The Effect of Recency to Human Mobility

Hugo Barbosa¹, Fernando Buarque de Lima Neto², Alexandre Evsukoff³, and Ronaldo Menezes¹

¹BioComplex Lab, Department of Computer Sciences & Cybersecurity, Florida Institute of Technology, USA
²Computational Intelligence Research Group, Polytechnic School, University of Pernambuco, Brazil
³COPPE, Federal University of Rio de Janeiro, Brazil
*hbarbosa@biocomplex.org

April 8, 2015

Abstract

A better understanding of how people move in space is of fundamental importance for many areas such as prevention of epidemics, urban planning and wireless network engineering, to name a few. In this work we explore a rank-based approach for the characterization of human trajectories and unveil a visitation bias toward recently visited locations. We test our hypothesis against different empirical data of human mobility. Also, we propose an extension to the Preferential Return mechanism to incorporate the new recency-based mechanism.

Introduction

A better understanding on the fundamental mechanisms of human mobility is of importance for many research fields such as epidemic modeling [1, 2, 3], urban planning [4, 5], and traffic engineering [6, 7, 8]. Although individual human trajectories can seem unpredictable and intricate for an external observer, in fact, human trajectories are very predictable [9, 10, 11, 12, 13, 14] and regular over space and time [15, 16, 17]. One characteristic of human motion, largely observed in empirical data, is the fact that we have the tendency to spend most of our time in just a few locations [15, 18, 19]. More generally, visitations frequencies distribution have been observed to be heavy tailed [18, 19].

However, the fundamental mechanisms responsible for shaping our visitation preferences are not fully understood. The preferential return (PR) mechanism, proposed by Song et al. [18], offered an elegant and robust model for the visitation frequency distribution. More precisely, it defines the probability $\Pi_i$ for
returning to a location \( i \) as \( \Pi_i \propto f_i \), where \( f_i \) is the visitation frequency of the location \( i \). It implies that the more visits a location receives, the more visits it is going to receive in the future, which in different fields goes by the names of Matthew effect \[20\], cumulative advantage \[21\], or preferential attachment \[22\].

Although the focus of the PR mechanism–as part of the Exploration and Preferential Return (EPR) individual mobility model–was to reproduce some of the scaling properties of human mobility, its general principles are grounded on plausible assumptions from the human behavior point of view. However, in the long term, the PR assumption as a property of human motion leads to two discrepancies. First, in the model, the earlier a location is discovered, the more visits it is going to receive. It implies that first visited location will most likely also be the most visited one. Second, if the cumulative advantage indeed holds true for human movements, people would not change their preferences, which is clearly not true.

In this work, we explore the visitation return patterns under a temporal perspective. We analyzed different ranking approaches and tested their respective correlations with the return probabilities. Our approach is based on the empirical evidences that the longer the time since the last visit to a location, the lower is the probability of observing a user at this location \[18, 15\]. The proposed approach aims at overcoming the limitations of the PR mechanism.

**Results**

**Data**

In this work, we used two mobility datasets: the first one \((D1)\) corresponds to 6 months of anonymized mobile-phone traces from a large metropolitan area in Brazil. This dataset is composed of 8,898,108 records from 30,000 users between January 1–June 30, 2014. The second dataset \((D2)\) is composed of 23,736,435 check-ins from 51,406 Brightkite users in 772,966 different locations. Unlike the mobile phone datasets, the Brightkite data has a spatial resolution in the range of a few meters. Given our interest here is on the individuals’ trajectories, in this analysis we considered only the data that give us information relating to the users’ displacement. Hence, we filtered out repeated observations in one place, resulting in a time series for each individual representing their trajectories over the observed period.

**Heterogeneities in human mobility**

The first analysis we performed was to measure the population-level heterogeneities represented by the different activity patterns. First we measured the number of observed displacements \((N)\) per user during the period. Notice that it does not necessarily represent the actual number of displacements, but rather
the number of jumps per user captured by the datasets. All the scaling parameters were estimated using the methods described by Clauset et al [24].

