Patterns of Human Mobility

Shijun Tang[1] and Sihai Tang[2]

[1]Department of Science and Mathematics
Alvernia University, Reading, PA 19607

[2]Department of Computer Science and Engineering
University of North Texas, Denton, TX 76203

Abstract— **Human movement usually has its own regular property. For example, most people will go to work/school, then go back home within a region. But sometimes some people may visit medium-range locations or few people may travel far places too. In this paper, we have investigated the distribution of human movements within different stages based on González et al studying the large data set. Corresponding to the given range (*i.e.* characteristic distance r_g), there exist three regions: near, medium and far movements. Our study indicates that the probability of human movement associated the three regions can be derived from the convolution between the statistics of individual trajectories and the population heterogeneity; its final form of the probability of human movement is still inverse power law distributions with a specific power exponent; that major activities of people mainly concentrate on the scope of less than 0.6r_g (characteristic distance r_g).**

Keywords — ***Human mobility, characteristic distance, patterns of human mobility, inverse power law distributions***

I. INTRODUCTION

Human mobility has spatial, temporal, and social characteristics. Human mobility pays a key role not only in modern economy and society, for instance, urban planning [1], traffic forecasting [2][3], but also in understanding biological and mobile epidemics [3][4][5][6] and mobile viruses [7] in interdisciplinary fields, such as biology, medicine, information technology, computer science, etc. Human mobility can simulate social networks and predict the performance of social /mobile networks, and determines the formation of several social networks, such as the type of people live, work, visitation and travel.

Human mobility is a very complex movement, including several types, *e.g*. long, medium and short distances, air, water and land transportation. Each trajectory of humans is very different. Many researchers have indicated that the emerging statistical pattern referring to humans has been often approximated with various random walk or diffusion models that typically obey Pareto inverse power law or Lévy flight distributions [8][9][10][11][12].

After studying the distribution of distances between consecutive sightings of nearly half-a-million bank notes, Brockmann *et al.* [13] proposed to analyze human traffic using the geographic circulation of money, and viewed the human trajectories as the Lévy flights with power law decay.

Cell phones can reflect human mobility very well since paths of mobile phones are directly corresponding to those of the phone users. So, cell phones are better than the current of bank notes [14][15][16[17] as an approximate tool of human movement which has been used to learn the structure of social networks and analyze patterns of human mobility and behavior. [14][18] explored the mobility patterns recorded from over 100,000 individuals selected randomly from a sample of more than 6 million anonymized mobile phone a six-month period, and concluded that the observed jump size distribution $P(\Delta r)$ is the convolution between the statistics of individual trajectories $P(\Delta r|r_g)$ and the population heterogeneity $P(r_g)$, where r_g is radius of gyration.

In this paper, we have further explored three cases under different slopes using the convolution between the statistics of individual trajectories $P(\Delta r|r_g)$ and the population heterogeneity $P(r_g)$, and provided the relationship among the observed jump size distribution, step size and radius of gyration.

II. METHODOLOGY

Our study in this paper is based on González *et al* [14] research, in which the data González *et al.* [14] used, includes data sets D_1 and D_2. The data set D_1 recorded over 100,000 individuals selected randomly from a sample of more than 6 million anonymized mobile phone users in a six-month period. Each time a user initiated or received a call or a text message, the location of the mobile phone tower routing the communication was recorded (González *et al.*[14]). The data

set D_2 that captured the location of 206 mobile phone users, recorded every two hours for an entire week. González et al. [14] measured the distance between user's positions at consecutive calls, capturing 16,264,308 displacements for the D_1 and 10,407 displacements for the D_2 data set.

González *et al.* [14] pointed out that $P(\Delta r)$ is the convolution between the statistics of individual trajectories $P(\Delta r|r_g)$ and the population heterogeneity $P(r_g)$, which is expressed as (1).

$$P(\Delta r) = \int_0^\infty P(\Delta r|r_g)P(r_g)dr_g \quad \text{..... (1)}$$

Where $P(\Delta r|r_g) \sim r_g^{-\alpha} F(\frac{\Delta r}{r_g})$ and $\alpha = 1.2 \pm 0.1$.

