Analysis of Mobility Patterns

Galen Wilkerson

March 13, 2013

Brockmann, D., *The Scaling Laws of Human Travel*

Compares human travel distances to *dispersal curves* found in biology, using movement of bank notes. [1]

Definitions

- *Levy Flight*: Dispersal curves lacking typical length scale.

$$P(r) \sim r^{(-1+\beta)},$$

  with $\beta < 2$.

- *Continuous Time Random Walk*: succession of random displacements δx and random waiting times δt_n, each drawn from $P(\delta x)$ and $\phi(\delta t)$.

Initial measurement of short-time trajectories

Analysed distribution of spatial displacement, at short-time ($T < 14$ days) trajectories for cities of three sizes:

1. $N > 120,000$,
2. $22,000 < N < 120,000$,
3. $N < 22,000$,

found $P(r) \sim r^{(-1+\beta)}$, with $\beta = 0.59 \pm 0.02$, for *all city sizes.*

**Did not correct for anisotropy (c.f. Gonzalez et al.)

Claims this power-law is intrinsic and universal property of dispersal.

Dispersal attenuation

Levy flights predict time to reach stationary, equilibrium distribution $T \approx 68$ days.

However, at $T > 100$ days, many notes (23%) still did not travel very far (< 800km). I.e. much slower dispersal than predicted.

Possible causes of dispersal attenuation:

1. spatial inhomogeneities - less likely to leave large cities
2. long periods of rest - fat tail in $\phi(t)$ leads to subdiffusion: ambivalence between scale-free displacements and scale-free rest periods.

To investigate, analysed distribution of waiting times:
$P_0^i(t)$ is the probability a bank note is reported at the initial location at time t.

Found

$$P_0(t) \sim At^{-\eta},$$

with $\eta = 0.6 \pm 0.03$.

For pure two-dimensional Levy flight,

$$t^{-2/\beta}, \beta \approx 0.6 \implies \eta = 3.33.$$

Observed 'decay' is very slow.
They conclude slow decay in time exponent is result of rests.
If $\phi(t) \sim t^{(-1+\alpha)}$, with $\alpha < 1$, then $\eta = \alpha \approx 0.60$.

Continuous time random walks (CTRW)

Position, elapsed time are:

$$x_N = \sum_n \delta x_n, \qquad t_N = \sum_n \delta t_n.$$

The probability density $W(x,t)$.

For finite-variance (σ^2) displacements, and finite mean (τ) waiting times, obtain

$$\partial W(x,t) = D\partial_x^2 W(x,t),$$

with $D = \sigma^2/\tau$.

However, since here both displacement and waiting times are fat-tailed, having σ^2 and τ infinite, obtain:

$$\partial_t^\alpha W(x,t) = D_{\alpha,\beta} \partial_{|x|}^\beta W(x,t),$$

with $D_{\alpha,\beta}$ as the diffusion coefficient.
This gives

$$W_r(r,t) = t^{-\alpha/\beta} L_{\alpha,\beta}(\frac{r}{t^{\alpha/\beta}}),$$

for universal scaling function $L_{\alpha,\beta}$, which implies

$$r(t) \sim t^{1/\mu},$$

where $\mu = \beta/\alpha$.
(i.e r scales with the ratio of $\mu = \beta/\alpha$.)

This ratio ($\mu \approx 1$) implies super-diffusive travel patterns (despite fat-tailed rest).

Plots $W_r(z,t)$ on log-log scale (Fig. 2a), $z = log_1 0r, \tau = log_1 0t$.
obtains scaling exponent $\mu = 1.05 \pm 0.02$, claims agrees with their model.

Claim data supports $W_r(r,t)$ model (above).

Claim bank notes passed at constant rate, implying that there is a 'lack of scale in human waiting-time statistics' (??).
(validate against transportation data)

Conclude that 'dispersal of bank notes and human travel behaviour can be described by a continuous-time random walk process that incorporates scale-free jumps as well as a long waiting times between displacements.

Gonzalez, M., et al. *Understanding individual human mobility patterns*

Human mobility demonstrates 'temporal and spatial regularity'. Characteristic travel distance and high probability of returning to frequent locations, using mobile phone call data. [2]
Animal trajectories follow Levy flights.
Address 'bursty' call pattern: validated call pattern against sample of traced location.

