# Optimal supply & Structure detection Santa Fe Institute Summer School, 2009

### Prof. Peter Dodds

Department of Mathematics & Statistics Center for Complex Systems Vermont Advanced Computing Center University of Vermont











#### Single Source

#### Distributed Sources

Facility leastion

Ciza danaitu law

A reasonable derivation Global redistribution networks

#### Structure Detection

etection

Hierarchy by division
Hierarchy by shuffling

Spectral methods Hierarchies & Missing Links

General structure detection

Final words

Deference

References





### Single Source

### Distributed Sources

Facility location

Size-density law

A reasonable derivation

Global redistribution networks

### Structure Detection

Hierarchy by division

Hierarchy by shuffling

Spectral methods

Hierarchies & Missing Links

General structure detection

### Final words

### References

### Single Source

#### Distributed Sources

### Structure

Detection





Structure Detection

### What's the best way to distribute stuff?

- Stuff = medical services, energy, nutrients, people, ...
- Some fundamental network problems:
- Q: How do optimal solutions scale with system size?



- Stuff = medical services, energy, nutrients, people, ...
- Q: How do optimal solutions scale with system size?



- Stuff = medical services, energy, nutrients, people, ...
- Some fundamental network problems:
- Q: How do optimal solutions scale with system size?



- Stuff = medical services, energy, nutrients, people, ...
- Some fundamental network problems:
  - Distribute stuff from single source to many sinks
- Q: How do optimal solutions scale with system size?



- Stuff = medical services, energy, nutrients, people, ...
- Some fundamental network problems:
  - Distribute stuff from single source to many sinks
  - 2. Collect stuff coming from many sources at a single sink
- Q: How do optimal solutions scale with system size?



Frame 3/78



## What's the best way to distribute stuff?

- Stuff = medical services, energy, nutrients, people, ...
- Some fundamental network problems:
  - Distribute stuff from single source to many sinks
  - 2. Collect stuff coming from many sources at a single sink
  - 3. Distribute stuff from many sources to many sinks
- Q: How do optimal solutions scale with system size?

## Sources

Detection

Frame 3/78



## What's the best way to distribute stuff?

- Stuff = medical services, energy, nutrients, people, ...
- Some fundamental network problems:
  - Distribute stuff from single source to many sinks
  - 2. Collect stuff coming from many sources at a single sink
  - 3. Distribute stuff from many sources to many sinks
  - 4. Redistribute stuff between many nodes
- Q: How do optimal solutions scale with system size?

## Detection

Frame 3/78



## What's the best way to distribute stuff?

- Stuff = medical services, energy, nutrients, people, ...
- Some fundamental network problems:
  - Distribute stuff from single source to many sinks
  - 2. Collect stuff coming from many sources at a single sink
  - 3. Distribute stuff from many sources to many sinks
  - 4. Redistribute stuff between many nodes
- Q: How do optimal solutions scale with system size?

### Basic Q for distribution/supply networks:

How does flow behave given cost:

$$C = \sum_{j} I_{j}^{\gamma} Z_{j}$$

where  $I_i$  = current on link iand  $Z_i = link j$ 's impedance?

 $\triangleright$  Example:  $\gamma = 2$  for electrical networks.

### Single Source

#### Distributed Sources

#### Structure Detection

Frame 4/78



### Basic Q for distribution/supply networks:

How does flow behave given cost:

$$C = \sum_{j} I_{j}^{\gamma} Z_{j}$$

where  $I_i$  = current on link iand  $Z_i = link j$ 's impedance?

**Example:**  $\gamma = 2$  for electrical networks.

### Single Source

#### Distributed Sources

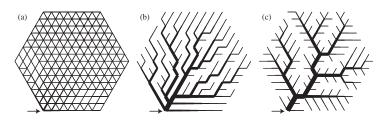
### Structure

Detection

Frame 4/78



## Single source optimal supply



(a)  $\gamma > 1$ : Braided (bulk) flow

(b)  $\gamma$  < 1: Local minimum: Branching flow

(c)  $\gamma$  < 1: Global minimum: Branching flow

From Bohn and Magnasco [3] See also Banavar et al. [1]

#### Optimal supply & Structure detection

### Single Source

## Distributed Sources

Facility location

A reasonable derivation Global redistribution networks

### Structure Detection

lierarchy by division

Spectral methods
Hierarchies & Missing Links

inal ...auda

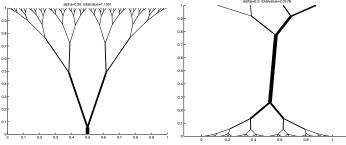
### nal words

References

Frame 5/78



## Optimal paths related to transport (Monge) problems:



Xia (2003) [24]

