, such as sending e-mails, printing and visiting web pages or the library [24, 25]. Those events are described by heavy-tailed processes [20, 26], challenging the traditional Poissonian modeling framework [27–31], with consequences on task completion in computer systems. In this section, we explore the interevent distribution of the calling activity of 6×10^6 mobile phone users during 1 month.

As many other human activities, the calling activity pattern is highly heterogeneous. While some users rarely use the mobile phone, others make hundreds or even thousands of calls each month. To analyze such different levels of activity, we group the users based on their total number of calls. Within each group, we measure the probability density function $P(\Delta T)$ of the time interval ΔT between two consecutive calls made by each user. As shown by the inset of figure 6, the tail of the distribution is shifted to longer interevent times for users with less activity. However, if we plot $\Delta T_a P(\Delta T)$ as a function of $\Delta T/\Delta T_a$, where ΔT_a is the average interevent time for the corresponding user, the data collapse into a single curve (figure 6). This indicates that the measured interevent distribution follows the expression $P(\Delta T) = 1/\Delta T_a \mathcal{F}(\Delta T/\Delta T_a)$, where $\mathcal{F}(x)$ is independent from the average activity level of

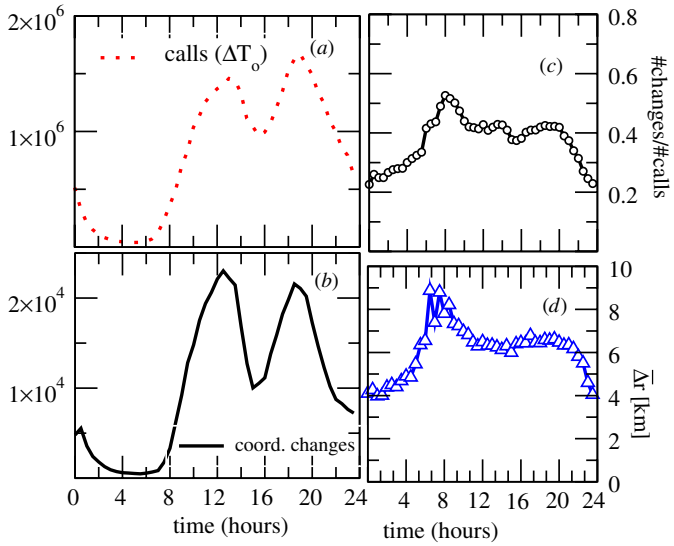

Figure 7. Travel behavior. (*a*), (*b*) Number of trips and consecutive calls that are reported within a fixed interevent time $\Delta T_o = 30$ min versus time of the day. (*c*) The ratio of the two quantities described in (*a*) and (*b*) shows that along the whole day $40 \pm 20\%$ of the people that is calling seems to be also traveling. (*d*) The average distance of travel within $\Delta T_o = 30$ min remains constant during the day within 6 ± 2 km, a reasonable value that may correspond to the combination between walk and motor transportation.

the population. This represents a universal characteristic of the system that surprisingly also coincides with results from e-mail communication [32]. The data are well fitted by

$$P(\Delta T) = (\Delta T)^{-\alpha} \exp(\Delta T/\tau_c), \tag{4}$$

where the power-law scaling with exponent $\alpha = 0.9 \pm 0.1$ is followed by an exponential cutoff at $\tau_c \approx 48$ days. Equation (4) is shown by a solid line in the inset of figure 6 and its scaled version is presented in the main panel of the figure using $\Delta T_a = 8.2$ h, which is the average interevent time measured for the whole population. This result, clearly different from that predicted by a Poisson approximation [26, 33, 34], would for instance affect the predictions of spreading dynamics through the network of calls [35].

To explore the interplay between human activity and mobility patterns, we fix the characteristic observation time to $\Delta T_o = 30$ min and collect only those consecutive calls that occur with this interevent time, recording also the time of the day in which they occurred (figure 7(*a*)). For each pair of calls, we count how many of them result in a change of coordinate, e.g. the user traveled in the 30 min time interval between the calls (figure 7(*b*)). The number of events that result in a change of location and the number of calls as a function of time capture the daily activity pattern of the users [36]. We find that both the call and the mobility pattern decrease at night and have clear peaks near noon and late evening. There is a factor of 30 between the largest and the smallest number of events (calls/changes of location) reported during the day. Interestingly, when we calculate the fraction of consecutive calls also resulting in a potential change of location, the quantity varies at most 40% during the whole day (figure 7(*c*)). This indicates that although the total activity varies strongly, the percentage of the people that are calling and traveling remains rather stable. More importantly,

the average distance traveled within $\Delta T_o = 30$ min is stable in the vicinity of $\Delta r = 6 \pm 2$ km (figure 7(*d*)), a value consistent for the combination between walk and motor transportation.

