# Dynamics on and of subway networks

CAMILLE ROTH

*CNRS* (CAMS, CNRS / EHESS)

Séminaire LIP6 – 2 mars 2012

# Dynamics [ON] and [OF] subway networks

based on two 2011 papers co-authored with

**Marc Barthélémy**, **Michael Batty** et **Soong Kang**

## [ON]

- **“Structure of urban movements : polycentric activity and entangled hierarchical flows”**

  PLoS One, 6(1):e15923

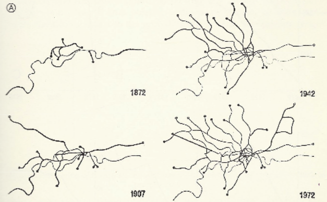



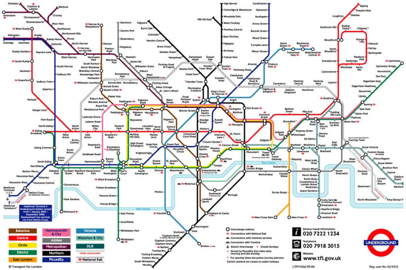

*Fig. 5.7.* Stages in the growth of the London underground rail network. Source: Cope, 1967.

## [OF]

- **“Long-time limit of world subway networks”**

  arXiv:1105.5294

# Dynamics [ON] subway networks





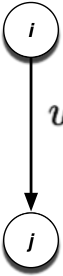
**flows of individuals :** proxy for activity and the city's structure

**data :** London subway

- Oyster card (anonymized) :
  $\rightarrow$ origin, destination, time
- 11m trips (1 week) for 2m individuals

# Origin-Destination Matrix

[ON]




$w_{ij}$ = number of rides from ***i*** to ***j***

Inflow of station ***i*** $= \sum_j w_{ji}$

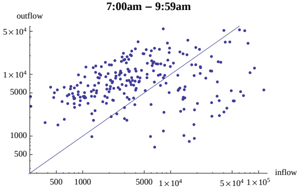
Outflow of station ***i*** $= \sum_j w_{ij}$

$\rightarrow$ complete, weighted graph...

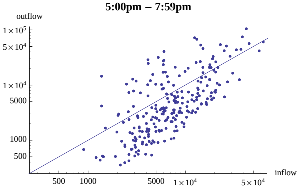
# Origin-Destination Matrix



## Temporal structure




**morning** : many sources, a few sinks



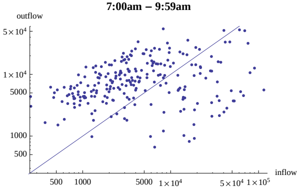

**evening** : a few sources, many sinks

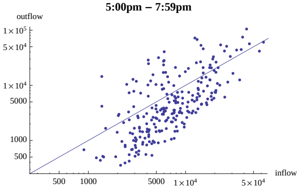
## Traffic distribution

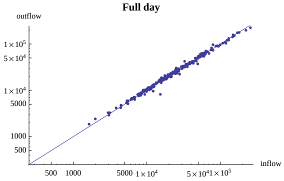
# Origin-Destination Matrix



## Temporal structure




**morning** : many sources, a few sinks

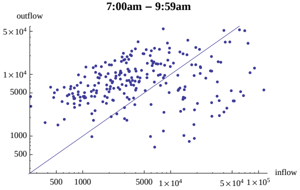
**evening** : a few sources, many sinks

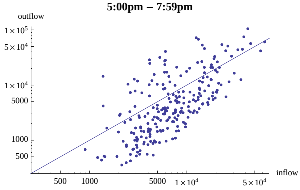
## Traffic distribution

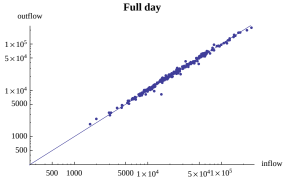
# Origin-Destination Matrix



## Temporal structure



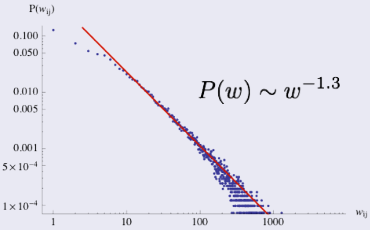




**morning** : many sources, a few sinks

**evening** : a few sources, many sinks

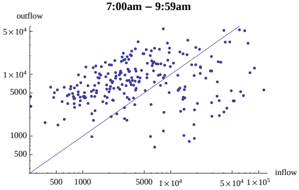
## Traffic distribution

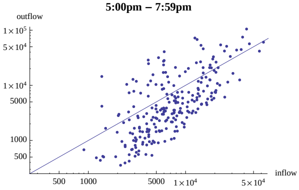


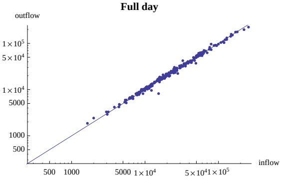
# Origin-Destination Matrix

[ON]

