Understanding individual human mobility patterns

Supplementary Material

Marta C. González, César A. Hidalgo, Albert-László Barabási

Contents

I. DATA

A. D_1 **Dataset**: This dataset was collected by a European mobile phone carrier for billing and operational purposes. It contains the date, time and coordinates of the phone tower routing the communication for each phone call and text message sent or received by 6 million costumers. The dataset summarizes 6 months of activity. To guarantee anonymity, each user is identified with a security key (hash code). Furthermore, we only know the coordinates of the tower routing the communication, hence a user's location within a tower's service area. Each tower serves an area of approximately 3 km^2. Due to tower coverage limitations driven by geographical constraints and national frontiers no jumps exceeding $\sim 1,000$ km can be observed in the dataset.

The research was performed on a random set of $100,000$ selected from those making or receiving at least one phone call or SMS during the first and last month of the study, translating to $16,364,308$ recorded positions. We removed all jumps that took users outside the continental territory. We did not impose any additional criterion regarding the calling activity to avoid possible selection biases in the mobility pattern.

B. D_2 **Dataset:** Some services provided by the mobile phone carrier, like pollen and traffic forecasts, rely on the approximate knowledge of customer's location at all times of the day. For customers that signed up for location dependent services, the date, time and the closest tower coordinates are recorded on a regular basis, independent of their phone usage. We were provided such records for $1,000$ users, among which we selected the group of users whose coordinates were recorded at every two hours during an entire week, resulting in 206 users for which we have $10,613$ recorded positions. Given that these users were selected based on their actions (signed up to the service), in principle the sample cannot be considered unbiased, but we have not detected any particular bias for this data set.

For each user in D_1 and D_2 we sorted the time resolved sequence of positions and constructed individual trajectories.

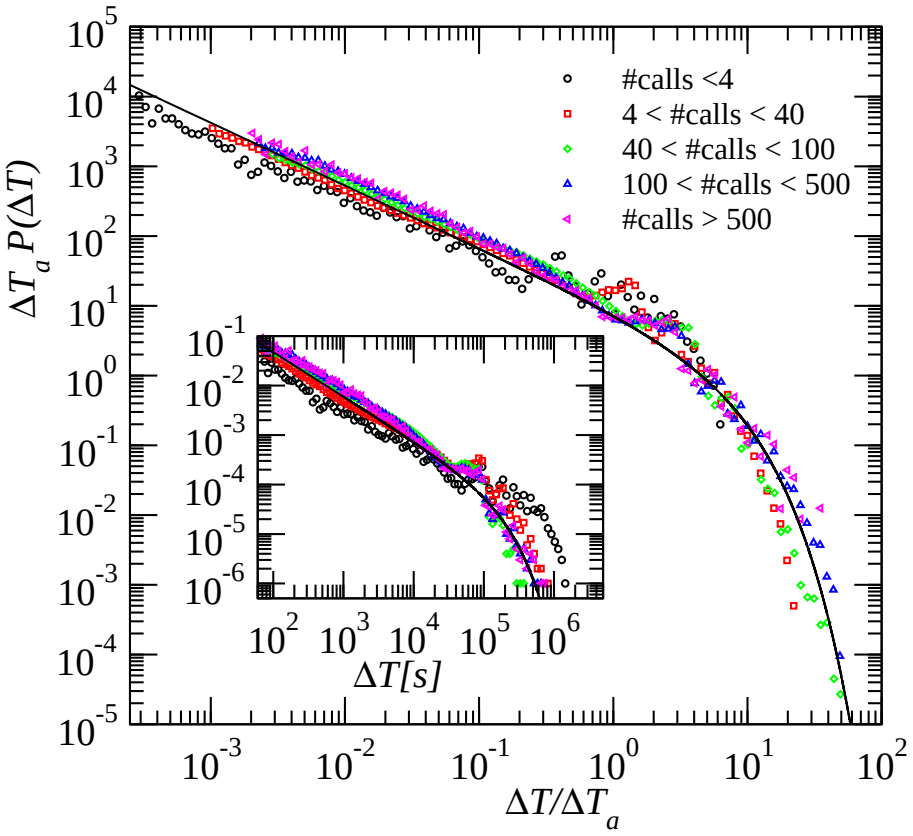

FIG. S1: Interevent time distribution $P(\Delta T)$ of calling activity. ΔT is the time elapsed between consecutive communication records (phone calls and SMS, sent or received) for the same user. Different symbols indicate the measurements done over groups of users with different activity levels (# calls). The inset shows the unscaled version of this plot. The solid line corresponds to Eq. (S1).

II. CHARACTERIZING INDIVIDUAL CALLING ACTIVITY

Communication patterns are known to be highly heterogeneous: some users rarely use the mobile phone while others make hundreds or even thousands of calls each month [1]. To characterize the dynamics of individual communication activity, we grouped users based on their total number of calls. For each user we measured the probability that the time interval between two consecutive calls is ΔT [2–4]. The inset of Fig. S1 shows that users with less activity tend to have longer waiting times between consecutive calls. By rescaling the axis with the average interevent time ΔT_a as $\Delta T_a P(\Delta T)$ and $\Delta T/\Delta T_a$ the obtained distributions collapse into a single curve (Fig. S1). Hence the measured interevent time distribution can be approximated by the expression $P(\Delta T)$

