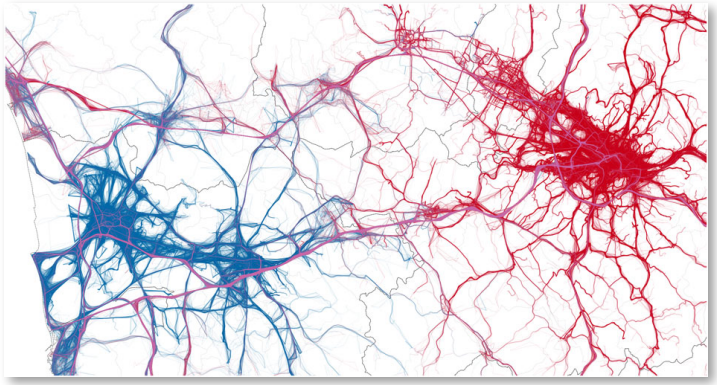
A Review of Human Mobility

Presenter: Tongtong Liu

Background

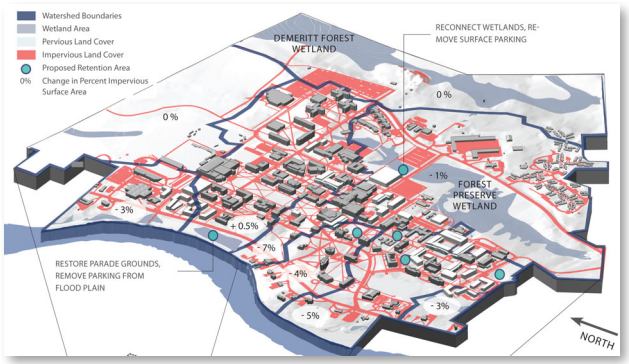
- Human mobility pattern



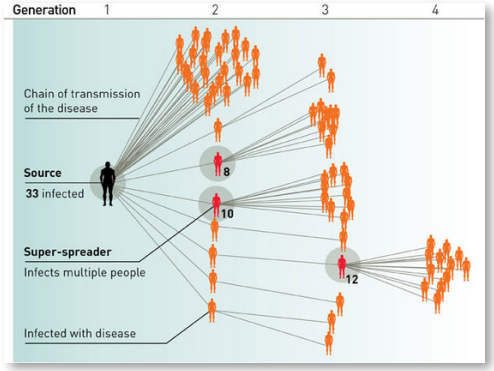
Migration

Commuting

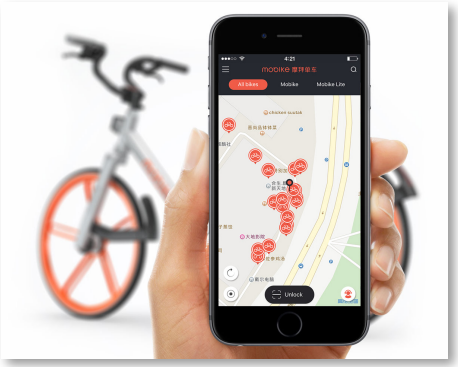
- Applications



Urban Planning



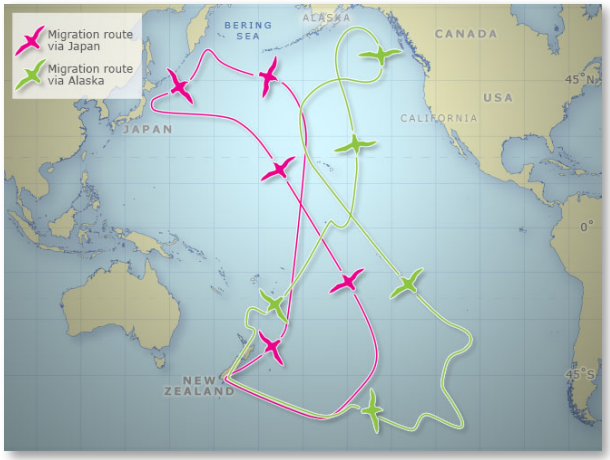
Epidemic spread



Mobike

Data Source

- Indirect data



Animal Migration

Bank Note

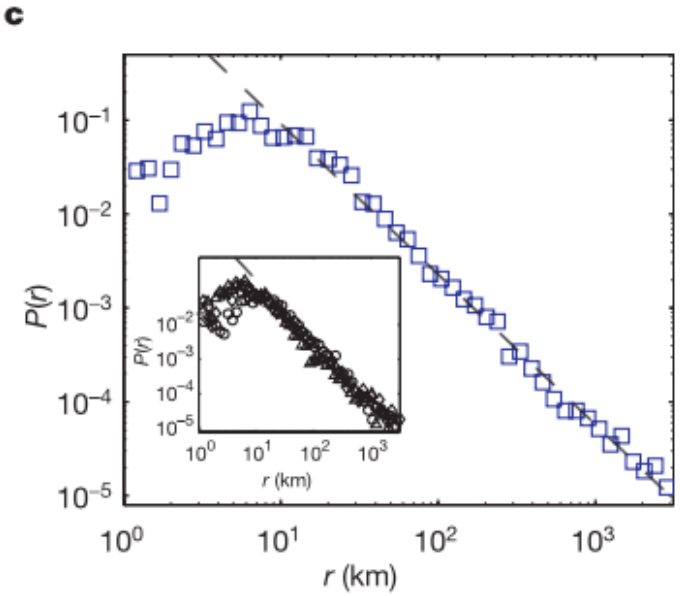
- Direct data
  - CDR: Call Detail Record
  - GPS: Global Positioning System
  - Cellular Data

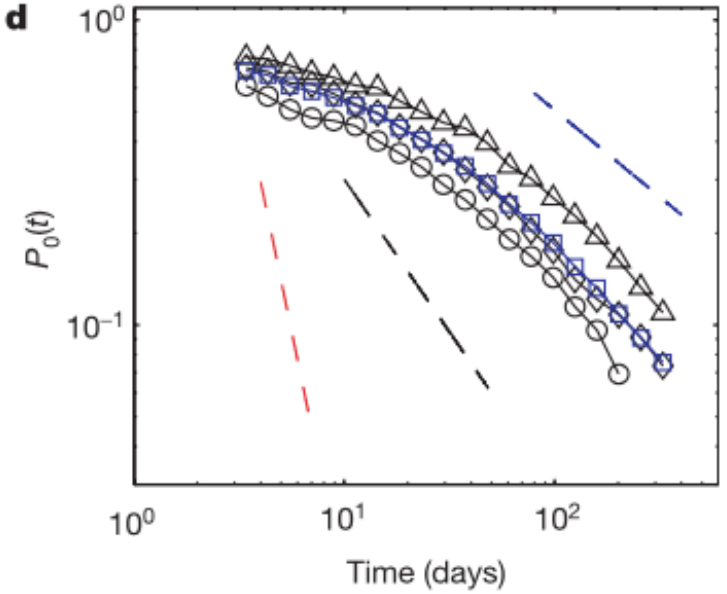
Random Walks