## Temporal structure



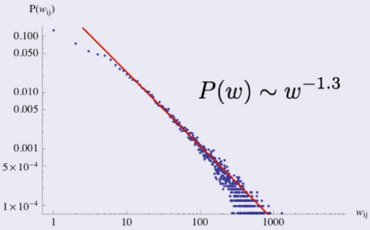



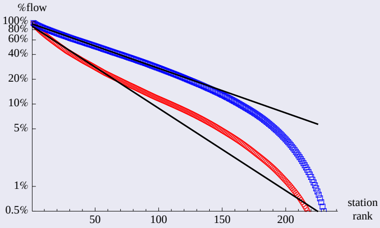




**morning** : many sources, a few sinks

**evening** : a few sources, many sinks

## Traffic distribution

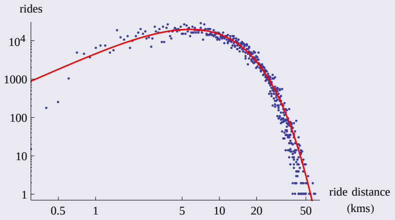


# Ride distribution



## Peaked distribution



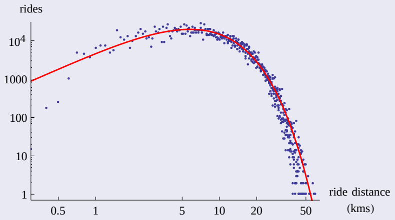

...as opposed to a typical law found in Brockman et al., 2006 (bank notes) and Gonzalez et al., 2008 (mobile phones) following :

$$P(d) \sim \frac{1}{d^\gamma} e^{-d/\kappa}$$

# Ride distribution



## Peaked distribution

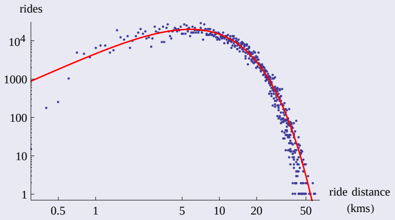



...as opposed to a typical law found in Brockman et al., 2006 (bank notes) and Gonzalez et al., 2008 (mobile phones) following :

$$P(d) \sim \frac{1}{d^{\gamma}} \mathbf{e}^{-d/\kappa}$$

# Ride distribution [ON]

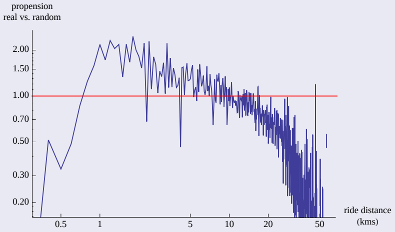
## Peaked distribution




...as opposed to a typical law found in Brockman et al., 2006 (bank notes) and Gonzalez et al., 2008 (mobile phones) following :

$$P(d) \sim \frac{1}{d^\gamma}\mathbf{e}^{-d/\kappa}$$

## Preference for short trips

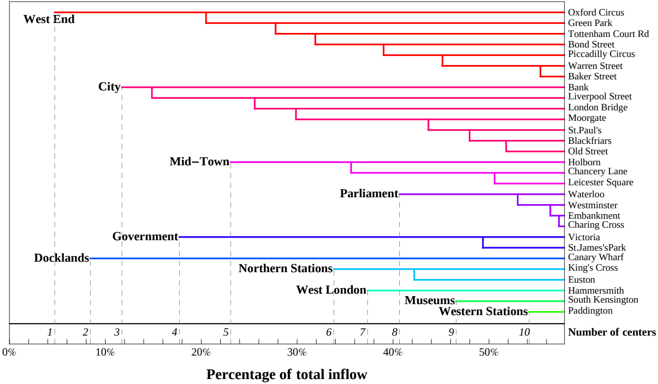


# Structure of centers



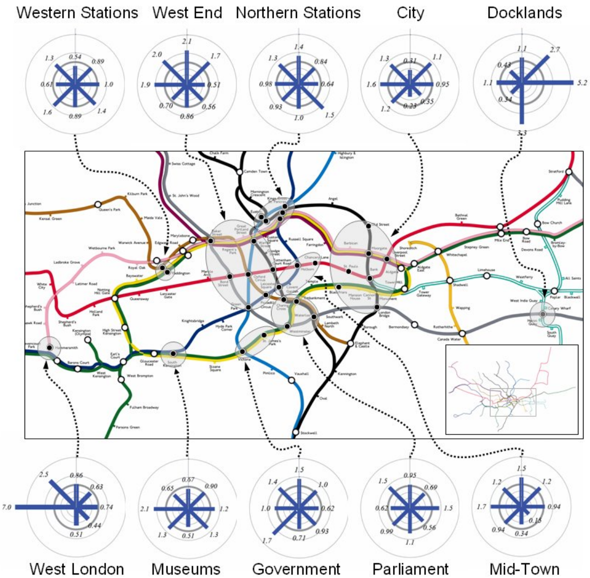
## Activity centers

- for each important **sink**, find a corresponding ‘**polycenter**’ in the vicinity
- rank stations by decreasing flow, group them geographically, stop when the total inflow reaches a certain value $W$ (here, $W = 60\%$)



# Structure of centers





# Structure of flows



How are the flows organized between sources and centers/sinks ?