The $p(N)$ of $D1$ and $D2$ are better approximated by truncated power-law distributions, defined as $p(x) = C x^{-\alpha} e^{-x/\tau}$ with $\alpha_{D1} \approx 1.000$ and $\tau_{D1} \approx 783$ observations whereas $\alpha_{D2} \approx 1.3$ and $\tau_{D2} \approx 923$ observations. (see Figure 1). This means that in both datasets, users tend to not move a lot, and highly mobile individuals are very rare. For instance, in $D1$, the daily average number of displacements is approximately 2.2 whereas in $D2$ it was approximately 1.7. The average number of jumps per month in $D1$ is 24.5 while in $D2$ it was 9.2. The lower average number of movements in $D2$ could be because Brightkite was a location-based social networking service, hence, movements related to social activities must be overrepresented in it. Nevertheless, given our focus is on individuals visitation preferences–rather than needs–it does not affect negatively our analysis.

Figure 1: The Number of observed displacements per user. The probability density function of the number of observed displacements during the observational period.

From the human mobility perspective, we extracted the number of distinct locations users have visited in the period (Figure 2). It depicts the probability $p(S)$ of a user having visited $S$ distinct locations at the end of the observational period. Both the curves are better approximated by truncated power laws with exponents $\alpha_{D1} \approx 1.00$, $\tau_{D1} \approx 52.63$ whereas $D2$ has exponents $\alpha_{D2} \approx 1.22$ and $\tau_{D2} \approx 200.0$. When we look at the CCDF in linear scale (inset of Figure 2) it becomes even more evident the fact that we spend most of our time in a very few locations. To illustrate, about 30% of the time, users in $D1$ were found at just 2 locations while in $D2$ this number was approximately 40%.
Figure 2: **Number of distinct visited locations.** The probability density function of the number of unique visited locations aggregated by users. Solid lines correspond to the truncated power law fits. The inset is the CCDF of the distribution in linear scale, illustrating the fact that people tend to concentrate most of their visits to just a few locations.

**Temporal patterns**

In a modern society, where most of the people have daily routines, part of our trajectories are constrained to a limited number of locations at regular time intervals. Human activity routines are responsible for part of the regularities human movements show. From the empirical data, we extracted the time interval (in hours) between two consecutive visits to a location. The distribution of time intervals is depicted in Figure 3. The plot unveils two important features of human movements: first, one can observe existence of peaks in 24h intervals representing the users’ daily routines; second, we can observe that the probability of returning to a location decreases with \( p(\Delta t) \propto \Delta t^{-\beta} e^{-\Delta t/\kappa} \) with \( \beta_{D1} \approx 1.429 \) and \( \kappa_{D1} \approx 2,347 \) hours and \( \beta_{D2} \approx 1.442 \) with \( \kappa_{D2} \approx 7,240 \) hours. Indeed a very rapid decay.

**A rank-based analysis of human visitation patterns**

As previously described, the PR mechanism suggests that the visitation probability of a particular location is proportional to the number of previous visits to it. Our claim is that the Zipf’s Law observed in visitation frequencies distribution is influenced by our tendency to return to recently visited locations.

To test such influence we compared the return probabilities from two ranking approaches: one based on the visitation frequencies \( (K_f) \), and the other \( (K_s) \)
Figure 3: Return probabilities as a function of the elapsed time $\Delta t$ since the most recent visit. Peaks are observed at 24h intervals, capturing the temporal regularity of which humans return to previously visited locations. Also, it is possible to see that the return probability decays very quickly as the time increases.

Based on the recency of the last visit to a location. Additionally, we performed the same analyses presented in here to a set of randomized variations of the datasets (see Supplementary Information).

In summary, the two ranks can be describe as:

- $K_s$ - recency-based rank. A location with $K_s = 1$ at time $t$ means that it was the previous visited location. $K_s = 2$ means that such location was the second most recent location visited up to time $t$ and so on, so forth.

- $K_f$ - is the frequency-based rank. A location with $K_f = 1$ at time $t$ means that it was the most visited location up to that point in time.

First we analyze the probability of return as a function of $K_s$. This analysis shows that such probability decays vary rapidly with $K_s$. More precisely, the probability $p(K_s)$ follows a truncated power law, similarly to $p(\Delta t)$. In fact, when we compare distribution exponents $\beta$ and $\alpha$ we see that the return probability decreases faster with $K_s$ than with $\Delta t$ for both datasets (see Figure 4).