- Data source: animals, bank notes
- Models: Levy flight, Continuous-Time Random Walk
- Distribution of jump sizes and waiting times

$$P(\Delta r) \sim \Delta r^{-(1+\alpha)}, 0 < \alpha \leq 2$$

$$P(\Delta t) \sim \Delta t^{-(1+\beta)}, 0 < \beta \leq 1$$





Drawbacks

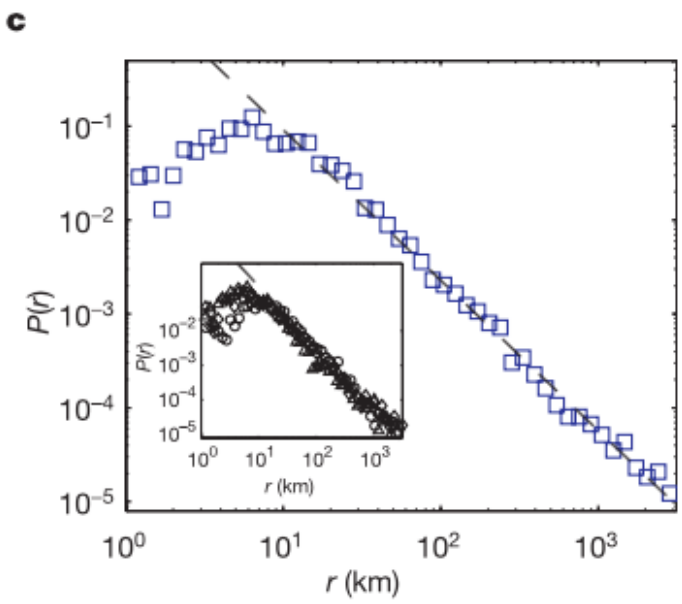
- Compared to animals, human mobility has much more constrains.
  - City transport lines, vehicles
  - Important locations, such as home, workplace
  - Weekday & weekend
- Indirect proxies don’t reflect the trajectory of only one person.
  - Each consecutive sighting of a bank note reflects the composite motion of two or more individuals.