4. Conclusions

Novel aspects of human dynamics and social interactions were addressed by means of mobile phone data with time and space resolution. This allowed us to study the mean collective behavior at large scales and focus on the occurrence of anomalous events. Considering a fixed time slice, we partitioned the space using a regular grid and studied the aggregated call activity inside each square bin forming the grid. We showed that anomalous events give rise to spatially extended patterns that can be meaningfully quantified in terms of standard percolation observables. By considering a series of consecutive time slices, we could investigate the rise, clustering and decay of spatially extended anomalous events, which could be relevant e.g. in real-time detection of emergency situations.

We also investigated patterns of calling activity at the individual level. We observed that the interevent time of consecutive calls is heavy-tailed, a finding that has implications for dynamics of spreading phenomena on social networks, and that agrees with results previously reported on other, related human activities. We also show that despite the complexity inherent in the interevent calling patterns, it is still possible to recover some characteristic values from the behavior of the population that are stationary during the day, such as the fraction of active traveling population and their average distance traveled.

In many ways, these results represent only a first step toward understanding human activity patterns. Our results indicate that the rich information provided by mobile communication data open avenues to addressing novel problems. These tools offer a chance to improve our understanding of complex networks as well [37–44], by potentially correlating the structure of social networks with the spatial layout of the users as nodes [45–51], thus contributing to a better understanding of the spatiotemporal features of network evolution.

Acknowledgments

This work was supported by the James S McDonnell Foundation 21st Century Initiative in Studying Complex Systems, the National Science Foundation within the DDDAS (CNS-0540348), ITR (DMR-0426737) and IIS-0513650 programs, as well as by the US Office of Naval Research Award N00014-07-C and the NAP Project sponsored by the National Office for Research and Technology (KCKHA005). Data analysis was performed on the Notre Dame Biocomplexity Cluster supported in part by NSF MRI Grant No DBI-0420980.

References

[1] González M C and Barabási A-L 2007 *Nature Phys.* **3** 224
[2] Onnela J-P, Saramäki J, Hyvönen J, Szabó G, Lazer D, Kaski K, Kertész J and Barabási A-L 2007 *Proc. Natl Acad. Sci.* **104** 7332
[3] Onnela J-P, Saramäki J, Hyvönen J, Szabó G, de Menezes M A, Kaski K, Barabási A-L and Kertész J 2007 *New J. Phys.* **9** 179
[4] Szabó G and Barabási A-L 2006 *Preprint* physics/0611177
[5] Ratti C, Pulselli R M, Williams S and Frenchman D 2006 *Environ. Planning* B **33** 727
[6] Ratti C, Sevtsuk A, Huang S and Pailer R 2007 *Location Based Services and TeleCartography* (Berlin: Springer) section V, p 433
[7] Palla G, Barabási A-L and Vicsek T 2007 *Nature* **446** 664
[8] Palla G, Barabási A-L and Vicsek T 2007 *Fluct. Noise Lett.* **7** L273