$= 1/\Delta T_a \mathcal{F}(\Delta T/\Delta T_a)$, where $\mathcal{F}(x)$ is independent of the average activity level of the population. This is a universal characteristic of the system and it agrees with earlier results on the temporal patterns of e-mail communication [5]. In addition, we find that the data in Fig. S1 is well approximated by

$$P(\Delta T) = (\Delta T)^{\alpha} \exp(\Delta T/\tau_c), \tag{S1}$$

where the power-law exponent $\alpha = 0.9 \pm 0.1$ is followed by an exponential cutoff of $\tau_c \approx 48$ days. Equation (S1) is shown by a solid line in the inset of Fig. S1 and its scaled version is presented in the main panel of Fig. S1. Here we used $\Delta T_a = 8.2$ hours, which is the average interevent time measured for the whole population. The heterogeneity in the communication pattern translates into heterogeneous sampling for the D_1 dataset. The D_2 dataset, with records at every two hours, obviously does not display this heterogeneity. Below we show that this temporal heterogeneity does not affect our results on the observed travel patterns.

III. OBSERVATIONS AT A FIXED INTEREVENT TIME

Given the widely varying distribution of the interevent times between two calls (an therefore the localization data), we need to investigate if the observed displacement statistics are affected by this sampling heterogeneity. Using the D_1 dataset, we calculated the displacement distribution $P(\Delta r)$ for consecutive calls separated by a time $\Delta T_o \pm 0.05 \Delta T_o$, where ΔT_o ranged from 20 min to one day. Figure S2 shows that for $\Delta T_o < 4$ h, the observed displacements are bounded by the maximum distance that users can travel in the ΔT_o time interval. For $\Delta T_o \geq 8$ hours we already observe $\Delta r_{max} \sim 1,000$ km, which corresponds to the largest displacement we could possible observe given the area under study (such large jumps likely are the result of airline travel). We observe that the resulting $P(\Delta r)$ distributions for different ΔT_o is again well approximated by a truncated power-law with an exponent $\beta = 1.75$. This agrees with the exponent found when we studied all consecutive calls (see Fig. 1C), suggesting that the use of consecutive calls is an accurate proxy to measure human displacement at large enough scales (> 1 km).

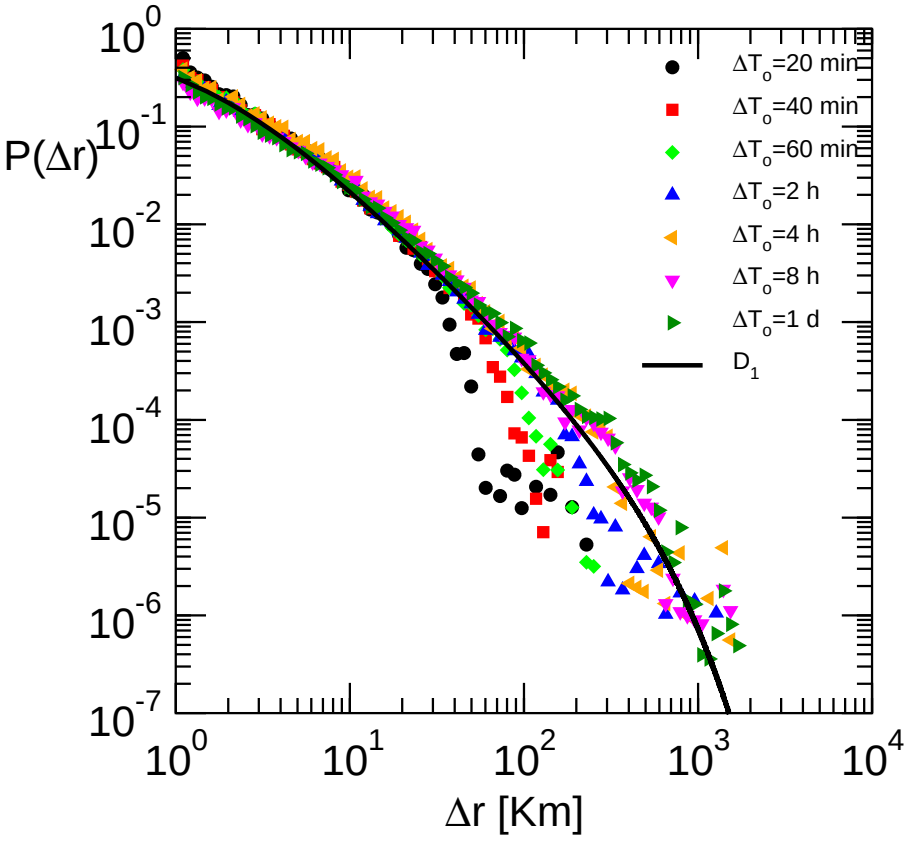

FIG. S2: Displacement distribution $P(\Delta r)$ for fixed inter event times ΔT_o based on the D_1 dataset. The cut-off of the distribution is set by the maximum distance users can travel for shorter inter event times, whereas for longer times the cutoff is given by the finite size of the studied area, as discussed in the manuscript. The black lines are from (1) reported in the manuscript, with the value $\kappa = 400$ km corresponding to D_1 (solid line).