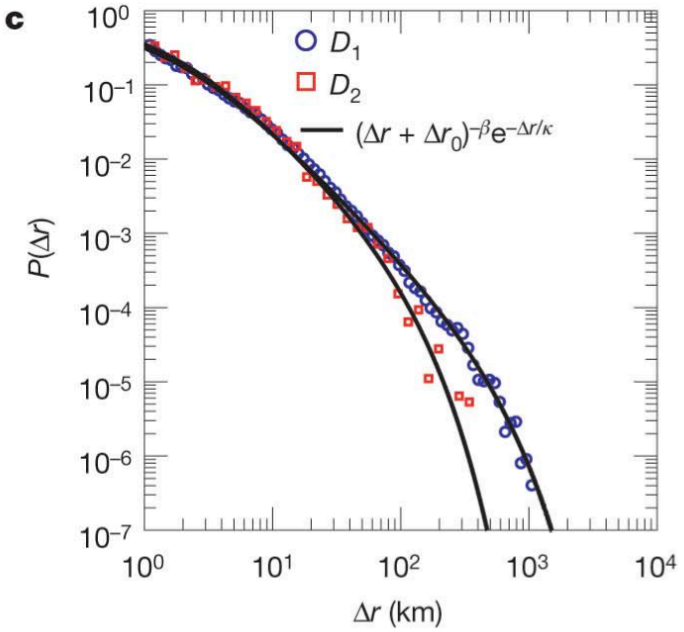
Test of Random Walks on CDR

- Distribution of jump sizes and waiting times
  - Power law -> Truncated power law