According to the Fig. 2b in González *et al.* [14]), we can find that the collapsed curve consists of three consecutive segments with different slopes. The slopes corresponding to three segments are -1.2, -0.6 and -4.8, respectively. The scope of Δr corresponding to three segments are $\Delta r < 0.6r_g$ $0.6r_g \leq \Delta r < 3r_g$ and $3r_g < \Delta r$.

Now, let's further explore and discuss patterns of human mobility from different perspectives as following.

Case I. $F(x) \sim x^{-\alpha}$ for $x < 1$. In fact, $\Delta r < r_g$ when $x < 1$ (*i.e.* $\frac{\Delta r}{r_g} < 1$). So, we can take $\Delta r < 0.6r_g$.

$$P(\Delta r|r_g) \sim r_g^{-1.2} (\frac{\Delta r}{r_g})^{-1.2}$$

That is, $P(\Delta r|r_g) = a_1(\Delta r)^{-1.2}$

Using the initial condition, a_1 =0.251 can be obtained. Thus,

$$P(\Delta r|r_g) = 0.251(\Delta r)^{-1.2} \quad \text{.....(2)}$$

Putting (2) into (1), we find that

$$P(\Delta r) = 0.251(\Delta r)^{-1.2} Q_1 \quad \text{.....(3)}$$

where

$$P(r_g) = (r_g + r_g^0)^{-\beta_r} \exp(-\frac{r_g}{k}) \quad \text{.....(4)}$$

and

$$Q_1 = \int_0^\infty (r_g + r_g^0)^{-\beta_r} \exp(-\frac{r_g}{k}) dr_g \quad \text{.....(5)}$$

Case II, for this stage between $0.6r_g$ and $3r_g$ and its slope can be computed as -0.6. Thus, under the condition: $0.6r_g \leq \Delta r < 3r_g$,

$$P(\Delta r|r_g) = a_2 r_g^{-0.6} (\Delta r)^{-0.6} \quad \text{..... .....(6)}$$

Using the initial condition, a_2=0.155 can be obtained. Putting Equ. (6) into Equ.(1), we find that

$$P(\Delta r) = 0.155 * (\Delta r)^{-0.6} Q_2 \quad \text{.....(7)}$$

where

$$Q_2 = \int_0^\infty r_g^{-0.6} (r_g + r_g^0)^{-\beta_r} \exp(-\frac{r_g}{k}) dr_g \quad \text{.....(8)}$$

Case III. The final stage is corresponding to the condition $3r_g < \Delta r$. The slope corresponding to the segments is -4.8. Thus,

$$P(\Delta r|r_g) = a_3 r_g^{3.6} (\Delta r)^{-4.8} \quad \text{.....(9)}$$

Using the one point in the final stage, a_3=0.02 can be obtained. Putting Equ. (9) into Equ. (1), we find that

$$P(\Delta r) = = 0.02(\Delta r)^{-4.8} Q_3 \quad \text{....(10)}$$

where

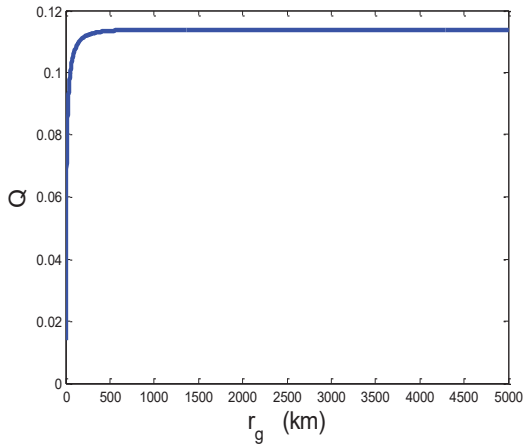
$$Q_3 = \int_0^\infty r_g^{3.6} (r_g + r_g^0)^{-\beta_r} \exp(-\frac{r_g}{k}) dr_g \quad \text{.....(11)}$$