IV. INTRINSIC REFERENCE FRAME FOR INDIVIDUAL TRAJECTORIES

A. Radius of gyration: The linear size occupied by each user's trajectory up to time t is characterized by its the radius of gyration defined as

$$r_g^a(t) = \sqrt{\frac{1}{n_c^a(t)} \sum_{i=1}^{n_c^a} (\vec{r_i^a} - \vec{r}_{cm}^a)^2}, \tag{S2}$$

where $\vec{r_i^a}$ represents the $i = 1, ..., n_c^a(t)$ positions recorded for user a and $\vec{r}_{cm}^a = 1/n_c^a(t) \sum_{i=1}^{n_c^a} \vec{r_i^a}$ is the center of mass of the trajectory.

B. Moment of Inertia: To compare different users' trajectories we need to study them in a common reference frame. Inspired by the mechanics of rigid bodies [6], we assign each user to an intrinsic reference frame calculated a posteriori from a user's trajectory. We can think of the number times a user visited a given location as the mass associated with that particular position. We denote a user's trajectory with a set of locations $\{(x_1, y_1), (x_2, y_2), ...(x_{n_c}, y_{n_c})\}$, where n_c is the number of positions available for the user. An object's moment of inertia is given by the average spread of an object's mass from a given axis. A two dimensional object can be characterized by a 2×2 matrix known as the *tensor of inertia*

$$\mathbf{I} = \begin{pmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{pmatrix}. \tag{S3}$$

We can calculate the inertia tensor for user's trajectories by using the standard physical formulas

$$I_{xx} \equiv \sum_{i=1}^{n_c} y_i^2 \tag{S4}$$

$$I_{yy} \equiv \sum_{i=1}^{n_c} x_i^2 \tag{S5}$$

$$I_{xy} = I_{yx} \equiv -\sum_{i=1}^{n_c} x_i y_i. \tag{S6}$$

Since the tensor $\mathbf{I}$ is symmetric, it is possible to find a set of coordinates in which $\mathbf{I}$ will be diagonal. These coordinates are known as the tensor's principal axes $(\hat{e}_1, \hat{e}_2)$. In this set of coordinates $\mathbf{I}$ takes the form

$$\mathbf{I_D} = \begin{pmatrix} I_1 & 0 \\ 0 & I_2 \end{pmatrix}, \tag{S7}$$

where I_1 and I_2 are the principal moments of inertia. They also correspond to the eigenvalues of $\mathbf{I}$ and can be calculated from the original set of points as

$$I_1 = \frac{1}{2}(I_{xx} + I_{yy}) - \frac{1}{2}\mu \tag{S8}$$

$$I_2 = \frac{1}{2}(I_{xx} + I_{yy}) + \frac{1}{2}\mu, \tag{S9}$$

with

$$\mu \equiv \sqrt{4\, I_{xy}\, I_{yx} + {I_{xx}}^2 - 2\, I_{xx}\, I_{yy} + {I_{yy}}^2} \tag{S10}$$

The corresponding eigenvectors determine the principal axes ($\hat{e}_1$ and $\hat{e}_2$), representing the symmetry axes of a given trajectory.

C. Rotation of user trajectories: We transform each user's principal axes $(\hat{e}_1, \hat{e}_2)$ to a common intrinsic reference frame $(\hat{e}_x, \hat{e}_y)$ calculating the angle between the axes $\hat{e}_x$ and $\hat{e}_1$, as

$$\cos(\theta) = \frac{\vec{v}_1.\hat{e}_x}{|\vec{v}_1|} \tag{S11}$$

where v_1, is the eigenvector associated with the eigenvalue I_1

$$\vec{v}_1 = \begin{bmatrix} -\frac{I_{xy}}{1/2\, I_{xx} - 1/2\, I_{yy} + 1/2\, \mu} \\ 1 \end{bmatrix}, \tag{S12}$$

resulting in

$$\cos(\theta) = -I_{xy}\,(1/2\, I_{xx} - 1/2\, I_{yy} + 1/2\, \mu)^{-1}\, \frac{1}{\sqrt{1 + \frac{{I_{xy}}^2}{(1/2\, I_{xx} - 1/2\, I_{yy} + 1/2\, \mu)^2}}}. \tag{S13}$$

After rotation by θ, we impose a conditional rotation of 180^o such that the most frequent position lays always in $x > 0$.

D. Example: Figure S3 shows the recorded trajectories of 3 users (u_1, u_2 and u_3), each characterized by a different radius of gyration: $r_g|_{u1} = 2.28$ km, $r_g|_{u2} = 29.02$ km, and $r_g|_{u3} = 313.72$ km. Using (S4), (S5) and (S6), we calculated the different components of the tensor of inertia. Equations (S12) and (S13) allow us to determine the intrinsic axes for each user $(\hat{e}_1, \hat{e}_2)$, which are displayed in Fig. S3a. Their respective angles are: $\theta|_{u1} = 127.67^o$, $\theta|_{u2} = 40.20^o$ and $\theta|_{u3} = 60.08^o$. Each set of points is rotated by $-\theta$, such that $(\hat{e}_x, \hat{e}_y)$ is the new intrinsic reference frame of each user's trajectory, as shown in Fig. S3b. The most frequent and the second most frequent positions of each user are marked as a blue and orange circle respectively. After rotating the trajectory of user 2, its most frequent position lays in $x < 0$, hence we apply an additional rotation of 180^o such that the most frequent position lays in $x > 0$. The purpose of this is to conserve the asymmetry of the user's visitation pattern. In the absence of the rotation the trajectories in Fig. S3a and B (also Fig. 3 in the manuscript) will appear to be symmetric. Given, however, that there is a significant difference in the most and the second most visited locations (see Fig. 2D in the paper), we need to perform the symmetry breaking rotation to emphasize its presence. For example, we