**Recency over frequency - the role of recent events in human mobility**

In this section we explore the two-dimensional density distribution of returns $p(K_f, K_s)$. The idea is to investigate the return probabilities as an outcome of the convolution between visitation frequencies and times, encoded in $K_f$ and $K_s$ simultaneously. If users have a stronger preference for recently visited locations we should observe:

1. lower values of $K_s$ must be frequently observed over a wider range of $K_f$. It
Figure 4: Comparison between the probability of return by recency and frequency ranks. The distributions of both ranks can be better approximated by truncated power laws (fit lines dashed). a, The recency-based rank of $D_1$ has exponents $\alpha_{K_s} \approx 1.644$ and exponential cut-off $\tau_{K_s} \approx 41.66$, whereas the frequency-based rank distribution has a better fit for $\alpha_{K_f} \approx 1.86$ with $\tau_{K_f} \approx 37$. b, The best fit for the return ranks distribution in $D_2$ is achieved with parameters $\alpha_{K_s} \approx 1.699$ and $\tau_{K_s} \approx 250$ for the recency rank whereas the frequency rank has parameters $\alpha_{K_f} \approx 1.625$ and $\tau_{K_f} \approx 125$

would suggest that we tend to return to recently visited locations even if we have not visited such location many times before (i.e. lower $K_f$ rank);

2. higher values of $K_f$ must deviate from lower $K_f$ values, suggesting that the probability of return to a location decays with time, even if it was a highly visited location.

To test these hypotheses, we analyzed the frequency of returns with ranks $(K_f, K_s)$ for all $K_f$ and $K_s$. For example, a visit to a location with ranks $(10, 3)$ means a return to the 10th most visited site after visiting three other locations. This return distribution is represented as a two-dimensional histogram for each of the datasets (Fig. 5). From the heatmaps, we can observe that returns to the most visited locations (e.g., $K_f \leq 7$) have shorter return trajectories. In other words, when it comes to our most visited locations, we tend to return to them after visiting very few locations. It can be seen by the rapid decrease in the returns frequencies when $K_s$ grows. For instance, in $D_1$, more than 86% of the returns to the most visited location occurred after visiting less than five other locations while for $D_2$, it was more than 91% (see Figure 5).

We can observe also that the recency increases the probability of return to less visited locations (e.g., $7 \leq K_f \leq 40$), expressed by a broader distribution of $K_f$ when $K_s$ is low (e.g., $K_s \leq 3$. For instance, a closer look at the bottom rows of the plots in Figure 5 shows that a recent visit to a location can increase the probability of returning to it up to 10 times in $D_2$ (see Figure 5).

When we compare $D_1$ and $D_2$ we can observe a slightly different pattern between them. First, the effect of recency is much stronger in $D_2$ than in $D_1$. 
Such difference can be rooted on the fact that the mobility data of $D_1$ is coarse-grained to a cell tower level. $D_2$, on the other hand, provides finer-grained mobility data, capturing changes in visitation preferences, even when the locations are in the same cell area. Further analyses based on randomized versions of the datasets have shown that indeed the recency effect can be observed in both $D_1$ and $D_2$ (see Supplementary Information).

Figure 5: **Two-dimensional return frequencies.** Each point represents a return step, whereas the color encodes the density of points. The ranks here were shifted to have the highest-ranked locations at $(0, 0)$. A point $(x, y)$ in the histogram represents a return to the $(x + 1)^{th}$ most visited location after $y + 1$ steps. a, looking at the return ranks distribution for $D_1$ we can observe that the recency influence is less pronounced in $D_1$ in comparison with $D_2$. b, On the other hand, the fine-grained data of $D_2$ a a strong influence of the recency.

**The Recency-based model**

An alternative explanation for the anomalies observed in the ranks distribution would be that they could be simply a byproduct or artifact of the rank-based approach. To test to what extent the patterns we observed in the rank distribution corresponds to an unforeseen mechanism of human mobility, we tested for the hypothesis that it emerges from the data when we build the sequence-based ranks of frequency-driven trajectories. If the last one indeed holds true, the same patterns must be observed in the synthetic data produced by the EPR model. To test our hypothesis, we compared the purely frequentist mechanism of the EPR against our new human mobility model where returns have a bias toward recently visited locations.

