Supplementary Information
Unraveling the origin of exponential law in intra-urban human mobility

Xiao Liang, Jichang Zhao, Li Dong, and Ke Xu

I. DISTANCE DISTRIBUTIONS IN URBAN AREAS

There are three candidate models (power law, exponential and truncated Pareto) to be compared for describing the empirical distribution of travel distances. The parameters of these models are determined by the method of maximum likelihood estimation (MLE). The Akaike weights are calculated, which are considered as relative likelihoods being the best model. The model with the largest Akaike weight should be selected as the best one. The details of methods can be found in [1, 2].

As shown in Fig. S1, the distribution of trip lengths in each city is better fitted by the exponential rather than the power law or truncated Paretos. Moreover, the Akaike weight of exponential model (w_{EXP} equals about 1.0) is significantly larger than the ones of the other two models in each city. These all demonstrate that the exponential law of collective human movements exists in urban areas of cities extensively.

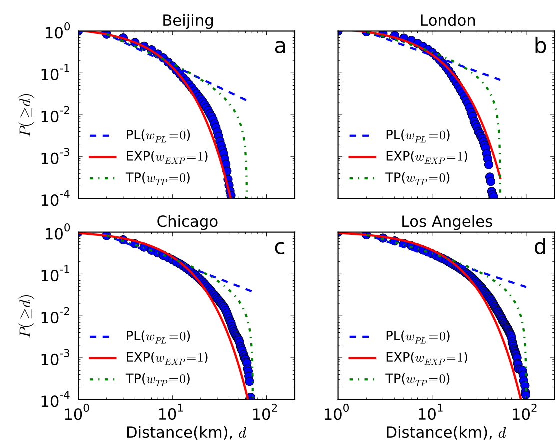

FIG. S1. The complementary cumulative distributions of travel distances in four cities. For each city, the fitted results of three models (power law (PL), exponential (EXP) and truncated Pareto (TP)) are plotted by the blue dashed line, the red solid line and the green dash-dot line respectively. w_{PL}, w_{EXP}, w_{TP} are the Akaike weights for the three models.

Furthermore, the fitted results of the exponential model in four cities are shown in Table S1. The 95% CI represents the 95% confidence interval of the exponential parameter. And the goodness of fit (GOF) is measured by the Kolmogorov-Smirnov statistic. The smaller the value is, the more similar the empirical distance distribution and the fitted exponential model are.

TABLE S1. The results of best fitted exponential model in four cities.

City	Parameter(EXP)	95% CI	GOF
Beijing	0.2283	(0.2282, 0.2284)	0.0843
London	0.2061	(0.2050, 0.2072)	0.0122
Chicago	0.1084	(0.1061, 0.1106)	0.0133
Los Angeles	0.0774	(0.0759, 0.0789)	0.0102

II. PROOF OF THE TRIP-LENGTH DISTRIBUTION

As shown in Fig. S2, the center is the point O and the non-increasing density distribution is $\rho(r)$ $(0 \leq r \leq R)$, where r is the distance to the center O and R is the size of a city. By using our model, we can estimate the displacement

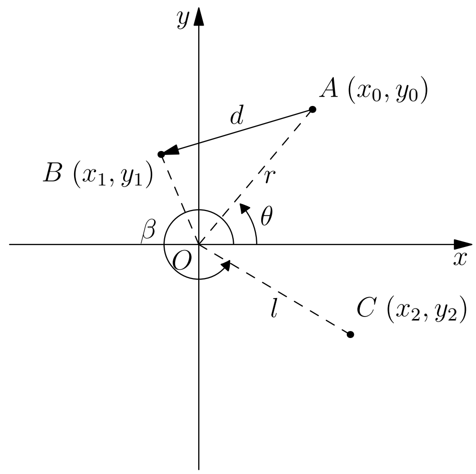

FIG. S2. Illustration to the proof of distance distribution.

distribution of human movements as follows:

$$P(d) = \sum_{d(A,B)=d} P(A \to B)$$

Supporting continuity of the density distribution, it can also be written by

$$P(d) = \iint \rho(A) \frac{\int_{s:d(A,B)=d} \rho(B) d^{-\sigma} \mathrm{d}s}{\iint_{C \neq B} \rho(C) d(C,A)^{-\sigma} \mathrm{d}S_2} \mathrm{d}S_1$$