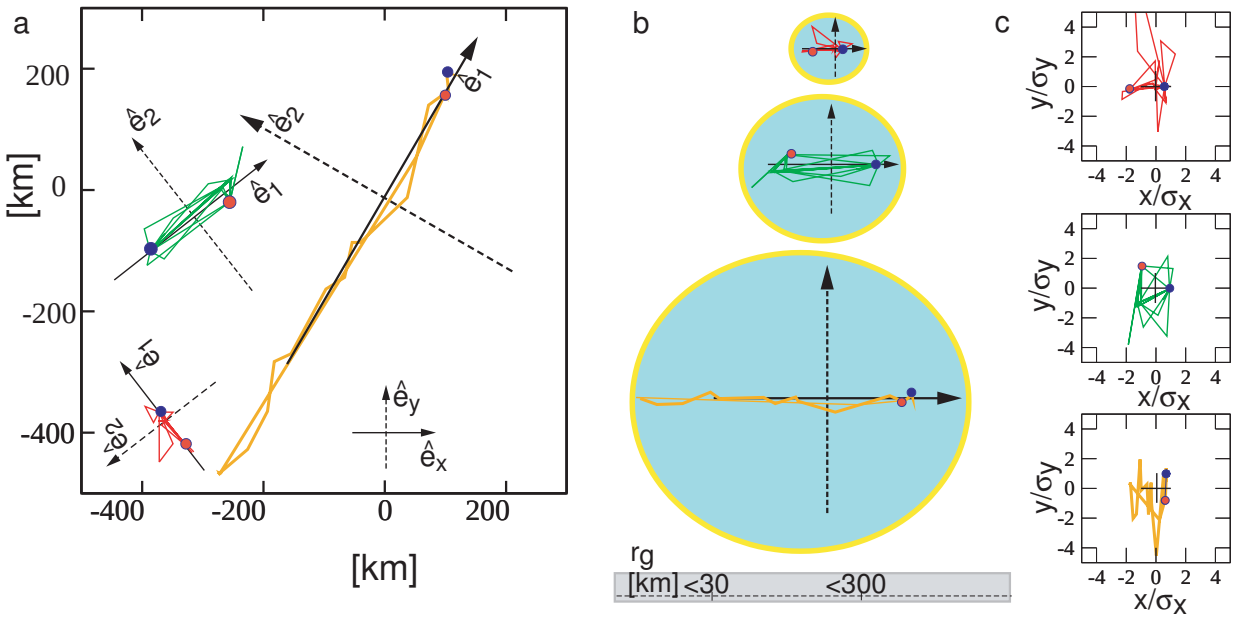

FIG. S3: Example of how to transform the user trajectories in a common reference frame. **a,** Initial trajectories of three users and their principal axes $(\hat{e}_1, \hat{e}_2)$. **b,** Each trajectory is rotated an angle $-\theta$ to align $\hat{e}_1$ with $\hat{e}_x$. An additional rotation with 180^o is required when the most frequent position (marked with a blue circle) lays in $x < 0$ after the rotation. This is the case of user 2 (green line). **c,** Positions (x, y) are scaled as $(x/\sigma_x, y/\sigma_y)$ after which the different trajectories have a quite similar shape.

found that for finite Lévy flights the rotation induced a slight but detectable anisotropy, capturing the fact that each finite trajectory has some inherent anisotropy.

We scale the trajectories on the intrinsic axes with the standard deviation of the locations for each user a

$$\sigma_x^a = \sqrt{\frac{1}{n_c^a}\sum_{i=1}^{n_c^a}(x_i^a - x_{cm}^a)^2}, \tag{S14}$$

$$\sigma_y^a = \sqrt{\frac{1}{n_c^a}\sum_{i=1}^{n_c^a}(y_i^a - y_{cm}^a)^2}. \tag{S15}$$

Note that the coordinate origin for each user is placed at the center of mass of the trajectory, $\vec{r_{cm}^a} = (0, 0)$. In this example, $\sigma_x|_{u1} = 2.24$ km, $\sigma_x|_{u2} = 28.76$ km, and $\sigma_x|_{u3} = 313.60$ km whereas $\sigma_y|_{u1} = 0.43$ km, $\sigma_y|_{u2} = 3.88$ km, and $\sigma_y|_{u3} = 8.49$ km. After scaling, the shapes of the three trajectories look similar (S3c), despite that we are showing users with significantly different mobility patterns and ranges. This is the underlying procedure that allows us to obtain

the universal density function $\tilde{\Phi}(x/\sigma_x, y/\sigma_y)$.