The recency-based model extends the Preferential Return mechanism endowing it with a mechanism capable of capturing the visitation bias towards recently visited locations. Besides that, all other ingredients of the EPR model were kept intact except for the temporal dimension. The reason for that is because the waiting-time distribution of the EPR model determines only when an
Figure 6: **Fraction of returns to the $K_f$ most visited location occurring after the visitation of $L$ different locations.** Another way to look at the recency effect is by analyzing the correlation between the number of different visited locations between two visits to a location. We can see that people tend to return to their most visited locations after visiting very few places. **a,** In $D1$, more than 86% of the returns to the most visited location occurred after visiting less than five other locations while for $D2$ (**b**), it was more than 91%.

individual is going to move (i.e., how much time it will wait still before the next jump) but not where to go. It is important to emphasize that the recency bias underlying our model is regarding the visitation path and is time-independent.

The model can be described as follows: first, a population of $N$ agents is initialized and scattered randomly over a discrete lattice with $70 \times 70$ cells, each one representing a possible location. The initial position of each agent is accounted as its first visit. At each time step agents can either visit a new location if probability $p_{\text{new}} = \rho S^\gamma$ where $\rho = 0.6$ and $\gamma = 0.6$ are control parameters–whose values were derived by Song et al from empirical data– and $S$ corresponds to the number of distinct locations visited thus far. With complementary probability $1 - p_{\text{new}}$ an agent return to a previously visited location. If the movement is selected to be a return, with probability $1 - \alpha$ the $i^{\text{th}}$ last visited location is selected from a Zipf’s law with probability $p(i) \propto k_s(i)^{-\eta}$ where $k_s(i)$ is the recency-based rank of the location $i$. The parameter $\eta$ controls the number of previously visited locations a user would remember when deciding to visit a location. With probability $\alpha$ the destination is selected based on the visitation frequencies with probability $\Pi_i \propto k_f(i)^{-1-\gamma}$ where $k_f(i)$ is the frequency rank of location $i$. Notice that when $\alpha = 1$ we recover the original preferential return behavior of the EPR model while when $\alpha = 0$, visitation returns will be based solely on the recency. We experimentally tested different parameters configuration for the model. Our analyses have shown that when $\alpha = 0$, the heavy tail of the visitation frequency disappears while for $\alpha = 1$ the power law of the recency distribution vanishes. It suggests that both mechanisms must be present in order to reproduce those two features. In practice different individuals could have different $\alpha$ values. However, extracting it from the empirical data is not
an easy task once it is hard to determine either the movement was driven by the recency or frequency. Nevertheless, we determined that $\alpha = 0.1$ (i.e., 10% of the movements influenced by the visitation frequencies) was enough to restore the recency and frequency ranks distributions. Also, for the Zipf’s Law distribution of the recency rank we used $\eta = 1.6$, extracted from the empirical data (see Supplementary Information for the parameter estimation process).

Visually, the synthetic data produced by the EPR model seems to have a good approximation with the empirical data (see Figure 8). However, when we compare the bottom-most rows of the histogram, it deviates from the empirical evidence, by not capturing the broader distribution of $p(k_f, k_s)$ for recently visited locations. On the other hand, the recency-based mechanism (RM) reproduced the recency influence as observed in the empirical data (Figure 8b). When we look at each variable individually we notice that the $K_s$ distribution as produced by the EPR model deviates from a power law, being better approximated by an exponential distribution. When we look at the $K_f$ distribution, the EPR model recovers its heavy tail, as expected (Figure 8d).

When we look at each individual’s trajectories over, let us say, one year, the visitation patterns and regularities become very evident and radical changes in visitation patterns—such as during a long vacation abroad or after starting a new job in another city—are very unlikely. In a large population, these events indeed occur, but their effect on the population scale are very diluted and, sometimes, transient. Within such limited time window, individuals indeed are predictable, and believing that one is going to be at one of its most visited locations is a reasonable guess. However, it is really unlikely that the individual’s preferences are the same for 10 or 20 years. A recently discovered restaurant is a more plausible destination than our former workplace. Some events in our lives have the potential to reshape not only our visitation patterns but also our preferences.
Figure 8: Model results–Comparison between the EPR model and the recency-based model. 

**a,** The analysis of the return ranks generated by the EPR model shows that it reproduces a pattern similar to the one observed from the empirical analysis, especially of $D_1$. 

**b,** On the other hand, on the presence of the recency mechanism, we can observe the same high probability of return to recently visited locations (i.e., low $K_s$) as observed on the empirical data.

**c,** When we look at the distribution of the frequency ranks, the Preferential Return mechanisms (red diamonds) successfully exhibited a power-law distribution, in agreement with the empirical observations. The activation of the recency mechanism does not affect the frequency rank distribution (purple hexagons).