- we consider flows from a station to a center/sink : $w_{iC} = \sum_{j \in C} w_{ij}$
- we sort them and explore the changes in the structure when going down the list

**Example map**

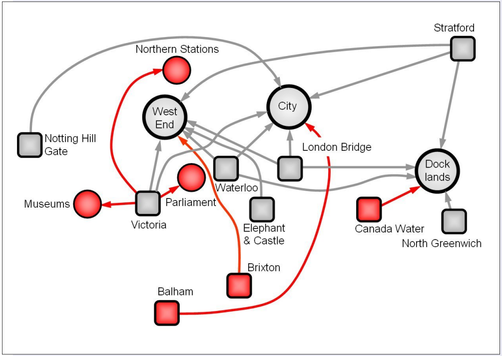
# Structure of flows



How are the flows organized between sources and centers/sinks ?

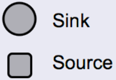
- we consider flows from a station to a center/sink : $w_{iC} = \sum_{j \in C} w_{ij}$
- we sort them and explore the changes in the structure when going down the list

**Example map**




When going from 20% (gray) of flow to 40% (red) :

1. largest sources go to smaller sinks,
2. smaller sources go to largest sinks.



# The transition matrix $T$ [ON]

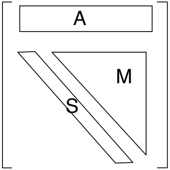
## $W \to W + \delta W$

- $T_{ij}$ : number of sources such that :

$$k_{\text{out}} = i(W) \quad \to \quad k_{\text{out}} = j(W + \delta W)$$

- $A_i$ : $\quad 0 \to i > 0$

T=



# The transition matrix $T$

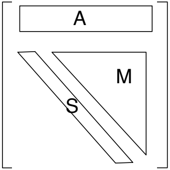


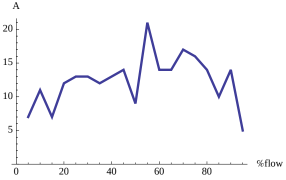
## $W \to W + \delta W$

- $T_{ij}$ : number of sources such that :

$$k_{\text{out}} = i(W) \quad \to \quad k_{\text{out}} = j(W + \delta W)$$

- $A_i$ : $\quad 0 \to i > 0$

T=


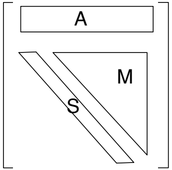
# The transition matrix $T$ [ON]

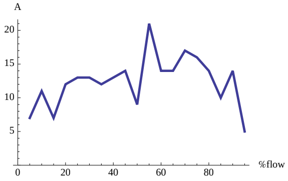
## $W \to W + \delta W$

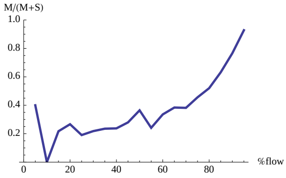
- $T_{ij}$ : number of sources such that :