$$P(\Delta r) = (\Delta r + \Delta r_0)^{-\beta} exp(-\Delta r/k)$$

  - Difference



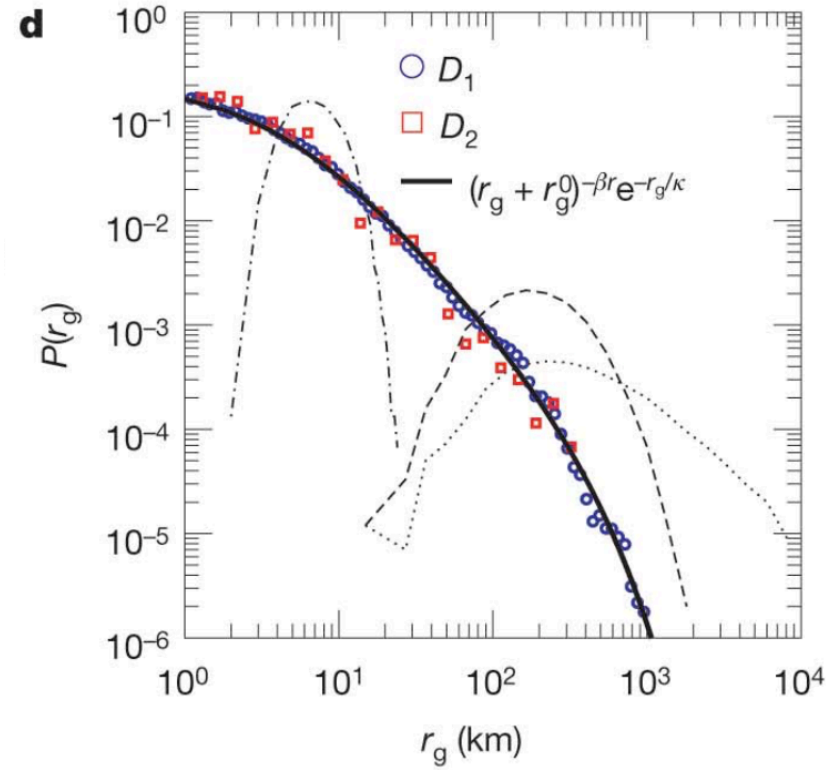


Test of Random Walks on CDR

- Radius of gyration

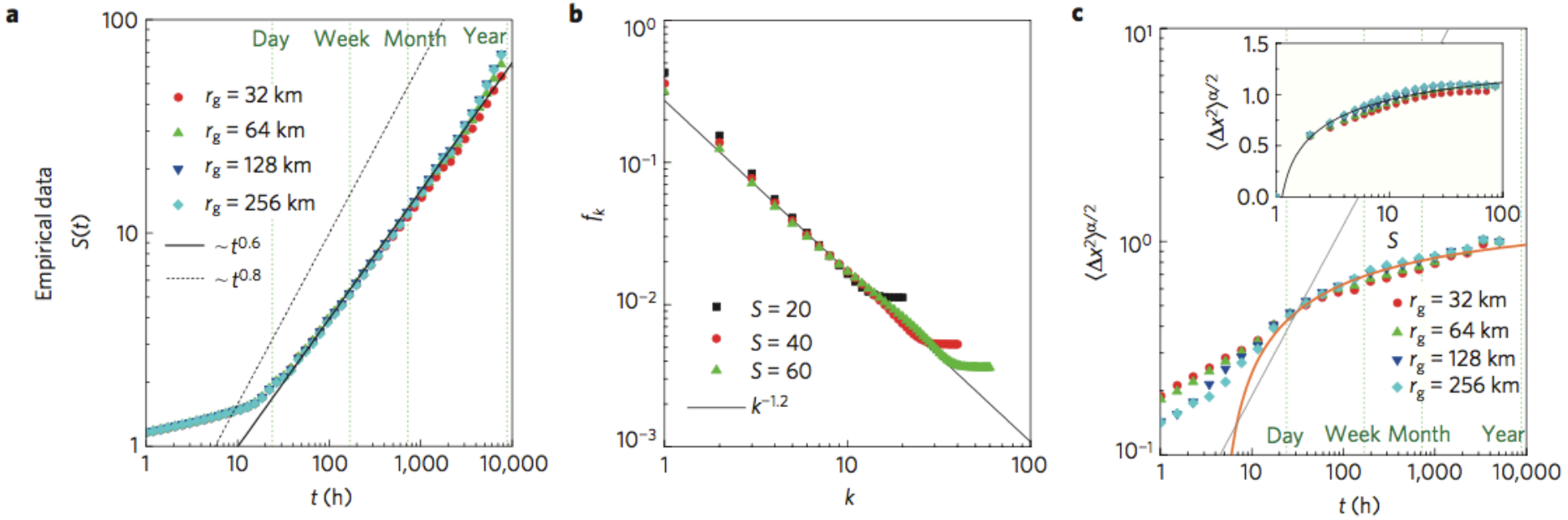
$$r_g = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(r_i - r_{cm})^2}$$

- The dotted, dashed and dot-dashed curves show P(rg) obtained from the standard null models (RW, LF and TLF, respectively)