**d,** However, when we look at the $K_s$ distribution, the EPR mechanism does not capture the power-law behavior observed on the empirical data.
In this work we explored this idea under a simple rank-based framework. We unveiled empirical evidences supporting the idea that human trajectories are biased towards recently visited locations. We also offered a different perspective for human mobility investigation, where the temporal dimension plays a role much more important than the inter-event times.

Methods

Datasets

Given that the coverage areas of mobile phone towers are far from uniform, the first step was to convert the antennas IDs in $D_1$ into more reasonable estimates of users’ locations. In fact, coverage areas frequently overlap and multiple towers can cover the same area—and they actually do. That is true specially for densely populated regions where the density of antennas is also high. Moreover, for commercial purposes, multiple communication antennas can be at the same location. In order to reduce the influence of these factors, we truncated the cell towers’ coordinates to the forth decimal place—which corresponds to antennas within less than 11 meters apart—and merged together those having the same (truncated) coordinates under the same id.

A rank-based characterization of human trajectories

In this work, we propose a rank-based approach to the analysis of human trajectories. For such, we defined two rank variables, namely the frequency rank ($K_f$) and the recency rank ($K_s$). Both ranks were measured in a expanding basis from the accumulated sub-trajectories. To illustrate, consider a particular user $x$ with a trajectory $T = [(l_1, l_2, \ldots, l_n), l_i \in [1, \ldots, N]]$ composed of $N$ steps to $S \leq N$ locations. For each step $j > 0$, we have the partial trajectory $T = [l_1, l_2, \ldots, l_{j-1}]$ composed of all the previous steps, with $l_{j-1}$ being the immediate preceding step. From the sub-trajectory $T$ we compute the frequency-based ranks $K_f$ of all locations visited so far. If the step $l_j$ is a return (i.e., $l_j \in T$) we say that the frequency rank of the location $l_j$ is the rank $k_f(j) = K_f[l_j]$.

The Exploration and Preferential Return Model

The Exploration and Preferential Return (EPR) individual mobility model, proposed by Song et al [18] is based on two components, the exploration and the preferential return. The exploration mechanism reproduces the scaling properties of the individuals’ jump length and radius of gyration distribution whereas the exploration reproduces the Zipf’s Law exhibited by the visitation frequencies to each location.

In the EPR model, at each time step an individual can either move or stay at the same location, according to the waiting time distribution

$$P(\Delta t) \sim \Delta t^{-1-\beta} \exp(\Delta t / \tau)$$

with $\beta = 0.8$ and $\tau \approx 17$ hours.

If the individual is going to move, it can either explore (i.e., visit a new location) with probability $\rho S^\gamma$ or return to a previously visited location with
complementary probability $1 - \rho S^\gamma$ where $\rho = 0.6$ and $\gamma = 0.6$ are the values selected according to the method described in Ref. [18], and $S$ is the number of previously visited locations. If the step corresponds to a return, a location is selected with a probability $P(f) \sim f^{-(1+1/\zeta)}$ to be proportional to the visitation frequency $f$ of such location and with $\zeta \approx 1.2 \pm 0.1$. If the next move corresponds to an exploration step, then a new location random location is selected according to a Lévy Flight

$$P(\Delta_r) \sim \Delta_r^{-1-\alpha} \exp(\Delta_r/\kappa)$$

with $\alpha \approx 0.55 \pm 0.05$ and $\kappa \approx 100$ km. Notice that the $\alpha$ parameter described here is not the same as the recency parameter of our model. In the context of this work, all the parameters proposed by Song et al [9] were kept intact.

**Author contributions**

Developed the ideas, methods and analyses: HB and RM. Empirical data analysis: HB and AE. Wrote the manuscript: HB, FBLN and RM.

**Additional information**

**Competing financial interests:** The authors declare no competing financial interests.

**References**

[1] Belik, V., Geisel, T. & Brockmann, D. Natural Human Mobility Patterns and Spatial Spread of Infectious Diseases. *Physical Review X* 1, 011001 (2011).

[2] Colizza, V., Barrat, A., Barthelemy, M., Valleron, A.-J. & Vespignani, A. Modeling the worldwide spread of pandemic influenza: baseline case and containment interventions. *PLoS medicine* 4, e13 (2007).