$$k_{\text{out}} = i(W) \quad \to \quad k_{\text{out}} = j(W + \delta W)$$

- $A_i$ : $\quad 0 \to i > 0$

T=


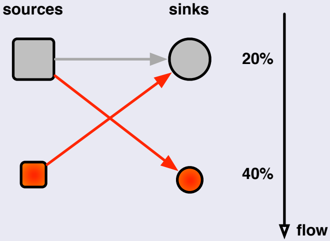



# Structure of flows



## Non-hierarchical structure

→ mixing of flows of different orders of magnitude



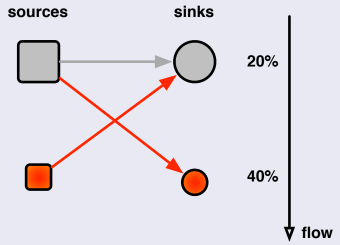

→ we divide polycenters in three groups according to their total inflow

# Structure of flows

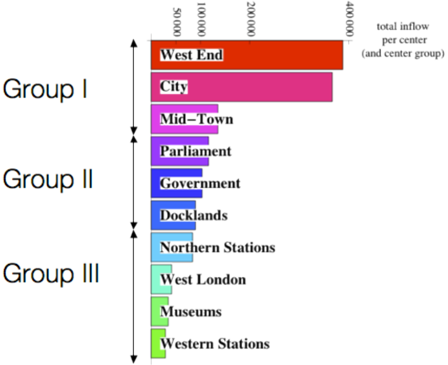


## Non-hierarchical structure

→ mixing of flows of different orders of magnitude




→ we divide polycenters in three groups according to their total inflow

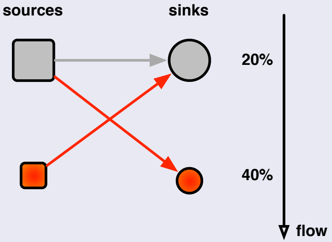


# Structure of flows

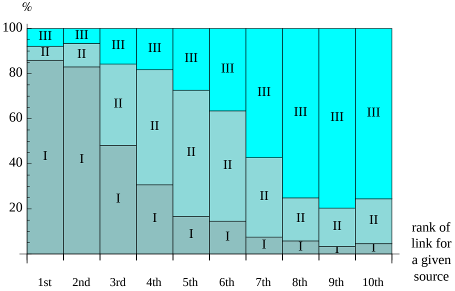


## Non-hierarchical structure

→ mixing of flows of different orders of magnitude




- to which groups do the important links of the sources go ?



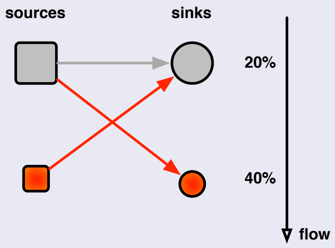

- for 80% of the sources the most important flow goes to group I
  → not optimized, congestion...

# Structure of flows

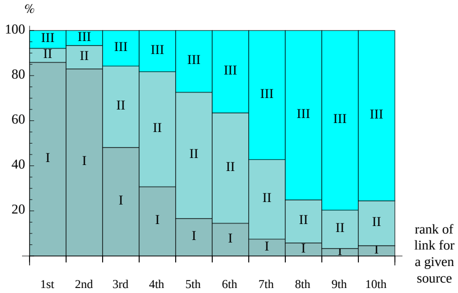
[ON]

## Non-hierarchical structure

→ mixing of flows of different orders of magnitude




- to which groups do the important links of the sources go ?




- for 80% of the sources the most important flow goes to group I
  → not optimized, congestion...

# Dynamics [ON] subway networks

## conclusions

- spatial features of the rides
  - consistence of recent results on ride/trip distribution ?
- spatial organization of activity
  - limited number of sources and sinks
  - organized in a hierarchical polycentric structure
    - → the notion of polycenter depends on the scale
- spatial organization of flows
  - iterative scheme
  - non hierarchical : most important flows always go to the same centers
    - → efficiency and congestion : London as a 'natural' city (as opposed to an 'artificial', optimized city)

# Dynamics [ON] subway networks

## conclusions

- spatial features of the rides
  - consistence of recent results on ride/trip distribution ?
- spatial organization of activity
  - limited number of sources and sinks
  - organized in a hierarchical polycentric structure
    - $\rightarrow$ the notion of polycenter depends on the scale
- spatial organization of flows
  - iterative scheme
  - non hierarchical : most important flows always go to the same centers
    - $\Rightarrow$ efficiency and congestion : London as a 'natural' city (as opposed to an 'artificial', optimized city)

# Dynamics [ON] subway networks

## conclusions

- spatial features of the rides
  - consistence of recent results on ride/trip distribution ?
- spatial organization of activity
  - limited number of sources and sinks
  - organized in a hierarchical polycentric structure
    - → the notion of polycenter depends on the scale
- spatial organization of flows
  - iterative scheme
  - non hierarchical : most important flows always go to the same centers
    - ⇒ efficiency and congestion : London as a ‘natural’ city (as opposed to an ‘artificial’, optimized city)

# Dynamics [OF] subway networks

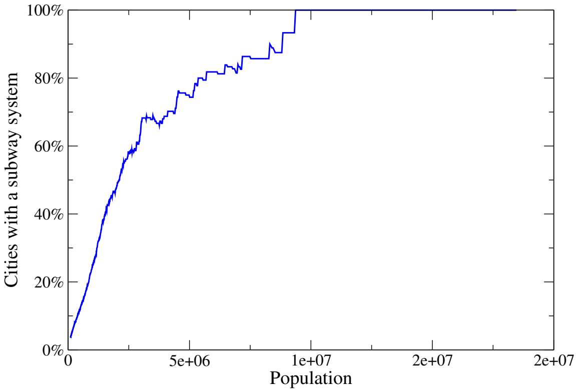



**FIGURE:** Percentage of cities with a subway system versus the population (UN data) [exponential fit : $f \sim 1 - \exp(-P/P_0)$ with $P_0 \approx 3$m].

# Dynamics [OF] subway networks

## specific issues

- spatial networks
  - nodes are located in space
  - network embedded in a space equipped with some metric...

    → with a 'cost' associated to the length of links : a long link must be compensated by something else (large degree, traffic, etc.)
- time evolving
  - new stations, new lines.
  - increase of the spatial extension.
- network-based measures and tools
  - typical "scale-free" measures are not immediately relevant
  - usual measures for planar networks

# Dynamics [OF] subway networks

## specific issues

- spatial networks
  - nodes are located in space
  - network embedded in a space equipped with some metric...