E. Spatial density function: For agent based modeling it is crucial to know the probability that an individual can be found at a position (x, y) during the day. As our results show, knowledge of the spatial density function $\tilde{\Phi}(x/\sigma_x, y/\sigma_y)$ represent the first step towards such a modeling effort. Indeed, using the density function $\tilde{\Phi}(x/\sigma_x, y/\sigma_y)$ for an ensemble of agents with r_g's following Eq.(3), each agent's position can be rescaled using Eq.(4) and the fact that $\sigma_x = 0.94 r_g^{0.97}$. The distribution of individuals in space can be arbitrary or more realistic if taken from census information. The three matrixes shown in Fig. 3B can be downloaded from: http://www.nd.edu/ mgonza16/DensityFunction/

V. SCALING RELATION BETWEEN EXPONENTS

Next we show that there is a consistent relationship among the different exponents describing the travel patterns of the population. The exponent β characterizing the distances traveled by the entire population is related to α, which characterizes distances traveled by individuals and β_r, that captures the distribution of the radius of gyration. We note that (1) should be the result of a convolution between (3) and $P(\Delta r|r_g)$, hence

$$P(\Delta r) = \int_0^\infty P(\Delta r|r_g) P(r_g) dr_g, \tag{S16}$$

using the expressions introduced in the manuscript this equation can be expanded as

$$P(\Delta r) = \int_0^\infty r_g^{-\alpha} F(\frac{\Delta r}{r_g})(r_g + r_g^0) e^{-r_g/\kappa} dr_g. \tag{S17}$$

Focusing on the asymptotic scaling behavior we drop the short length cutoff r_g^0 and extract the leading term by performing the substitution $r_g = \Delta r x$. Finally the scaling is given by

$$P(\Delta r) \approx \Delta r^{-\alpha-\beta_r+1} \int_0^\infty x^{-\alpha} F(\frac{1}{x}) e^{x\Delta r/\kappa} dx, \tag{S18}$$

indicating that $\beta = \alpha + \beta_r - 1$. Note, however, that the integral in (S18) also depends on Δr, therefore the scaling relationship is valid only to the leading order and further corrections may result from the integral. This correction cannot be evaluated in the absence of an analytical approximation for $F(x)$. For our data we find $\beta = 1.75 \pm 0.15$, $\beta_r = 1.65 \pm 0.15$ and $\alpha = 1.2 \pm 0.1$, indicating that the scaling relation, within error bars, is satisfied, and that there is a systematic difference between β and β_r of magnitude $\alpha - 1$.

VI. TIME DEPENDENCE OF THE RADIUS OF GYRATION

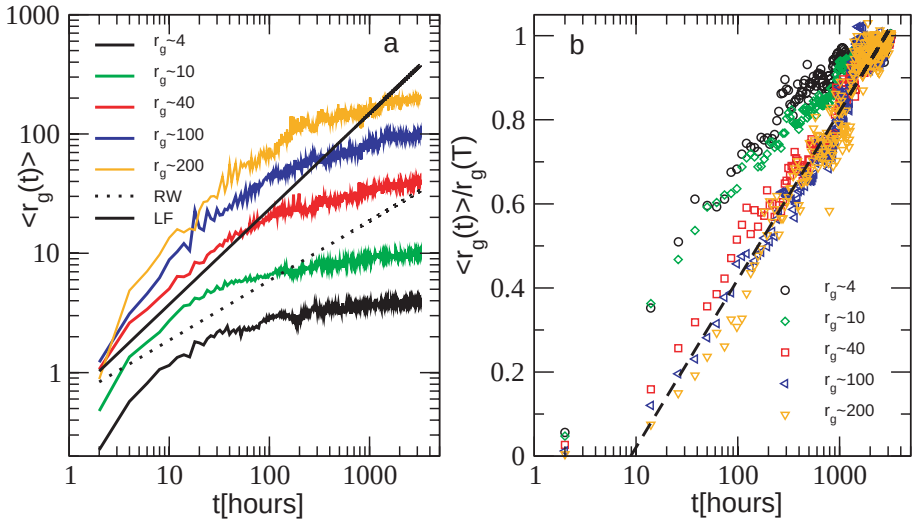

FIG. S4: Time evolution of the radius of gyration $\langle r_g(t)\rangle$ vs. time, for various groups of users with different asymptotic $r_g(T)$ after $T = 6$ months. **a,** In a log-log scale the black lines correspond to the powers of time for the random walk and Lévy flight models, which are in contrast with the time dependence of $\langle r_g(t)\rangle$ measured for the mobile phone users. **b,** In a log-linear scale, note that the $r_g = 4$ and $r_g = 10$ visibly deviate from the large r_g curves. This is not surprising, as for these two curves the recorded distances are comparable to the average tower distances (both curves appear to saturate under 10 km, while the average area of reception for a tower is about 3 km^2). Thus, small travel distances are overestimated due to the measurement resolution. Curves with $r_g > 10$ km are less affected by the tower resolution and all these appear to collapse in the same behavior once rescaled with $r_g(T)$, and they are all approximated with a logarithmic time dependence. The straight line is not a fit, but it is shown only as a guide to the eye.

Figure 2A in the manuscript shows three groups of users chosen according to the asymptotic $r_g(T)$ after $T = 6$ months. In Fig. S4 we show that the same time dependence is observed for a more strictly selected grouping of the users, choosing five different groups of users with very similar asymptotic radius of gyration: $r_g(T) \pm 0.05 r_g(T)$. Given the high daily and weekly-based

fluctuations in the phone usage patterns, we averaged $\langle r_g(t) \rangle$ over 168 different initial conditions, *i.e.* we started the measurements at every 6 hours during one week. This averaging not only removed the dependence on the initial conditions, but also significantly reduced the noise in the curves.