[9] Pastor-Satorras R and Vespignani A 2001 *Phys. Rev. Lett.* **86** 3200
[10] Eubank S, Guclu H, Kumar V S A, Marathe M, Srinivasan A, Toroczkai Z and Wang N 2004 *Nature* **429** 180
[11] Vibpoud C, Bjonstadt O, Smith D L, Simonsen L, Miller M A and Grenfell B T 2006 *Science* **312** 447
[12] Colizza V, Barrat A, Barthelemy M, Valleron A-J and Vespignani A 2007 *PLoS Med.* **4** e13
[13] González M C and Herrmann H J 2004 *Physica* A **340** 741
[14] González M C, Herrmann H J and Araújo A D 2005 *Physica* A **356** 100
[15] Candia J 2006 *Phys. Rev.* E **74** 031101
[16] Candia J 2007 *Phys. Rev.* E **75** 026110
[17] Candia J 2007 *J. Stat. Mech.* P09001
[18] Madey G, Szabó G and Barabási A-L 2006 *Lecture Notes in Computer Science* vol 3993 ed V N Alexandrov, G D van Albada, P M A Sloot and J Dongarra (Berlin: Springer) p 417
[19] Schoenharl T, Bravo R and Madey G 2007 *Int. J. Int. Contr. Syst.* **11** 209
[20] Barabási A-L 2005 *Nature* **435** 207–11
[21] Vázquez A 2007 *Physica* A **373** 747
[22] Dezsö Z, Almaas E, Lukács A, Rácz B, Szakadát I and Barabási A-L 2006 *Phys. Rev.* E **73** 066132
[23] Helbing D, Treiber M and Kesting A 2006 *Physica* A **363** 62
[24] Oliveira J G and Barabási A-L 2005 *Nature* **437** 1251
[25] Harder U and Paczuski M 2006 *Physica* A **361** 329
[26] Vázquez A, Oliveira J G, Dezsö Z, Goh K-I, Kondor I and Barabási A-L 2006 *Phys. Rev.* E **73** 036127
[27] Vázquez A 2005 *Phys. Rev. Lett.* **95** 248701
[28] Gabrielli A and Caldarelli G 2007 *Phys. Rev. Lett.* **98** 20
[29] Blanchard P and Hongler M O 2007 *Phys. Rev.* E **75** 026102
[30] Daly E and Porporato A 2007 *Phys. Rev.* E **75** 011119
[31] Hidalgo C 2006 *Physica* A **369** 877
[32] Goh K-I and Barabási A L 2006 *Preprint* physics/0610233
[33] Feller W 1966 *An Introduction to Probability Theory and its Applications* vol II (New York: Wiley)
[34] Laherrère J and Sornette D 1998 *Eur. Phys. J.* B **2** 525
[35] Vázquez A, Rácz B, Lukács A and Barabási A-L 2007 *Phys. Rev. Lett.* **98** 158702
[36] Golder S A, Wilkinson D and Huberman B A 2007 *Proc. 3rd Int. Conf. on Communities and Technologies (CT2007) (East Lansing, MI)* pp 28–30
[37] Barabási A-L and Albert R 1999 *Science* **286** 509
[38] Albert R and Barabási A-L 2002 *Rev. Mod. Phys.* **74** 47
[39] Dorogovtsev S N and Mendes J F F 2003 *Evolution of Networks: from Biological Nets to the Internet and WWW* (Oxford: Oxford University Press)
[40] Pastor-Satorras R and Vespignani A 2004 *Evolution and Structure of the Internet* (Cambridge: Cambridge University Press)
[41] Boccaletti S, Latora V, Moreno Y, Chavez M and Hwang D-U 2006 *Phys. Rep.* **424** 175
[42] Newman M, Barabási A-L and Watts D J 2006 *The Structure and Dynamics of Networks* (Princeton, NJ: Princeton University Press)
[43] Caldarelli G 2007 *Scale-Free Networks* (Oxford: Oxford University Press)
[44] Caldarelli G and Vespignani A (ed) 2007 *Large Scale Structure and Dynamics of Complex Networks* (Singapore: World Scientific)
[45] Manna S S and Sen P 2002 *Phys. Rev.* E **66** 066114
[46] Yook S H, Jeong H and Barabási A-L 2003 *Proc. Natl Acad. Sci. USA* **99** 13382
[47] Barrat A, Barthélémy M and Vespignani A 2005 *J. Stat. Mech.* P05003
[48] Grinstein G and Linsker R 2006 *Phys. Rev. Lett.* **97** 130201
[49] González M C, Lind P G and Herrmann H J 2006 *Phys. Rev. Lett.* **96** 088702
[50] Lind P G, Andrade J S Jr, da Silva L R and Herrmann H J 2007 *Phys. Rev.* E **76** 036117
[51] Lind P G, Andrade J S Jr, da Silva L R and Herrmann H J 2007 *Europhys. Lett.* **78** 68005

Endnotes

(1) Author: Please provide PACS number(s) in full form.
(2) Author: Please be aware that the colour figures in this article will only appear in colour in the Web version. If you require colour in the printed journal and have not previously arranged it, please contact the Production Editor now.