    $\rightarrow$ with a 'cost' associated to the length of links : a long link must be compensated by something else (large degree, traffic, etc.)
- time evolving
  - new stations, new lines.
  - increase of the spatial extension.
- network-based measures and tools
  - typical "scale-free" measures are not immediately relevant
  - usual measures for planar networks

# Dynamics [OF] subway networks

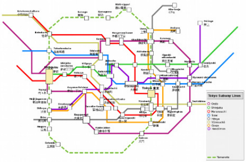
## specific issues

- spatial networks
  - nodes are located in space
  - network embedded in a space equipped with some metric...

    $\rightarrow$ with a 'cost' associated to the length of links : a long link must be compensated by something else (large degree, traffic, etc.)
- time evolving
  - new stations, new lines.
  - increase of the spatial extension.
- network-based measures and tools
  - typical "scale-free" measures are not immediately relevant
  - usual measures for planar networks

# Spatial features

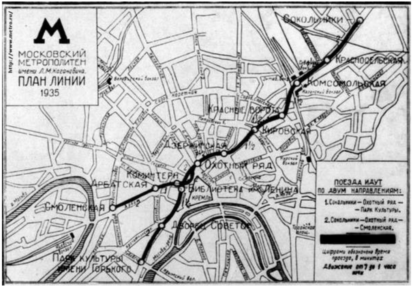
[OF]

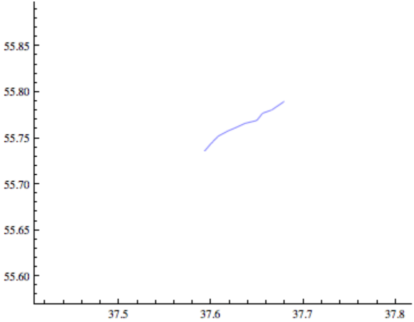





# Evolution

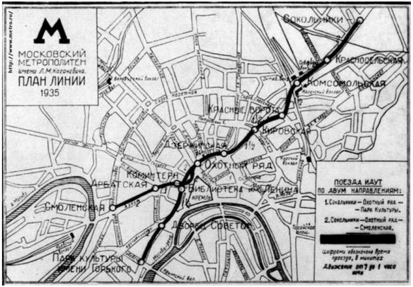
[OF]

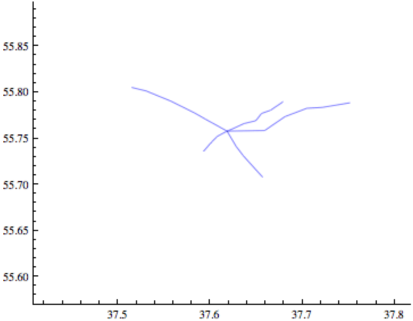



Moscow metro 1936

# Evolution [OF]

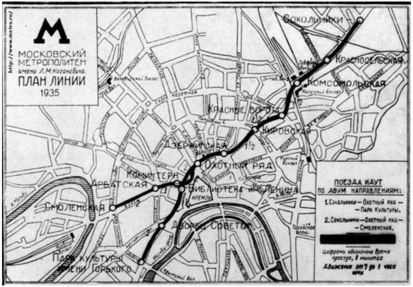


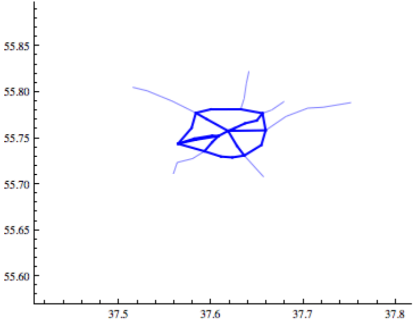




Moscow metro 1950

# Evolution



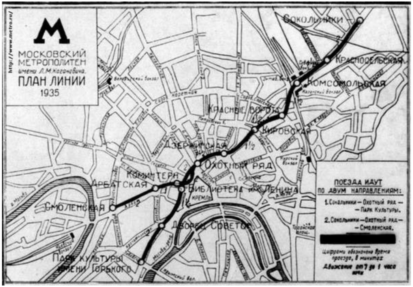


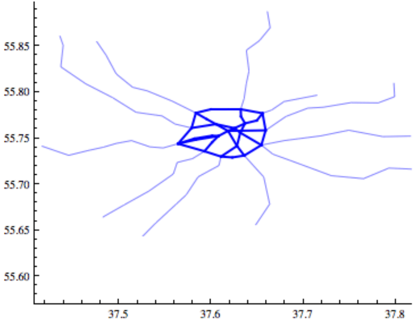




Moscow metro 1959

# Evolution



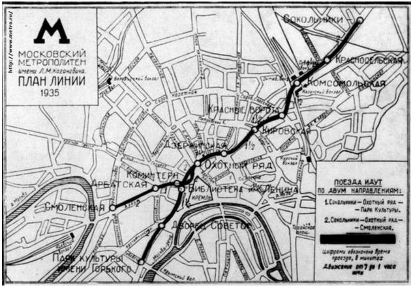


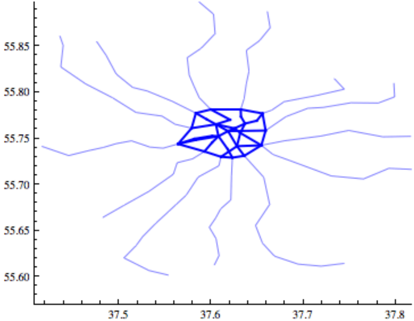




Moscow metro 1981

# Evolution

[OF]