Definitions

- *Levy flight*: "a random walk for which step size Δr follows a power-law distribution $P(\Delta r) \sim r^{-(1+\beta)}$".

- *Radius of gyration*: (the linear size occupied by each user's trajectory up to time t)

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

with $\vec{r}_i^a$ as the $i = 1, ..., n_c^a(t)$ positions for user a and

$$\vec{r}_{cm}^a = \frac{1}{n_c^a(t) \sum_{i=1}^{n_c^a} \vec{r}_i^a}.$$

as the center of mass of the trajectory.

"It is calculated as the root mean square distance of the objects' parts from either its center of gravity or a given axis." (Wikipedia)

"Conceptually, the radius of gyration is the distance that, if the entire mass of the object were all packed together at only that radius, would give you the same moment of inertia. That is, if you were to take the entire mass of the disk-shaped flywheel with some radius and pack it into a narrow donut whose radius is the flywheel's radius of gyration, they'd both have the same moment of inertia. With the same moment of inertia, they will behave very similarly when you spin them." (http://van.physics.illinois.edu/)

Spatial distribution

Distribution of displacements well-modelled by truncated power-law:

$$P(\Delta r) = (\Delta r + \Delta r_0)^{-\beta} e^{(-\Delta r/\kappa)}, \tag{1}$$

with $\beta = 1.75 \pm 0.15, \Delta r_0 = 1.5$km, and cutoff values $\kappa|_{D1} = 400$ km and $\kappa|_{D2} = 80$ km.

Claim "not far" (?? yes, but on log scale ??) from Brockmann's $\beta = 1.59$.

Three possible explanations for (1):

A. Each individual follows Levy trajectory with step size described by (1)

B. (1) captures population-based heterogeneity between individuals.

C. both (A) and (B).

To investigate, calculated the *radius of gyration* of each individual, i.e. the characteristic distance traveled by the user when observed up to time t.

Found distribution of radius of gyration:

$$P(r_g) = (r_g + \Delta r_g^0)^{-\beta_r} e^{(-r_g/\kappa)},$$

with $r_g^0 = 5.8$km, $\beta_r = 1.65 \pm 0.15$, and $\kappa = 350$km.

Using $P(\Delta r)$ and $P(r_g)$, claim they can rule out hypothesis A, above.

??

RW = random walk
LF = Levy flight
TLF = truncated Levy flight

Effects of time

For LF and TLF,

$$r_g(t) \sim t^{3/(2+\beta)}.$$

For RW,

$$r_g(t) \sim t^{1/2}.$$

Logarithmic increase describes r_g.
(So which, TLF, LF, or RW?)

Claim single jump size distribution (normalized by r_g),
and thus obtained

$$P(\Delta r | r_g) \sim r_g^{-\alpha} F(\Delta r / r_g),$$

with $\alpha \approx 1.2 \pm 0.1$ and $F(x) \sim x^{-\alpha}$.

Claim travel patterns can be characterized and bounded by r_g, since large displacements are fewer than expected in Levy flights.

Note that probs are related by

$$P(\Delta r) = \int_0^\infty P(\Delta r | r_g) P(r_g) dr_g,$$

implying the relation between the $P(r_g)$ and $P(\Delta r | r_g)$ exponents: $\beta = \beta_r + \alpha - 1$, implying hypothesis C (above).

Probability of return

In 2-dimensionsal random walk, return probability should follow

$$F_{pt} \sim \frac{1}{(t \ln^2(t))}.$$

Found peaks in return probability indicating recurrence/periodicity. (Fig. 2c)

Plotted visit rank (L) of locations vs. $P(L)$, found $P(L) \sim 1/L$, claim comes from regularity of travel patterns.(Fig. 2d)

Plotted $\Phi_a(x, y)$, the probability density function. (Fig. 3a) for $r_g \leq 3, 20 \leq r_g \leq 30 and r_g > 100$km , showing increasing anisotropy.

Defined anisotropy ratio $S = \sigma_x/\sigma_y$,
found

$$S \sim r_g^{-\eta},$$

for $\eta \approx 0.12$.

Plotted $\tilde{\Phi}_a(x/\sigma_x, y/\sigma_y)$, with σ the standard deviation in the individual's reference frame, showing similar travel patterns for all r_g sets.