### Single Source

### Distributed Sources

acility location

A reasonable derivation Global redistribution networks

### Structure

Detection

Hierarchy by shuffling

Hierarchies & Missing Links

Final words

References

Frame 6/78



#### Optimal supply & Structure detection

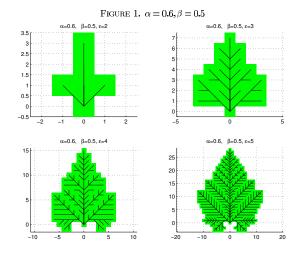
### Single Source

### Distributed Sources

### Detection

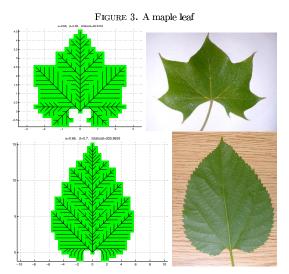
Frame 7/78





Xia (2007) [23]

## Growing networks:



Single Source

### Distributed Sources

Facility location

A reasonable derivation Global redistribution networks

#### Structure Detection

erarchy by division ierarchy by shuffling

Spectral methods Hierarchies & Missing Link

General structure dete

Final words

References

Frame 8/78



A reasonable derivation Global redistribution networks

Structure Detection

Hierarchy by division
Hierarchy by shuffling

Spectral methods Hierarchies & Missing Links

Eta al consula

References

### Frame 9/78



### An immensely controversial issue...

► The form of river networks and blood networks: optimal or not? [22, 2, 7]

### Two observations:

- Self-similar networks appear everywhere in nature for single source supply/single sink collection.
- Real networks differ in details of scaling but reasonably agree in scaling relations.

## An immensely controversial issue...

► The form of river networks and blood networks: optimal or not? [22, 2, 7]

### Two observations:

- Self-similar networks appear everywhere in nature for single source supply/single sink collection.
- Real networks differ in details of scaling but reasonably agree in scaling relations.

### Single Source

### Distributed Sources

Size-density law

A reasonable derivation Global redistribution networks

#### Structure Detection

Hierarchy by shuffling Spectral methods Hierarchies & Missing Links

eneral structure detection

#### inal words

References

Frame 9/78



Structure
Detection
Hierarchy by division

Hierarchy by shuffling Spectral methods Hierarchies & Missing Links

General structure detec

Final words

References

nelelelices

### An immensely controversial issue...

► The form of river networks and blood networks: optimal or not? [22, 2, 7]

### Two observations:

- Self-similar networks appear everywhere in nature for single source supply/single sink collection.
- ► Real networks differ in details of scaling but reasonably agree in scaling relations.

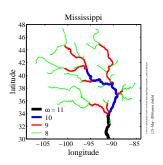
Frame 9/78



- ▶ Label all source streams as order  $\omega = 1$ .
- Follow all labelled streams downstream
- ▶ Whenever two streams of the same order  $(\omega)$  meet, the resulting stream has order incremented by 1  $(\omega + 1)$ .
- If streams of different orders  $\omega_1$  and  $\omega_2$  meet, then the resultant stream has order equal to the largest of the two
- Simple rule:

$$\omega_3 = \max(\omega_1, \omega_2) + \delta_{\omega_1, \omega_2}$$

where  $\delta$  is the Kronecker delta



### Single Source

## Distributed Sources

Facility location
Size-density law
A reasonable derivation

#### Structure Detection

Hierarchy by division Hierarchy by shuffling Spectral methods Hierarchies & Missing Links

inal words

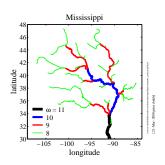
References



- ▶ Label all source streams as order  $\omega = 1$ .
- Follow all labelled streams downstream
- ▶ Whenever two streams of the same order  $(\omega)$  meet, the resulting stream has order incremented by 1  $(\omega + 1)$ .
- If streams of different orders  $\omega_1$  and  $\omega_2$  meet, then the resultant stream has order equal to the largest of the two
- ▶ Simple rule:

$$\omega_3 = \max(\omega_1, \omega_2) + \delta_{\omega_1, \omega_2}$$

where  $\delta$  is the Kronecker delta



### Single Source

### Distributed Sources

Facility location
Size-density law
A reasonable derivation

#### Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

inal words

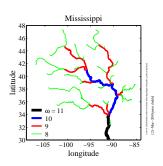
References



- ▶ Label all source streams as order  $\omega = 1$ .
- Follow all labelled streams downstream
- ▶ Whenever two streams of the same order  $(\omega)$  meet, the resulting stream has order incremented by 1  $(\omega + 1)$ .
- If streams of different orders  $\omega_1$  and  $\omega_2$  meet, then the resultant stream has order equal to the largest of the two
- ▶ Simple rule:

$$\omega_3 = \max(\omega_1, \omega_2) + \delta_{\omega_1, \omega_2}$$

where  $\delta$  is the Kronecker delta



### Single Source

### Distributed Sources

Size-density law
A reasonable derivation

#### Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

inal words

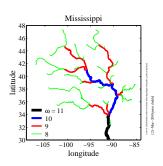
References



- ▶ Label all source streams as order  $\omega = 1$ .
- Follow all labelled streams downstream
- ▶ Whenever two streams of the same order  $(\omega)$  meet, the resulting stream has order incremented by 1  $(\omega + 1)$ .
- If streams of different orders  $\omega_1$  and  $\omega_2$  meet, then the resultant stream has order equal to the largest of the two.
- Simple rule:

$$\omega_3 = \max(\omega_1, \omega_2) + \delta_{\omega_1, \omega_2}$$

where  $\delta$  is the Kronecker delta.



### Single Source

### Distributed Sources

Facility location
Size-density law
A reasonable derivation

### Structure Detection

Hierarchy by division Hierarchy by shuffling Spectral methods Hierarchies & Missing Links

### inal words

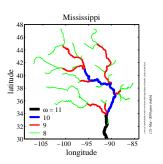
References



- ▶ Label all source streams as order  $\omega = 1$ .
- Follow all labelled streams downstream
- ▶ Whenever two streams of the same order  $(\omega)$  meet, the resulting stream has order incremented by 1  $(\omega + 1)$ .
- If streams of different orders  $\omega_1$  and  $\omega_2$  meet, then the resultant stream has order equal to the largest of the two.
- Simple rule:

$$\omega_3 = \max(\omega_1, \omega_2) + \delta_{\omega_1, \omega_2}$$

where  $\delta$  is the Kronecker delta.



### Single Source

### Distributed Sources

Size-density law
A reasonable derivation

#### Structure Detection

Hierarchy by division Hierarchy by shuffling Spectral methods Hierarchies & Missing Links

inal words

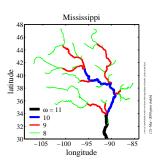
References



- ▶ Label all source streams as order  $\omega = 1$ .
- Follow all labelled streams downstream
- ▶ Whenever two streams of the same order  $(\omega)$  meet, the resulting stream has order incremented by 1  $(\omega + 1)$ .
- If streams of different orders  $\omega_1$  and  $\omega_2$  meet, then the resultant stream has order equal to the largest of the two.
- Simple rule:

$$\omega_3 = \max(\omega_1, \omega_2) + \delta_{\omega_1, \omega_2}$$

where  $\delta$  is the Kronecker delta.



Single Source

Distributed Sources

Size-density law
A reasonable derivation

Global redistribution networks

Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Link

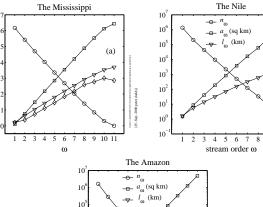
nal words

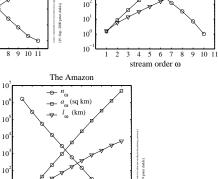
References



### Horton's laws in the real world:

10<sup>1</sup> 10





8 9 10 11

stream order ω

#### Optimal supply & Structure detection

### Single Source

### Distributed Sources

## Detection

Frame 11/78



relation:	scaling relation/parameter: [6]
$\ell \sim {\sf L}^{\sf d}$	d
$T_k = T_1(R_T)^{k-1}$	$T_1 = R_n - R_s - 2 + 2R_s/R_n$
	$R_T = R_s$
$n_{\omega}/n_{\omega+1}=R_n$	$R_n$
$ar{a}_{\omega+1}/ar{a}_{\omega}=R_a$	$R_a = \frac{R_n}{R_n}$
$ar{\ell}_{\omega+1}/ar{\ell}_{\omega}= extbf{ extit{R}}_{\ell}$	$R_\ell = R_s$
$\ell \sim a^h$	$h = \log \frac{R_s}{\log R_n}$
$oldsymbol{a} \sim oldsymbol{\mathcal{L}}^{oldsymbol{D}}$	D = d/h
${m L}_{\perp} \sim {m L}^{m H}$	H = d/h - 1
$P(a) \sim a^{- au}$	au = 2 - h
$P(\ell) \sim \ell^{-\gamma}$	$\gamma = 1/h$
$oldsymbol{\wedge} \sim oldsymbol{a}^eta$	$\beta = 1 + h$
$\lambda \sim \mathcal{L}^{arphi}$	$arphi= extsf{d}$