III. EXPERIMENTAL RESULTS AND ANALYSIS

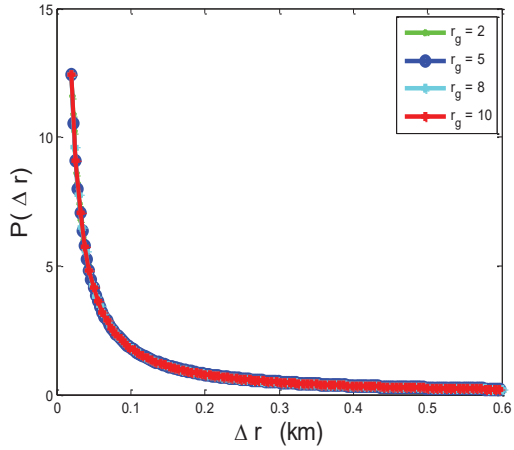
Based on the above descriptions, the Equ. (3), (7) and (10) associated with the above three cases can be further calculated. First, using numerical method, the Q_1, Q_2 and Q_3 can be obtained from Equ. (5), (8) and (11), respectively. For the infinitely great value (∞) -- the upper bound of integral, we tried 3000 km, 4000 km and 5000 km. The final value of integral Q_1 (Q_2) is the same when taking 3000 km, 4000 km and 5000 km as the upper bound of integral. The final value after integral of Q_3 has little difference when taking 3000 km, 4000 km and 5000 km as the upper bound of integral. Anyway, the final value of integral (Q_1, Q_2 and Q_3) will be one constant once taking one of 3000 km, 4000 km and 5000 km as the upper bound of integral. Therefore, the choice of upper bound of integral does not affect the nature of human mobility. The obtained integrals Q_1, Q_2 and Q_3 are shown in Fig. 1(a), 2(a) and 3(a), respectively. Taking into

account the actual data from the paper (González *et al.* 2008), we have taken 5000 km as the upper bound of integral.

In fact, we can find from Equ. (3), (7) and (10) that the final form of the observed jump size distribution is still determined by the inverse power law distributions after dividing three different stages, which are expressed as $P(x) \sim x^{-\alpha}$ and characterized by a specific power exponent α corresponding to different stage.

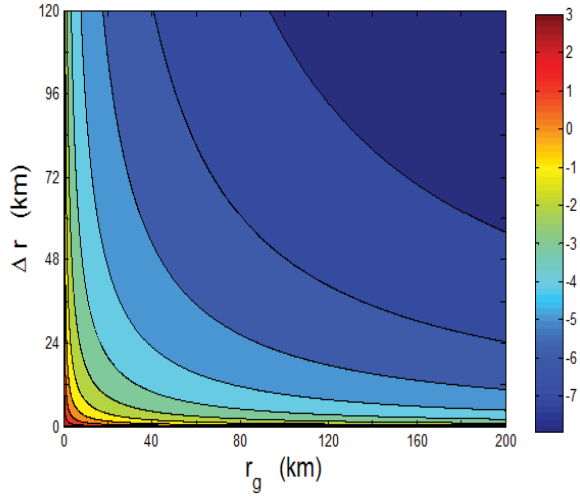


(*a*)



(*b*)

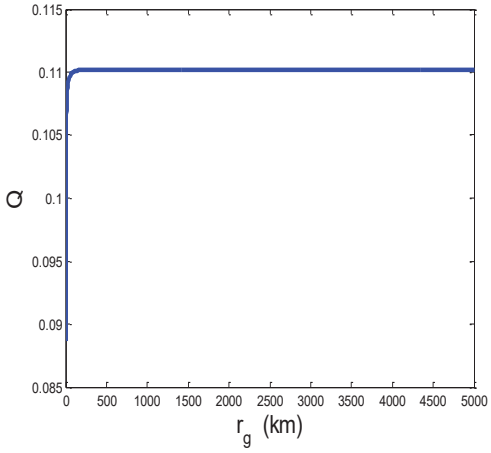
Fig. 1(b), 2(b) and 3(b) exhibit the $P(\Delta r)$ varying with Δr at different r_g. Since the scope of jump size is different at three cases, the observed jump size distribution exists big discrepancy at the same characteristic distance r_g. For smaller movement of human, there is a bigger probability, and the probability quickly decreases with increment of jump.