Let

$$I_1 = \int_{s:d(A,B)=d} \rho(B) d^{-\sigma} \mathrm{d}s$$

and

$$I_2 = \iint_{C \neq B} \rho(C) d(C,A)^{-\sigma} \mathrm{d}S_2.$$

Therefore,

$$\begin{aligned} I_1 &= \int_{s:(x_1-x_0)^2+(y_1-y_0)^2=d^2} \rho(\sqrt{x_1^2+y_1^2}) d^{-\sigma} \mathrm{d}s \\ &\quad (x_1 = x_0 + d\cos\alpha, y_1 = y_0 + d\sin\alpha) \\ &= \int_0^{2\pi} d^{1-\sigma} \rho(\sqrt{x_0^2+y_0^2+d^2+2d(x_0\cos\alpha + y_0\sin\alpha)}) \mathrm{d}\alpha, \end{aligned}$$

$$\begin{aligned} I_2 &= \iint_{(x_2,y_2)\neq(x_0,y_0)} \frac{\rho(\sqrt{x_2^2+y_2^2})}{(\sqrt{(x_2-x_0)^2+(y_2-y_0)^2})^{\sigma}} \mathrm{d}x_2 \mathrm{d}y_2 \\ &\quad (x_2 = l\cos\beta, y_2 = l\sin\beta) \\ &= \int_0^{2\pi} \mathrm{d}\beta \int_0^R \frac{l\rho(l)}{(\sqrt{l^2+x_0^2+y_0^2-2l(x_0\cos\beta+y_0\sin\beta)})^{\sigma}} \mathrm{d}l. \end{aligned}$$

Consequently, $P(d)$ can be represented as

$$
\begin{aligned}
P(d) &= \iint \rho(\sqrt{x_0^2+y_0^2})\frac{I_1}{I_2}\mathrm{d}x_0\mathrm{d}y_0 \\
&\quad (x_0 = r\cos\theta, y_0 = r\sin\theta) \\
&= \int_0^{2\pi}\mathrm{d}\theta\int_0^R r\rho(r)\frac{U}{V}\mathrm{d}r,
\end{aligned}
$$

where

$$
\begin{aligned}
U &= \int_0^{2\pi} d^{1-\sigma}\rho(\sqrt{r^2+d^2+2dr\cos(\theta-\alpha)})\mathrm{d}\alpha, \\
V &= \int_0^{2\pi}\mathrm{d}\beta\int_0^R l\rho(l)(\sqrt{l^2+r^2-2rl\cos(\beta-\theta)})^{-\sigma}\mathrm{d}l.
\end{aligned}
$$

It is noticed that the denominator V has nothing to do with d. Assuming the non-increasing density function

$$\rho(r) = Ce^{-\lambda r}(\lambda \geq 0),$$

the numerator U satisfies

$$U \geq C\int_0^{2\pi} d^{1-\sigma}e^{-\lambda(r+d)}\mathrm{d}\alpha = 2\pi C d^{1-\sigma}e^{-\lambda r}e^{-\lambda d}.$$

In a similar way,

$$U \leq C\int_0^{2\pi} d^{1-\sigma}e^{-\lambda(d-r)}\mathrm{d}\alpha = 2\pi C d^{1-\sigma}e^{\lambda r}e^{-\lambda d}.$$

As a result, when $\lambda = 0$, that is a uniform density function,

$$C_1 d^{1-\sigma} \leq P(d) \leq C_2 d^{1-\sigma}.$$

And when $\lambda > 0$, that is a negative exponential density funciton,

$$C_1 d^{1-\sigma}e^{-\lambda d} \leq P(d) \leq C_2 d^{1-\sigma}e^{-\lambda d}.$$

[1] Edwards AM, et al. (2007) Revisiting Lévy flight search patterns of wandering albatrosses, bumblebees and deer. *Nature* 449:1044–1048.

[2] Edwards AM, Freeman MP, Breed Ga, Jonsen ID (2012) Incorrect likelihood methods were used to infer scaling laws of marine predator search behaviour. *PLoS ONE* 7:e45174.