The log-log scale in Fig. S4a allows to see in detail the early behavior of the curves, indicating that a power law does not offer a good fit to the data. As we show in the log-linear plot in Fig. S4b, we find that the radius of gyration increases logarithmically in time, which is in strong contrast with the power law dependence expected for Lévy flights ($r_g(t)|_{LF/TLF} \sim t^{3/(2+\beta)}$) and random walks ($r_g(t)|_{RW} \sim t^{1/2}$). This indicates that the average radius of gyration of mobile phone users has a manifestly slower dependence than the predicted power laws, a behavior that may appear similar to a saturation process. Note that the $r_g \sim 4, 10$ curves appear to deviate from the logarithmic behavior. We belive that this is due to the spatial resolution offered by the tower density: we cannot reliably and systematically resolve jumps in the vicinity of a few kms, given that we record a motion only when a person moves between towers, that are often a few kilometers apart. The $r_g > 10$ km curves, given the distances involved, are not affected by the granularity of the data collection process, and they all follow the logarithmic behavior.

VII. STATISTICAL TESTS OF FITTING DISTRIBUTIONS

Given the fat tailed distributions observed for human travel patterns, it is important to see if the data is statistically consistent with the best fits. The purpose of this section is to support our findings with rigorous statistical tests. In the past year there has been significant attention devoted to the question of how to statistically measure the goodness of fit for a power law [7–10]. This was prompted partly by the need to quantify the validity of the Lévy flight finding in animal travel patterns. Note, however, that there is a significant difference between the data quality available in the animal and human travel patterns. Indeed, the mammalian data was available for short time periods for only a few animals, providing only a small number of observed individual displacements. Given the scarcity of data, precise statistical tools are needed to extract the proper fit. In contrast, the data analyzed in Ref. [11] as well as in this paper contain millions of displacements. Thus we are in a regime where typically traditional statistical tools, designed to deal with limited information, are less crucial. Yet, appropriate statistical tools can be used to explore the goodness

of the fit.

In this respect, it is often believed that statistical methods can validate a particular fit. The truth is, as emphasized in a recent publication [10], that these tools can only tell if a particular fit is consistent with the data, and rule some fits out, rather than validate a particular fit. A second important observation is that, given the high interest devoted to power laws, recently the issue of fitting a power law has been addressed in detail, developing the proper statistical tools to address the goodness of the fit [7–10]. The same tools are not available for truncated power laws, however, thus limiting the available methods to address their statistical relevance. In general we find that all the fits that we used in the paper pass the Kolmogorov-Smirnov test for the goodness of fit (Sect. V.A) and that a power law offers a much better approximation overall than an exponential function (Sect. V.B). Note that given the vast amount of data and the really good fit offered by the truncated power law, this last conclusion is hardly surprising.

A. Kolmogorov-Smirnov goodness of fit test

We tested whether the empirical data could come from the fitted distributions by performing a stringent variant of the Kolmogorov-Smirnov (KS) goodness of fit test [10]. The KS statistics is a simple way to compare whether two distributions are the same. In this case, we use it to test the hypothesis: *Could the empirically observed distributions come from the distribution found as its best fit*. For this we generated synthetic data starting from the fitted distribution and then use the KS test to see whether the empirical data we have behaves as well as the synthetic data generated from the fitted distribution.

We use two variants of the KS statistics to compare empirical data with the fitted distribution and synthetic data with the fitted distribution. The first method is the standard KS statistics and is given by:

$$KS = \max\left(|F - P|\right) \tag{S19}$$

where F is the cumulative distribution of the best fit and P is the cumulative distribution of the empirical or synthetic data. The regular KS statistic is not very sensitive on the edges of the cumulative distribution. Hence, we also used the weighted KS statistics defined as:

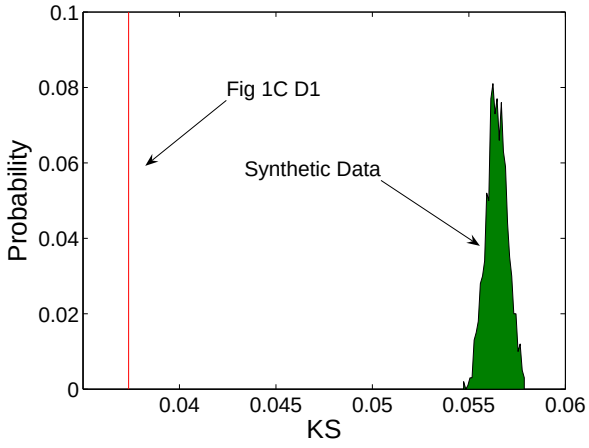

FIG. S5: KS test for Fig 1C D_1

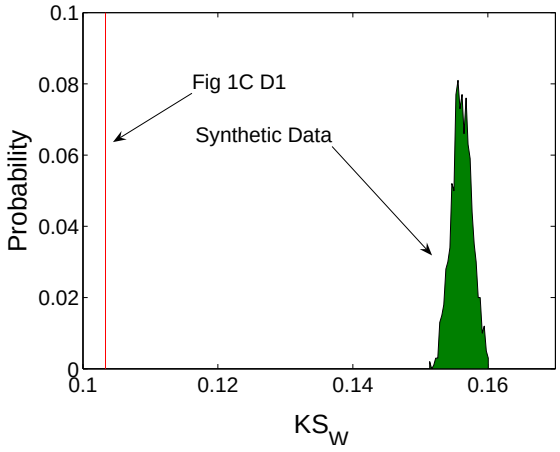

FIG. S6: KS_W test for Fig 1C D_1

$$KS_W = \max \frac{|F - P|}{\sqrt{P(1-P)}} \tag{S20}$$

To test whether the empirical data behaves as good as the synthetic data we calculated the KS and KS_W statistics between the empirical data and its best fit and compared these values with those obtained by calculating KS and KS_W for $1,000$ synthetic data sets generated from the best fit. If the values obtained for KS and KS_W for the empirical data behave as good or better than those obtained for the synthetic data, then we can conclude that the empirical data is statistically consistent with its best fit. The results of the KS test can be summarized using a $p - value$