[3] Balcan, D. & Vespignani, A. Phase transitions in contagion processes mediated by recurrent mobility patterns. *Nature physics* 7 (2011).

[4] Toole, J. L., Ulm, M., González, M. C. & Bauer, D. Inferring land use from mobile phone activity. *Proceedings of the ACM SIGKDD international workshop on urban computing* 1–8 (2012). [arXiv:1207.1115v1](http://arxiv.org/abs/1207.1115v1).

[5] Lenormand, M., Gonçalves, B., Tugores, A. & Ramasco, J. J. Human diffusion and city influence. *arXiv* 1–17 (2015). URL [http://arxiv.org/abs/1501.07788](http://arxiv.org/abs/1501.07788) [1501.07788]

[6] Kitamura, R., Chen, C., Pendyala, R. & Narayanan, R. Micro-simulation of daily activity-travel patterns for travel demand forecasting. *Transportation* 25–51 (2000).
[7] Jung, W., Wang, F. & Stanley, H. Gravity model in the Korean highway. *EPL (Europhysics Letters)* 1–13 (2008). [arXiv:0710.1274v1]

[8] Krajzewicz, D., Hertkorn, G., Wagner, P. & Rössel, C. SUMO (Simulation of Urban MObility) An open-source traffic simulation Car-Driver Model (2011).

[9] Song, C., Qu, Z., Blumm, N. & Barabási, A.-L. Limits of predictability in human mobility. *Science (New York, N.Y.)* 327, 1018–21 (2010).

[10] Wang, D., Pedreschi, D., Song, C., Giannotti, F. & Barabási, A.-L. Human mobility, social ties, and link prediction. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD ’11*, 1100 (ACM Press, New York, New York, USA, 2011).

[11] Yang, Y., Herrera, C., Eagle, N. & González, M. C. Limits of predictability in commuting flows in the absence of data for calibration. *Scientific reports* 4, 5662 (2014). URL [http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4092333&tool=pmcentrez&rendertype=abstract](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4092333&tool=pmcentrez&rendertype=abstract) [arXiv:1407.6256v1]

[12] Sadilek, A. & Krumm, J. Far Out: Predicting Long-Term Human Mobility. *Aaai* 814–820 (2012).

[13] Krumme, C., Llorente, A., Cebrian, M., Pentland, A. S. & Moro, E. The predictability of consumer visitation patterns. *Scientific reports* 3, 1645 (2013).

[14] Lu, X., Wetter, E., Bharti, N., Tatem, A. A. J. & Bengtsson, L. Approaching the limit of predictability in human mobility. *Scientific reports* 3, 2923 (2013).

[15] González, M. C., Hidalgo, C. A. & Barabási, A.-L. Understanding individual human mobility patterns. *Nature* 453, 479–482 (2008). [10.1038/nature06827](https://doi.org/10.1038/nature06827)

[16] Brockmann, D., Hufnagel, L. & Geisel, T. The scaling laws of human travel. *Nature* 439, 462–5 (2006).

[17] Hasan, S., Schneider, C. M., Ukkusuri, S. V. & González, M. C. Spatiotemporal Patterns of Urban Human Mobility. *Journal of Statistical Physics* 151, 304–318 (2012).

[18] Song, C., Koren, T., Wang, P. & Barabási, A.-L. Modelling the scaling properties of human mobility. *Nature Physics* 6, 818–823 (2010).

[19] Schneider, C. M. *et al.* Unravelling daily human mobility motifs. *Journal of the Royal Society, Interface / the Royal Society* 10, 20130246 (2013).

[20] Merton, R. K. The Matthew Effect in Science. *Science (New York, N.Y.)* 159, 56–63 (1968).
[21] Price, D. A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science and Technology* 18, 239–255 (1976).

[22] Barabási, A.-L. & Albert, R. Emergence of scaling in random networks. *Science* 286, 509–512 (1999).

[23] Grabowicz, P., Ramasco, J., Gonçalves, B. & Eguíluz, V. Entangling mobility and interactions in social media. *PLoS One* 9, e85422 (2014).

[24] Cho, E., Myers, S. A. & Leskovec, J. Friendship and mobility. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD ’11*, 1082 (ACM Press, New York, New York, USA, 2011).

[25] Clauset, A., Shalizi, C. R. & Newman, M. E. J. Power-Law Distributions in Empirical Data. *SIAM Review* 51, 661–703 (2009).