Moscow metro 1995

# Evolution

## Some simple properties

| City | $t_0$ | $\overline{v}$ | $\sigma_v$ | $f$ |
|---|---|---|---|---|
| Beijing | 1971 | 3.34 | 7.74 | 79% |
| Tokyo | 1927 | 1.8 | 3.4 | 58% |
| Seoul | 1974 | 11.2 | 14.9 | 20% |
| Mexico | 1969 | 3.7 | 5.9 | 55% |
| New York city | 1878 | 3.3 | 8.3 | 68% |
| Shanghai | 1995 | 14.9 | 20.2 | 31% |
| Moscow | 1936 | 1.7 | 1.9 | 43% |
| Chicago | 1901 | 1.91 | 6.24 | 71% |
| London | 1863 | 2.3 | 3.8 | 48% |
| Paris | 1900 | 2.6 | 5.1 | 60% |
| Madrid | 1919 | 2.3 | 4.6 | 59% |
| Berlin | 1901 | 1.6 | 3.3 | 65% |
| Barcelona | 1914 | 1.4 | 4.8 | 78% |
| Osaka | 1934 | 1.4 | 4.1 | 79% |

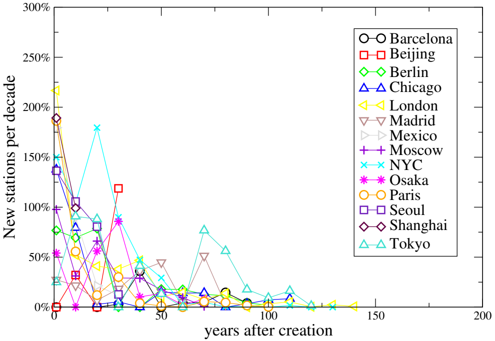
**TABLE:** $t_0$ is the initial year considered here for the different subways networks. $\overline{v}$ is the average velocity (number of stations built per year), $\sigma_V$ is the standard deviation of $v$, and $f$ is the fraction of years of inactivity (no stations built).

(i) large velocities at small times ; (ii) large fluctuations from year to year ; (iii) $f$ on average 60%

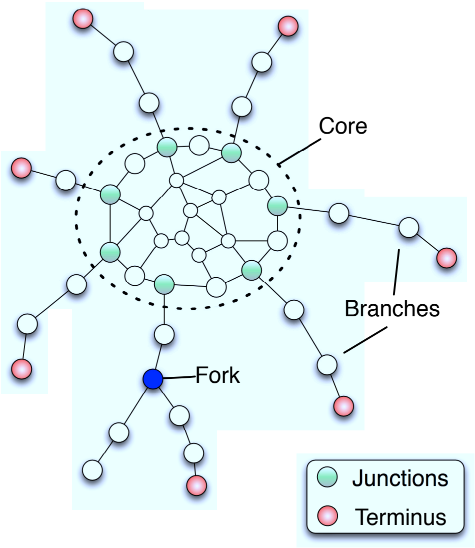
# Evolution



- bursty growth of the number of stations, with periods of inactivity
- convergence to a stationary state



# Universal template [OF]




simple algorithm :

- start from line tails
- progress along branch as long as node degree is 2
- ignore "simple forks"

# Universal template



## simple measures

- **fraction of branch nodes**

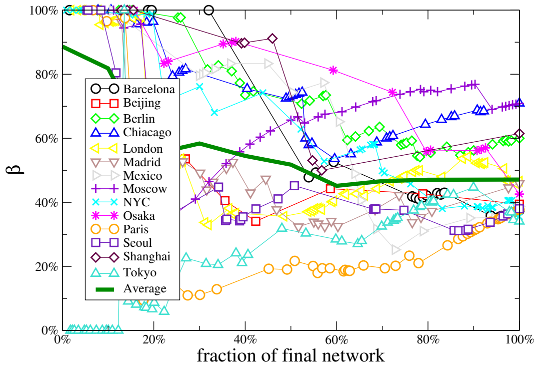
$$\beta = \frac{N_{\text{branches}}}{N_{\text{branches}+N_{\text{core}}}}$$