Claim cross-section $\Phi(x, \sigma_x, 0)$ shows universal $\tilde{\Phi}(\tilde{x}, \tilde{y})$ probability distribution (Fig. 3d).
(?? yes, but only showed at y = 0 ??)
Claim this supports Hypothesis C above.
Final note that $\tilde{\Phi}(\tilde{x}, \tilde{y})$ is r_g independent.

Noulas, A., et al., *A tale of many cities: universal patterns in human urban mobility*

Smartphone social media check-ins allow detailed analysis of mobility patterns. [3]
Look at costs (gravity model) versus intervening opportunities.

Data

Foursquare data from GPS-enabled smartphones
925,030 users
over 6 months
5 million places

34 cities
10-meter resolution

Gravity Models versus Intervening Opportunities

Gravity model assumes trip patterns driven primarily by time, energy cost of mobility.

Perhaps physical distance doesn't capture universal patterns. (Due to differences in patterns between large cities.)

Probability of traveling to a destination a result of intervening opportunities, depends indirectly on physical distance.

Obtain power law for longer trips (but dataset is huge, many cities!), of $\beta = 1.50$.

Claim is "almost identical" to Brockmann (1.59) and Gonzalez (1.75) results. (Hmm, remember is order of magnitude, we need error bars.)

For shorter trips, they claim there is no power law, obtaining $r_0 = 18.42$ and $\beta = 4.67$, with $p - value = 1.0$.

They detect an abrupt cut-off at $\delta_m \in [5, 30]$ where distribution changes from roughly uniform to decreasing, perhaps corresponding to the borders of the city.

Place density

> Stouffer: 'number of persons travelling a given distance directly proportional to the number of opportunities at that distance and inversely proportional to the number of intervening opportunities.'

plot place density vs. average distance of displacements - inversely proportional
$R^2 = 0.59$

Absolute city area vs. average distance of displacements reveals little correlation.

Two models of mobility:

1. Rank model:

   The chance of visiting a place is inversely proportional to the number of intervening destinations between the user and the destination.

$$Pr[u \to v] \propto \frac{1}{rank_u(v)^a},$$

where

$$rank_u(v) = |w : d(u,w) < d(u,v)|.$$

More detail: pairwise transition probability $u \to v$:

$$P_{uv} = \frac{rank_u(v)^{-\alpha}}{\sum_{u \in U} rank_u(v)^{-\alpha}}$$

where $rank_u(v) = |w \in U : d(u,w) < d(u,v)|$
(the number of destinations closer to u than v),
and $rank_u(u) = 0$ for every $u \in U$.
This gives:

$$P_u(\Delta r) := \sum_{v: d(u,v) \in |\Delta r, \Delta r + \epsilon|} P_{uv},$$

with ϵ the spatial resolution.

$$P(\Delta r) = \frac{1}{M} \sum_{u \in U} P_u(\Delta r)$$

2. Gravity model:

- deterrence by distance $d(u,v)$ between two places u, v.
- m_u: gravitational mass of place u by number of nearby (inside r_u) settlements

p(transition$_{u \to v}$):

$$P_g[u \to v] \propto \frac{m_u m_v}{d(u,v)^b}$$

Claim output of rank model matches real data (Figure 8) (seems to)

Claim distribution of places determines mobility patterns.

Randomization of places to test above hypothesis:

KL divergence between between empirical displacements H and dist. R obtains by rank-distance model:

$$D_{KL}(H||R) = \sum_i H(i) ln \frac{H(i)}{R(i)}.$$

Note potential demographic bias in Foursquare data.

Used Clauset, Shalizi, Newman fitting methods.

References

[1] D Brockmann, L Hufnagel, and T Geisel, *The scaling laws of human travel*, Nature **439** (2006), no. 7075, 462–465.

[2] Marta C Gonzalez, Cesar A Hidalgo, and Albert-Laszlo Barabasi, *Understanding individual human mobility patterns*, Nature **453** (2008), no. 7196, 779–782.

[3] Anastasios Noulas, Salvatore Scellato, Renaud Lambiotte, Massimiliano Pontil, and Cecilia Mascolo, *A tale of many cities: universal patterns in human urban mobility*, PloS one **7** (2012), 1–8.