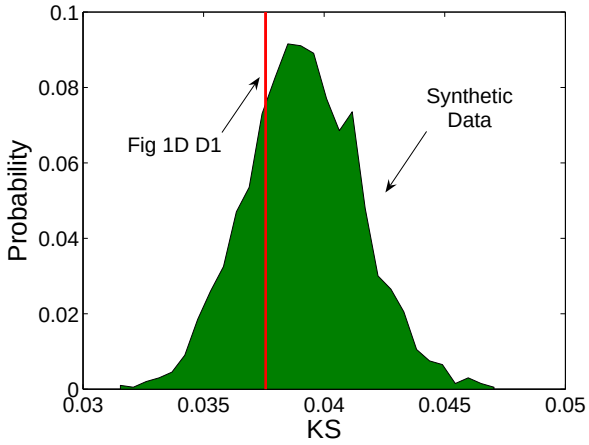

FIG. S7: KS test for Fig 1D D_1

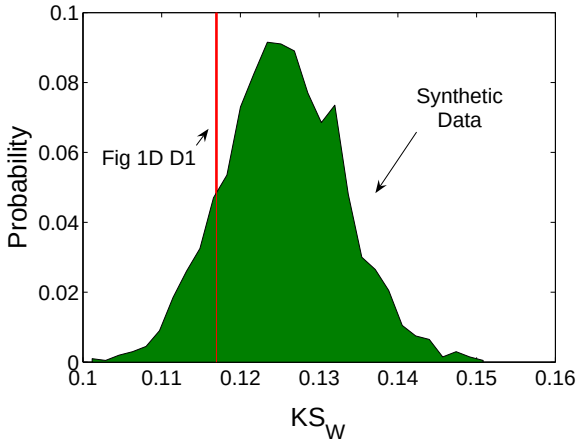

FIG. S8: KS_W test for Fig 1D D_1

by integrating the distribution of KS values generated with the synthetic data from the value representing the empirical distribution. When integrating such distributions from left to right we can interpret the $p-value$ as the probability that the observed data was the result of its best fit. A $p-value$ close to 1 will indicate that the empirical distribution matches its best fit as good as synthetic data generated from the fit itself [10], whereas a relative small $p-value$ (typically taken $p<0.01$) would suggest that the empirical distribution can not be the result of its best fit.

Passing the KS test does not rule out the possibility that the empirical data could be fitted as well or even better with some other function. In such a case, given the size of our samples, we

believe that such an exercise would be technical rather than practical and that different functional forms will closely resemble each other on the range were the fit was made once the fitting parameters have been fixed.

1. KS statistics for Fig 1C, D_1: Figure S5 compares the KS values obtained for the empirical data presented in Fig 1C, D_1 of the paper with those obtained for $1,000$ distributions of synthetic data generated to comply with Eq. (1). Figure S6 shows the same for KS_W. In both cases we find that the empirical data passes the KS test, in fact behaving better than the synthetic data. Indeed $p(KS) = 1$ and $p(KS_W) = 1$.

2. KS statistics for Fig 1D, D_1: Figure S7 compares the KS values obtained for the empirical data presented in Fig 1D, D_1 of the paper with those obtained for 2000 distributions of synthetic data generated to comply with Eq. (3). Figure S8 shows the same for KS_W. In both cases we find that the empirical data passes the KS test, behaving as some of the best examples of the synthetic data, obtaining $p(KS) = 0.62$ and $p(KS_W) = 0.82$.

B. Maximum Likelihood Estimates: Comparing power-laws and exponentials

The Maximum Likelihood Method is a powerful way of estimating the fitting parameters best describing a empirical distribution. The method can also be used to compare the relative likelihood of two fits. In this section we are interested in testing whether a broad distribution, such as a power-law, is a better fit than an exponential for many of the distributions presented on the paper. It is not our intention to claim that the distributions presented here are in fact power-laws but to build suggestive evidence testing if the data presented in the paper is better fitted by a broad distribution, such as the power-law, rather than a narrow distribution such as an exponential. In fact, as we discuss in the manuscript and the previous section, we believe that the best fit to our data is given by a truncated power law rather than a simple power law. We do not present here the maximum likehood estimate of the truncated power-law fit because being its cumulative a Whittaker function, it is not readily to be handled with this method. For those interested in the accuracy of the best fit we suggest reading section VII A of this supplementary material.

The details of the maximum likelihood method have been widely published. Those interested in performing such fits could find help in [10] and [7].