- avg. core degree

$$\langle k_{\text{core}} \rangle = 2\frac{E_{\text{core}}}{N_{\text{core}}}$$

- spatial extension

$$\eta = \frac{\bar{D}_{\text{branches}}}{\bar{D}_{\text{core}}}$$



# Universal template [OF]

## simple measures

- **fraction of branch nodes**

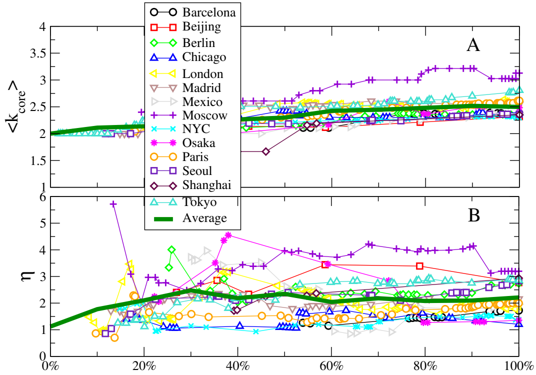
$$\beta = \frac{N_{\text{branches}}}{N_{\text{branches}}+N_{\text{core}}}$$

- **avg. core degree**

$$\langle k_{\text{core}} \rangle = 2\frac{E_{\text{core}}}{N_{\text{core}}}$$

- **spatial extension**

$$\eta = \frac{\bar{D}_{\text{branches}}}{\bar{D}_{\text{core}}}$$



# Universal template



## simple measures

- **fraction of branch nodes**

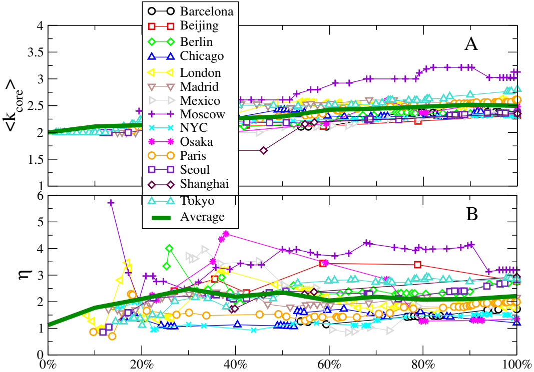
$$\beta = \frac{N_{\text{branches}}}{N_{\text{branches}+N_{\text{core}}}}$$

- **avg. core degree**

$$\langle k_{\text{core}} \rangle = 2\frac{E_{\text{core}}}{N_{\text{core}}}$$

- **spatial extension**

$$\eta = \frac{\bar{D}_{\text{branches}}}{\bar{D}_{\text{core}}}$$




→ limiting shape :

- **branch vs. core** is relatively constant, in terms of number of stations and spatial extension
- continuing, **slow densification** of the core

# Scaling core/branches



- old result for Paris
  (Benguigui & Daoud 1991)
  short scale, $N(r) \sim r^2$,
  long scale, $N(r) \sim r^{.5}$
- reinterpret it in terms of "branch/core"
  defining $r_{\mathrm{core}}$ as
  $N(r = r_{\mathrm{core}}) = N_{\mathrm{core}}$

| City | $N_{\mathrm{core}}$ | $r_{\mathrm{core}}$ (kms) |
|---|---|---|
| Beijing | 63 | 4.4 |
| Tokyo | 151 | 6.4 |
| Seoul | 243 | 11.6 |
| Mexico | 90 | 4.7 |
| Shanghai | 57 | 3.7 |
| Moscow | 39 | 5.9 |
| London | 142 | 7.3 |
| Paris | 186 | 4.2 |
| Madrid | 113 | 4.4 |
| Berlin | 68 | 5.5 |
| Barcelona | 57 | 3.5 |
| Osaka | 46 | 3.6 |

# Scaling core/branches



- old result for Paris
  (Benguigui & Daoud 1991)
  short scale, $N(r) \sim r^2$,
  long scale, $N(r) \sim r^{.5}$
- reinterpret it in terms of "branch/core"
  defining $r_{\text{core}}$ as
  $N(r = r_{\text{core}}) = N_{\text{core}}$

| City | $N_{\text{core}}$ | $r_{\text{core}}$ (kms) |
|---|---|---|
| Beijing | 63 | 4.4 |
| Tokyo | 151 | 6.4 |
| Seoul | 243 | 11.6 |
| Mexico | 90 | 4.7 |
| Shanghai | 57 | 3.7 |
| Moscow | 39 | 5.9 |
| London | 142 | 7.3 |
| Paris | 186 | 4.2 |
| Madrid | 113 | 4.4 |
| Berlin | 68 | 5.5 |
| Barcelona | 57 | 3.5 |
| Osaka | 46 | 3.6 |