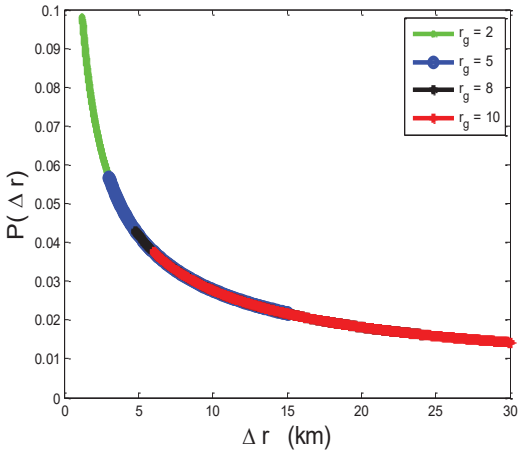
(*c*)

Fig. 1. shows the results from Equ. (3) at case I. (*a*). the integral value Q_1 of Equ. (5) at case I. (*b*). $P(\Delta r)$ varying with Δr at different r_g. (*c*). $P(\Delta r)$ varying with Δr and r_g.

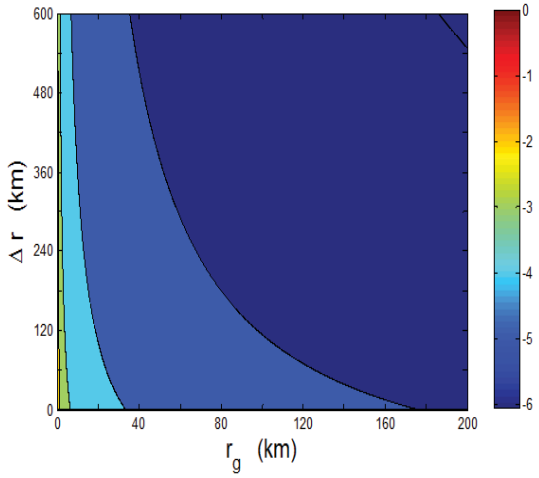
Fig. 1(c), 2(c) and 3(c) indicate that logarithm of probability of human movement $log(P(\Delta r))$ varies with different jump Δr and characteristic distance r_g. The color bar at the right side denotes the intensity of probability of human movement.



(*a*)



(*b*)



(*c*)

Fig. 2. shows the results from Equ. (7) at case II. (a). the integral value Q_2 of Equ.(8) at case II. (b). $P(\Delta r)$ varying with Δr at different r_g. (c). $P(\Delta r)$ varying with Δr and r_g.

In Fig. 1(c), there is the different probability (*i.e.* different color) with the increments of characteristic distance r_g and human mobility Δr at the same time. At certain characteristic distance r_g, the corresponding probability has different color (*i.e.* intensity value) in the whole scope of human mobility Δr. That is, the probability becomes small when the size of jump increasing at certain r_g. And at certain human movement Δr, different color (*i.e.* intensity value) in the whole scope of characteristic distance r_g suggests that the probability also becomes small with the increase of r_g.

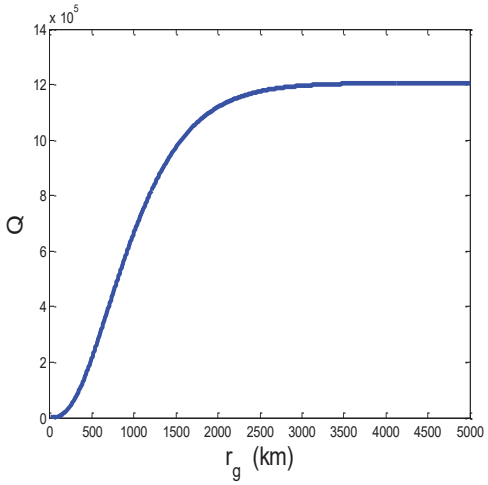
In Fig. 2(c) and 3(c), at certain characteristic distance r_g, the corresponding probability has the same color (*i.e.* intensity value) almost in the whole scope of human mobility Δr. But, the larger probability is only at the proximity of smaller r_g. So, Fig. 1(c), 2(c) and 3(c) show that all larger $P(\Delta r)$ locates at the proximity of smaller both Δr and characteristic distance r_g.