  - RW: Random Walk
  - LF: Levy Flight
  - TLF: Truncated Levy Flight



Test of Random Walks on CDR

- RW models cannot explain observed scaling properties.



a. The number of distinct locations: $S(t) \sim t^{\mu}$

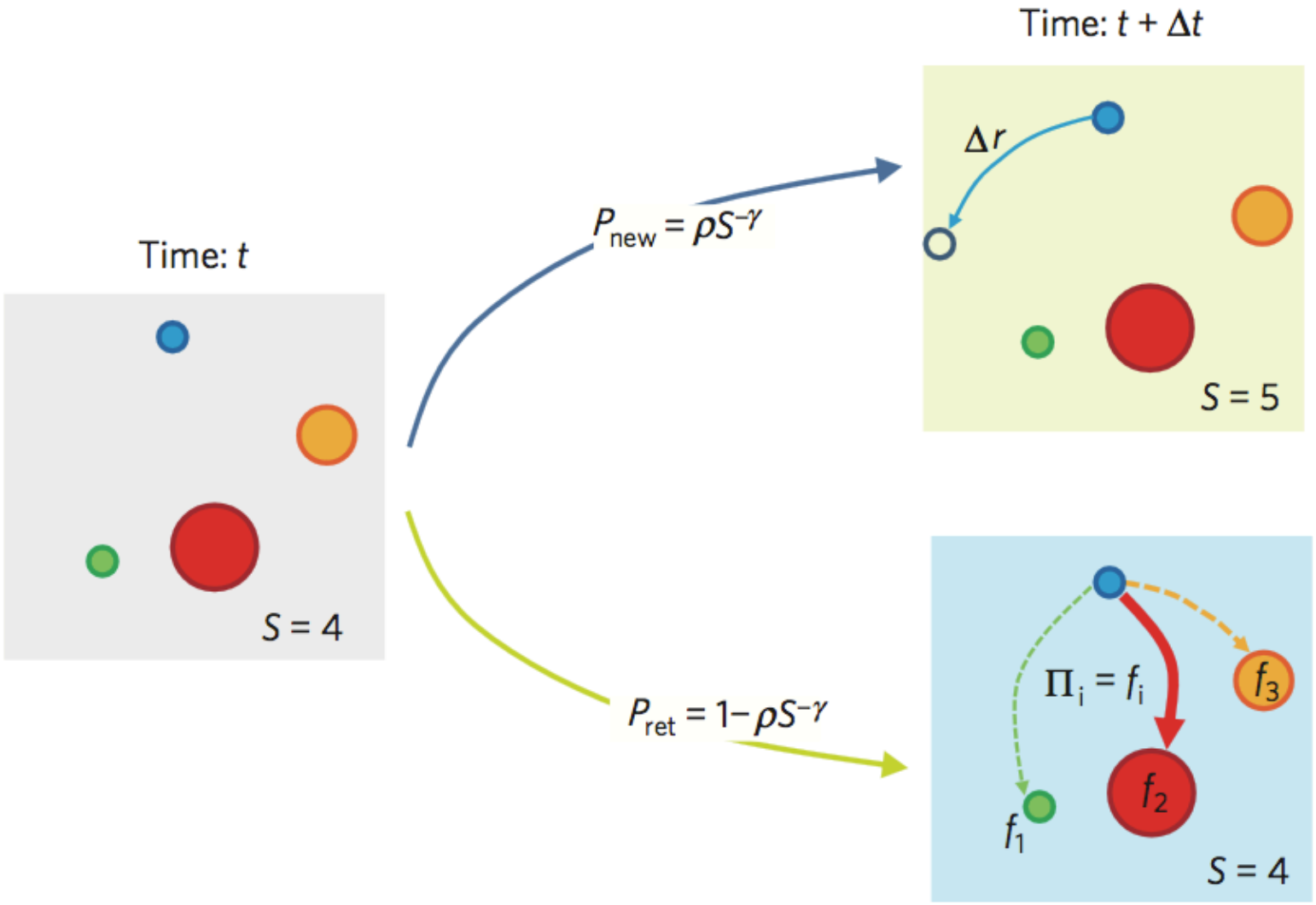
b. Visitation frequency of the kth most visited location: $f_k \sim k^{-\xi}$