# Scaling core/branches



- old result for Paris
  (Benguigui & Daoud 1991)
  short scale, $N(r) \sim r^2$,
  long scale, $N(r) \sim r^{.5}$
- reinterpret it in terms of "branch/core"
  defining $r_{\text{core}}$ as
  $N(r = r_{\text{core}}) = N_{\text{core}}$

| City | $N_{\text{core}}$ | $r_{\text{core}}$ (kms) |
|---|---|---|
| Beijing | 63 | 4.4 |
| Tokyo | 151 | 6.4 |
| Seoul | 243 | 11.6 |
| Mexico | 90 | 4.7 |
| Shanghai | 57 | 3.7 |
| Moscow | 39 | 5.9 |
| London | 142 | 7.3 |
| Paris | 186 | 4.2 |
| Madrid | 113 | 4.4 |
| Berlin | 68 | 5.5 |
| Barcelona | 57 | 3.5 |
| Osaka | 46 | 3.6 |

# Scaling core/branches

[OF]

- old result for Paris

  (Benguigui & Daoud 1991)

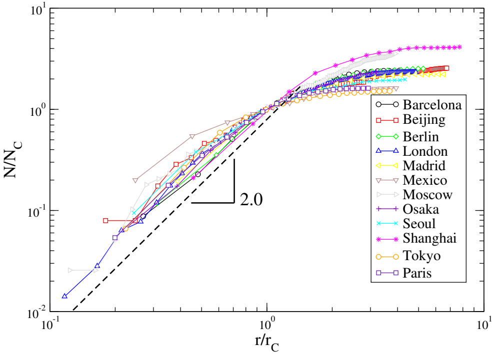
  short scale, $N(r) \sim r^2$,

  long scale, $N(r) \sim r^{.5}$

- reinterpret it in terms of “branch/core”

  defining $r_{core}$ as

  $N(r = r_{core}) = N_{core}$



# Scaling core/branches



- old result for Paris

  (Benguigui & Daoud 1991)

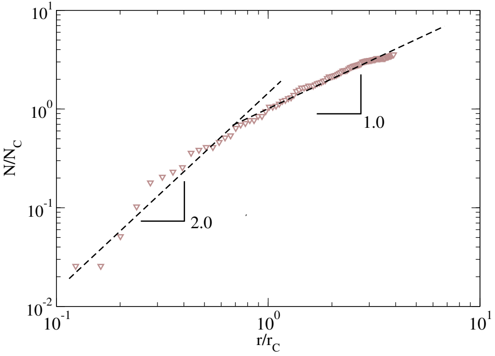
  short scale, $N(r) \sim r^2$,

  long scale, $N(r) \sim r^{.5}$

- reinterpret it in terms of "branch/core"

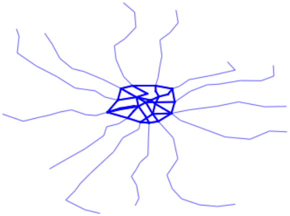
  defining $r_{core}$ as

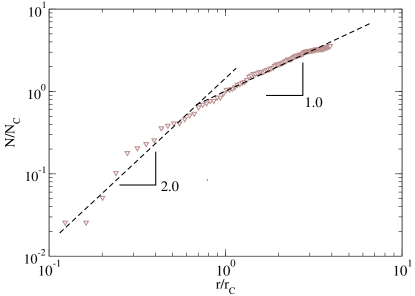
  $N(r = r_{core}) = N_{core}$



# Scaling core/branches

[OF]





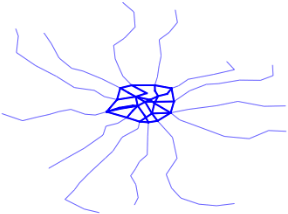

$$N(r) \sim \begin{cases} \rho_C \pi r^2 & \text{for} \quad r \ll r_{core} \\ \rho_C \pi r_{core}^2 + N_{branches} \int_{r_{core}}^{r} \frac{1}{\Delta(r)} dr & \text{for} \quad r_{core} \ll r < r_{max} \end{cases}$$

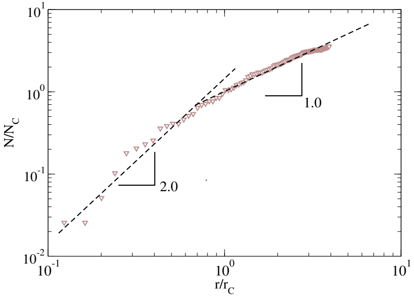
$$\Delta(r) \sim r^{\tau} \Rightarrow N(r \gg r_{core}) \sim r^{1-\tau}$$

$\tau \simeq 0.05$ (Moscow), $\tau \simeq 0.5$ (Paris)

# Scaling core/branches








$$N(r) \sim \begin{cases} \rho_C \pi r^2 & \text{for} \quad r \ll r_{core} \\ \rho_C \pi r_{core}^2 + N_{branches} \displaystyle\int_{r_{core}}^{r} \frac{1}{\Delta(r)} dr & \text{for} \quad r_{core} \ll r < r_{max} \end{cases}$$

$$\Delta(r) \sim r^\tau \Rightarrow N(r \gg r_{core}) \sim r^{1-\tau}$$

$\tau \simeq 0.05$ (Moscow), $\tau \simeq 0.5$ (Paris)