According to the above experimental results and analyses, three categories of reflecting human mobility patterns can be discribed as:

[1] Near case — whether Δr big or not, it seems like any movement (Δr) is following the same law: $P(\Delta r) \sim \frac{1}{\Delta r^{\alpha}}$.

[2] Medium case — with r_g increasing, the starting value of Δr will be increasing, and the maximum value of $P(\Delta r)$ will be decreasing. However, all $P(\Delta r)$ varying with Δr are overlapping on the same curve, suggesting that all $P(\Delta r)$ changing with Δr obeys the same law although the starting value of Δr and maximum $P(\Delta r)$ are different.

[3] Far case — for small r_g, there exists only small moving range Δr which obeys $P(\Delta r) \sim \frac{1}{\Delta r^{\alpha}}$. With the r_g increasing, the maximum value of $P(\Delta r)$ decreases. But, the starting value of Δr will be incremental. The curve of P(Δr) varying with Δr corresponding to different small r_g is discrete. As r_g increases, the curve of P(Δr) varying with Δr will overlap and finally become the same curve, implying that entire $P(\Delta r)$ varying with Δr will follows the same law for all r_g although the entire curve of $P(\Delta r)$ with Δr experiences the change from discrete, overlapping curves to identical ones.



(*a*)

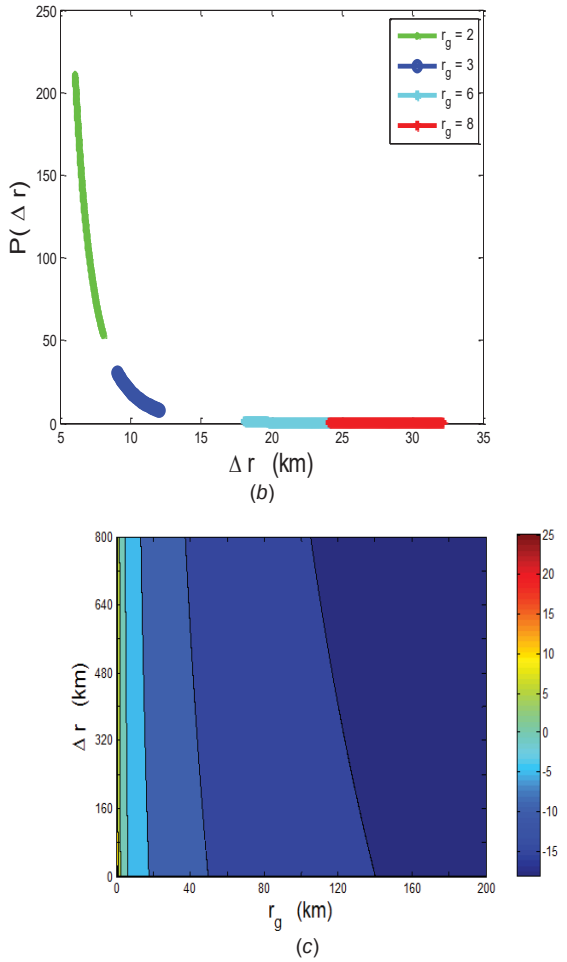


Fig. 3. shows the results from (10). (a) at case III. (*a*). the integral value Q_3 of Equ.(11) at case III. (b). $P(\Delta r)$ varying with Δr at different r_g. (c). $P(\Delta r)$ varying with Δr and r_g.

On the whole, patterns of human mobility can be classified into three different scenarios: far, medium and near movements, which are corresponding to their own exponent, respectively. The probability $P(\Delta r)$ will be decreasing when r_g goes far place from near one. Coefficient A is different in three different scenarios. However, the law of probability $P(\Delta r)$ with Δr has the same type: $P(\Delta r) = \frac{A}{\Delta r^{\alpha}}$.

IV. Conclusion

In this paper, we have explored the probability density function P(Δr) varying with step size Δr at different radius of gyration r_g, and the relationship among P(Δr), Δr and r_g based on González et al. work [14]. We have further studied on three cases under different slopes.

From the experimental results and analyses, we have concluded that although using the convolution between the statistics of individual trajectories and the population heterogeneity, the obeyed law of human mobility is still inverse power law distributions with a specific power exponent. Once characteristic distance r_g is selected, the people movement can be classified three categories: 0~0.6r_g , 0.6r_g~3r_g and more than 3r_g which follow the inverse power law distributions $P(x) \sim x^{-\alpha}$ with different power exponents α.