c. Ultraslow diffusion of the mean square displacement: $\Delta x^2(t) \sim t^{v}$

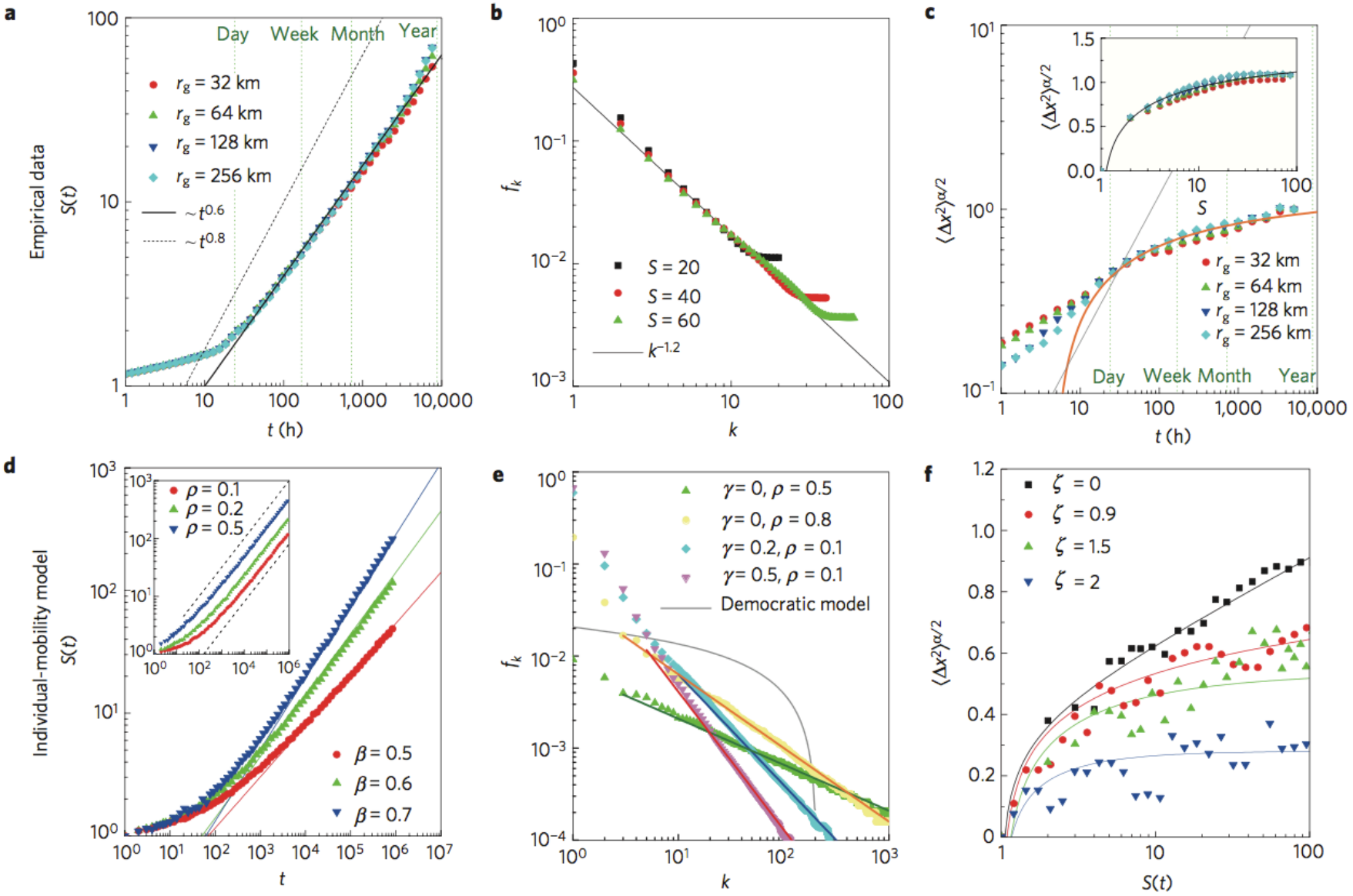
EPR Model

- EPR: Exploration and Preferential Return
- Exploration:
  - Random walk: the next diffusive step is in independent of the previously visited locations.
  - $S(t) \sim t^{\mu}$ indicates that the tendency to explore additional locations decreases with time.
- Preferential Return:
  - Random walk: the visitation probability is random and uniform in space
  - Human show significant propensity to return to the locations they visited frequently before.

EPR Model


Time: $t + \Delta t$
Δr
$P_{new} = \rho S^{-\gamma}$
Time: t
$S = 5$
$S = 4$
$P_{ret} = 1 - \rho S^{-\gamma}$
$\Pi_i = f_i$
f_3
f_2
f_1
$S = 4$