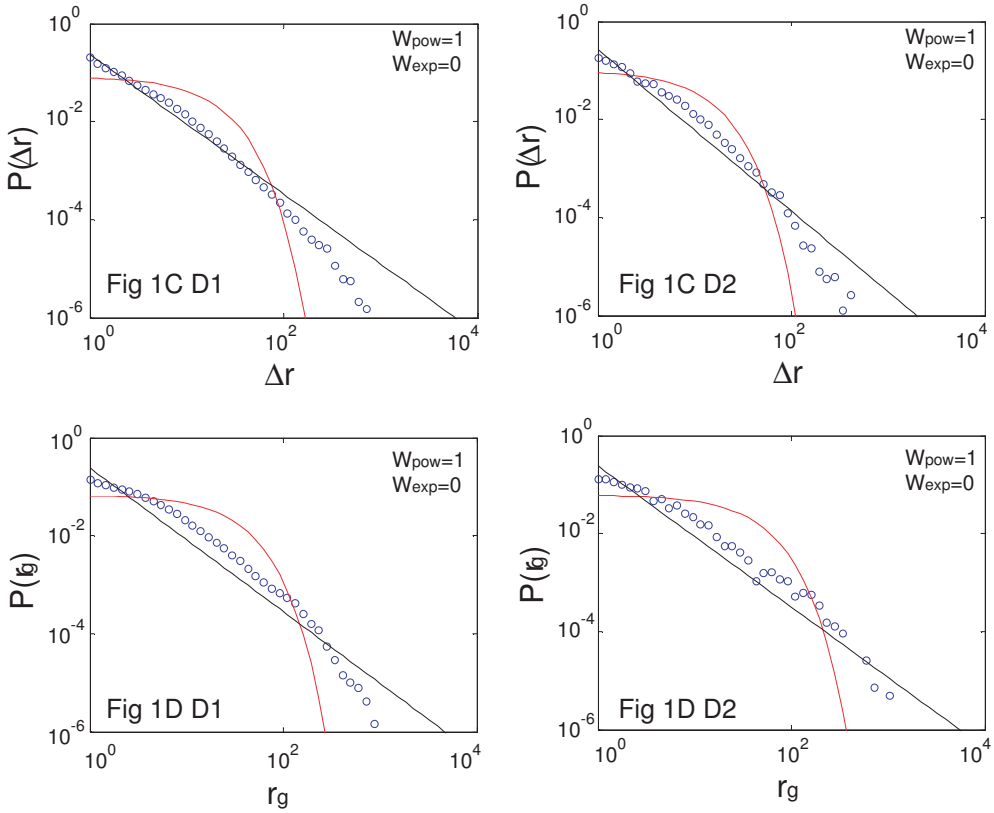

FIG. S9:

Testing for broad and narrow distributions: The relative likelihood between two distributions can be calculated from its relative weights

$$W_i = \frac{e^{-\Delta_i}}{e^{-\Delta_1} + e^{-\Delta_2}}, \tag{S21}$$

where $\Delta_i = \min\left(AIC_i - AIC_{min}\right)$. Here AIC is the Akaike information criteria associated with the power-law ($i = 1$) and exponential ($i = 2$) fits. The AIC can be expressed as a function of the log-likelihood [7]

$$AIC_i = 2\log\max(L_i) + 2K_i$$

where K_i is the number of parameters used in the fit and L_i is the likehood of a particular value of a fitting parameter.

Here we test two fits, a power-law $Ax^{-\beta}$ and an exponential $Ae^{-\mu x}$, where A is a normalization constant. In Fig. S9 we present the results of the Maximum Likelihood Fits for the distributions introduced in Fig. 1C and Fig. 1D of the paper. Note, we use the data log binned data, which has been recommended as the most accurate plotting method for this kind of analysis [8]. In each case we find $W_{pow} >> W_{exp}$ indicating that a power law is a more likely fit than an exponential, indicating that the distribution of displacements, as well as r_g, are better approximated by a broad rather than a narrow distribution.

---

[1] Onnela, J.-P., Saramäki, J., Hyvönen, J., Szabó, G., Lazer, D., Kaski, K., Kertész, K. & Barabási A.L. Structure and tie strengths in mobile communication networks. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 7332-7336 (2007).

[2] Barabási, A.L. The origin of bursts and heavy tails in human dynamics. *Nature* **435**, 207-211 (2005).

[3] Vázquez, A., Oliveira, J.G., Dezsö, Z., Goh,K.-I, Kondor, I. & Barabási, A.-L. Modeling bursts and heavy tails in human dynamics. *Physical Review E* **73**, 036127 (2006).

[4] Hidalgo, C. & Barabási, A.L. Inter-event time of uncorrelated and seasonal systems. *Physica A* **369**, 877-883 (2007).

[5] Goh, K.-I. & Barabási, A.L. Burstiness and memory in complex systems. *Eurphysics Letters* **81** 48002 (2008).

[6] Goldstein, H. *Classical Mechanics*. (Addison-Wesley, 1959).

[7] Edwards, A.M., Phillips, R.A., Watkins,N.W., Freeman, M.P., Murphy, E.J., Afanasyev, V., Buldyrev, S.V., da Luz, M.G.E., Raposo, E. P., Stanley, H. E., & Viswanathan, G. M. Revisiting Lvy flight search patterns of wandering albatrosses, bumblebees and deer. *Nature* **449**, 1044-1049 (2007).

[8] Sims D.W., Righton D. & Pitchford, J.W. Minimizing errors in identifying Lévy flight behaviour of organisms. *Journal of Animal Ecology* **76**, 222-229 (2007).

[9] Pueyo, S. & Jovani, R. Comment on A keystone mutualism drives pattern in a power function. *Science* **313**, 1739c (2006).

[10] Clauset, A., Rohila Shalizi, C. & Newman, M.E.J. Power-law distributions in empirical data. *arXiv:physics:/07061062*.

[11] Brockmann, D., Hufnagel, L. & Geisel, T. The scaling laws of human travel. *Nature* **439**, 462-465

(2006).