Thus, human mobility has its own similarity which does not change with respect to scale. For arbitrary characteristic distance r_g, the model always contains three categories (e.g. near, medium and far movements). That is, the choice of the characteristic distance r_g has not the important influence on the patterns of human mobility. Therefore, patterns of human mobility are that most people make movement within the $0.6r_g$ which takes up the majority. Some people travel medium-range distances between 0.6r_g and 3r_g. Few people move to far places (greater than 3r_g).

References

[1] Horner M.W, O'Kelly M.E.S. (2001). Embedding economies of scale concepts for hub networks design. J Transp Geogr, 9, 255–265.

[2] Kitamura R, Chen C, Pendyala R.M, Narayaran R. (2000). Micro-simulation of daily activity-travel patterns for travel demand forecasting. Transportation, 27, 25–51.

[3] Balcana D, Colizzac V, Goncalves B, Hu H, José J, Ramasco J, Vespignani A. (2009). Multiscale mobility networks and the spatial spreading of infectious diseases. PNAS, 106, 21484-21489.

[4] Eubank H, Guclu S, Kumar V.S.A , Marathe M, Srinivasan A, Toroczkai Z, Wang N. (2004). Controlling epidemics in realistic urban social networks. Nature, 429, 180–184.

[5] Hufnagel L, Brockmann D, Geisel T. (2004) Forecast and control of epidemics in a globalized world. *Proc Natl Acad Sci USA,* 101, 15124–15129.

[6] Funk S, Salathé M, Jansen V.A.A. (2010). Modelling the influence of human behaviour on the spread of infectious diseases: a review. J R Soc Interface, **7,** 1247-1256.

[7] Kleinberg J. (2007). The wireless epidemic. Nature, 449, 287–288.

[8] Viswanathan G.M, Afanasyev V, Buldyrev S.V, Murphy E.J, Prince P.A, Stanley H.E. (1996). Lévy flight search patterns of wandering albatrosses. Nature, 381, 413.

[9] 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. (2007). Revisiting Lévy flight search patterns of wandering albatrosses, bumblebees and deer. Nature (London), 449, 1044-1048.

[10] Ramos-Fernández G, Mateos J.L, Miramontes O, Cocho G, Larralde H, Ayala-Orozco B. (2004). Levy walk patterns in the foraging movements of spider monkeys (Ateles geoffroyi). Behav Ecol Sociobiol, 55, 223.

[11] Sims D.W, Southall E.J, Humphries N.E, Hays G.C, Bradshaw C.J.A, Pitchford J.W, James A, Ahmed M.Z, Brierley A.S, Hindell M.A, Morritt D, Musyl MK, Righton D, Shepard E.L.C, Wearmouth V.J, Wilson R.P, Witt M.J, Metcalfe J.D. (2008). Scaling laws of marine predator search behavior. Nature (London), 451, 1098.

[12] Rhee I, Shin M, Hong S, Lee K, Kim S.J, Chong S. (2011). On the levy-walk nature of human mobility. IEEE/ACM Trans Netw, 19, 630.

[13] Brockmann D, Hufnagel L, Geisel T. (2006). The scaling laws of human travel. Nature 439: 462–465.

[14] González M.C, Hidalgo C.A, Barabási A.L. (2008). Understanding individual human mobility patterns. Nature (London), 453, 779.

[15] Hidalgo C.A, Rodriguez-Sickert C. (2008). The dynamics of a mobile phone network. Physica A, 387, 3017-3024.

[16] González M.C, Lind P.G, Herrmann H.J. (2006). System of Mobile Agents to Model Social Networks. Phys Rev Lett, 96, 088702.

[17] Lambiotte R, Blondel V.D, Kerchove C.D, Huens E, Prieur C, Smoreda Z, Dooren P.V. (2008). Geographical dispersal of mobile communication networks. Physica A, 387,5317-5325.

[18] Eagle N, Pentland A.S. (2006). Reality mining: sensing complex social systems. *Personal Ubiquitous Computing,* 10, 255–268.