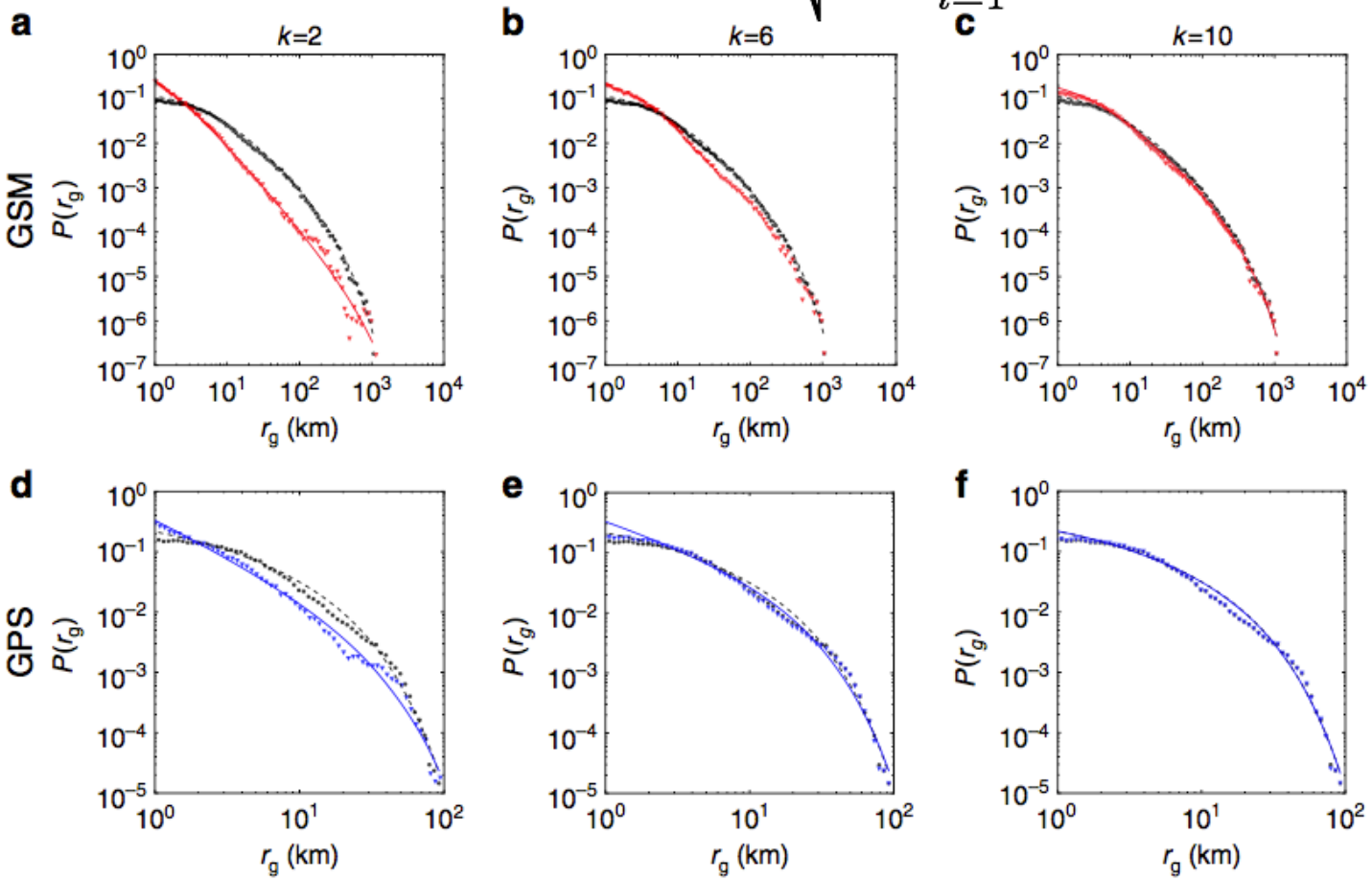
EPR Model


a
Empirical data
S(t)
Day
Week
Month
Year
r_g = 32 km
r_g = 64 km
r_g = 128 km
r_g = 256 km
~t^0.6
~t^0.8
t (h)
b
f_k
S = 20
S = 40
S = 60
k^-1.2
k
c
⟨Δx²⟩^α/2
S
r_g = 32 km
r_g = 64 km
r_g = 128 km
r_g = 256 km
Day
Week
Month
Year
t (h)
d
Individual-mobility model
S(t)
ρ = 0.1
ρ = 0.2
ρ = 0.5
β = 0.5
β = 0.6
β = 0.7
t
e
f_k
γ = 0, ρ = 0.5
γ = 0, ρ = 0.8
γ = 0.2, ρ = 0.1
γ = 0.5, ρ = 0.1
Democratic model
k
f
⟨Δx²⟩^α/2
ζ = 0
ζ = 0.9
ζ = 1.5
ζ = 2
S(t)

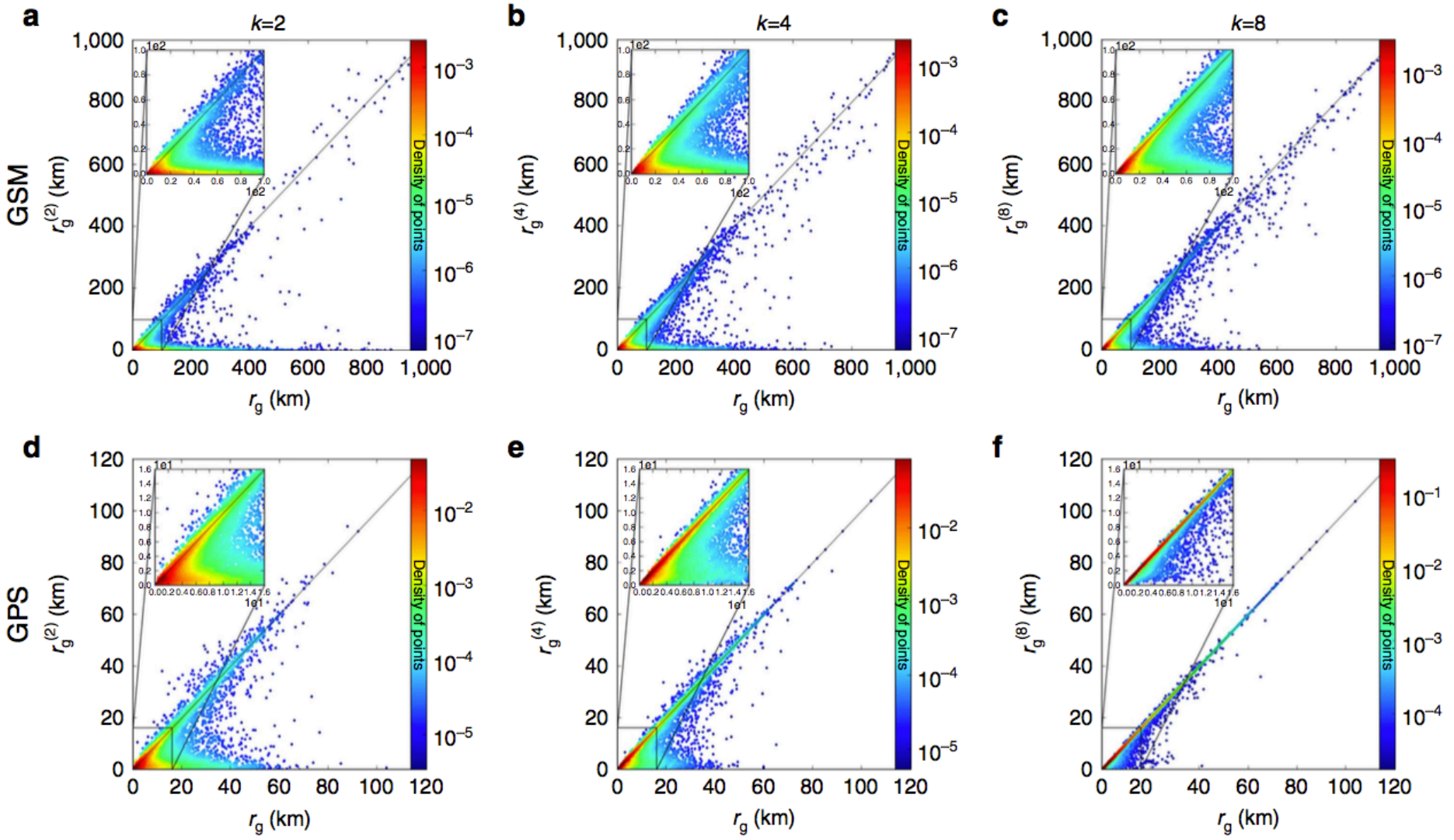
d-EPR Model

- Returners and explorers: well-separated
- k-radius of gyration:

$$r_g^{(k)} = \sqrt{\frac{1}{N_k}\sum_{i=1}^{k} n_i (r_i - r_{cm}^{(k)})^2}$$

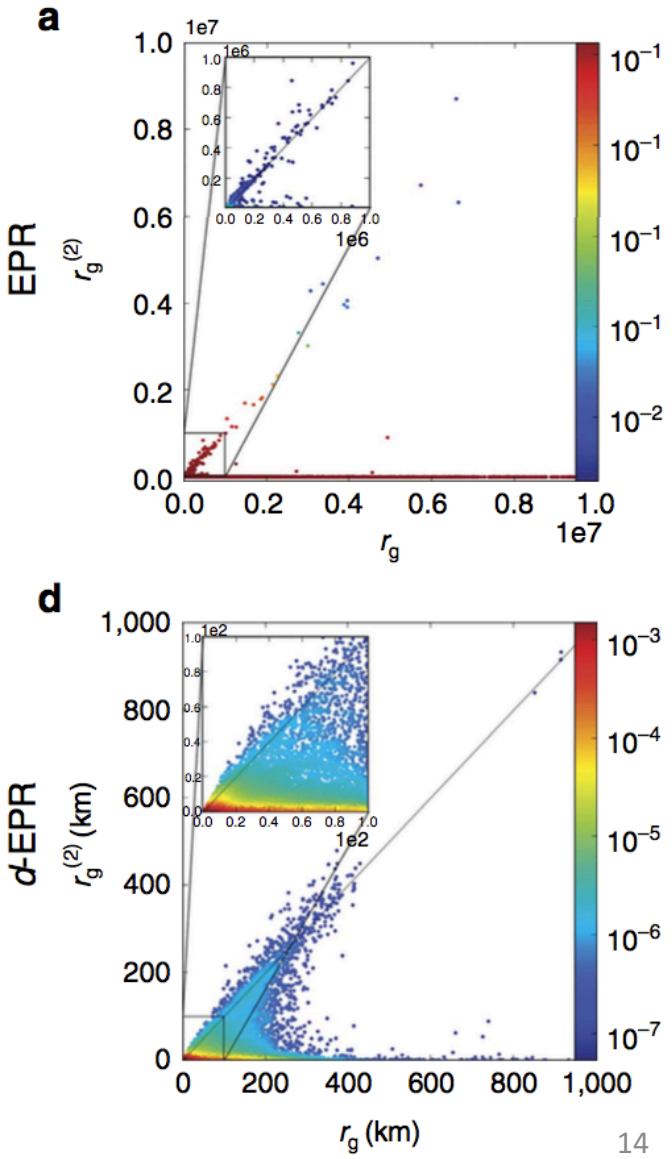


d-EPR Model


a
b
c
d
e
f
GSM
GPS
k=2
k=4
k=8
$r_g^{(2)}$ (km)
$r_g^{(4)}$ (km)
$r_g^{(8)}$ (km)
r_g (km)
Density of points

d-EPR Model

- EPR Model: Fail to reproduce
  - Reason: Individuals can travel arbitrarily large distances
  - i.e. select Δr from $P(\Delta r)$
- d-EPR Model:
  - Select a new location to visit depending on
  - 1. its distance Δr
  - 2. its relevance (the overall number of calls placed by all individuals from that location)
  - gravity model



Opportunities

- Cellular Data:
  - Finer granularity than CDR
  - Less sampling bias than GPS
- Ideas:
  - Difference between different granularity
  - Properties about trajectories
  - Less generality, more specificity
    - Time: periodicity
    - Space: urban functional zones
  - …