Single Source

Distributed Sources

racility localion Size-density law A reasonable derivation

Global redistribution networks

Structure Detection

Hierarchy by shuffling Spectral methods Hierarchies & Missing Links

Hierarchies & Missing Links General structure detection

Final words

References

Only 3 parameters are independent... [6]

Frame 12/78



Parameter:	Real networks:
$R_n$	3.0-5.0
$R_a$	3.0-6.0
$R_\ell = R_T$	1.5–3.0
$T_1$	1.0–1.5
d	$1.1\pm0.01$
D	$1.8\pm 0.1$
h	0.50-0.70
au	$\boldsymbol{1.43 \pm 0.05}$
$\gamma$	$1.8 \pm 0.1$
Н	0.75-0.80
$\beta$	0.50-0.70
arphi	$\boldsymbol{1.05 \pm 0.05}$

### Single Source

### Distributed Sources

Detection





### Data from real blood networks

Network	R <sub>n</sub>	$R_r^{-1}$	$R_\ell^{-1}$	$-\frac{\ln R_r}{\ln R_n}$	$-rac{\ln R_\ell}{\ln R_n}$	$\alpha$
West et al.	_	_	-	0.5	$0.3\bar{3}$	0.75
rat (PAT)	2.76	1.58	1.60	0.45	0.46	0.73
cat (PAT)	3.67	1.71	1.78	0.41	0.44	0.79
(Turcotte et al. [21])						
dog (PAT)	3.69	1.67	1.52	0.39	0.32	0.90
pig (LCX)	3.57	1.89	2.20	0.50	0.62	0.62
pig (RCA)	3.50	1.81	2.12	0.47	0.60	0.65
pig (LAD)	3.51	1.84	2.02	0.49	0.56	0.65
human (PAT)	3.03	1.60	1.49	0.42	0.36	0.83
human (PAT)	3.36	1.56	1.49	0.37	0.33	0.94

### Single Source

## Distributed

Frame 14/78



### Fundamental biological and ecological constraint:

$$P = c M^{\alpha}$$

P =basal metabolic rate

M =organismal body mass





### Single Source

### Distributed Sources

Facility location

A reasonable derivation Global redistribution networks

### Structure

Detection

Hierarchy by shuffling
Spectral methods

Hierarchies & Missing Links General structure detection

### inal words

References

Frame 15/78



### Single Source

### Distributed Sources

### Detection





 $P = c M^{\alpha}$ 

P =basal metabolic rate

Fundamental biological and ecological constraint:



Frame 15/78



Detection

Frame 16/78



1964: Troon, Scotland: 3rd symposium on energy metabolism.  $\alpha =$  3/4 made official . . .



1964: Troon, Scotland:

 $\alpha = 3/4$  made official . . .

3rd symposium on energy metabolism.

### Single Source

### Distributed Sources

### Detection

... 29 to zip.



Frame 16/78





### Distributed Sources

acility location

A reasonable derivation Global redistribution

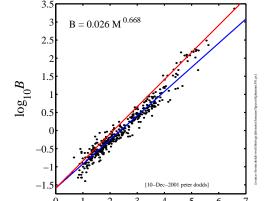
#### Structure Detection

Hierarchy by division
Hierarchy by shuffling

Hierarchies & Missing Links

Final words

References



 $\log_{10} M$ 

- Heusner's data (1991) [11]
- ▶ 391 Mammals
- ▶ blue line: 2/3
- ▶ red line: 3/4.
- ► (*B* = *P*)

Frame 17/78



## Some regressions from the ground up...

range of <i>M</i>	Ν	$\hat{lpha}$
$\leq$ 0.1 kg	167	$\boldsymbol{0.678 \pm 0.038}$
$\leq$ 1 kg	276	$\boldsymbol{0.662 \pm 0.032}$
$\leq$ 10 kg	357	$\textbf{0.668} \pm \textbf{0.019}$
$\leq$ 25 kg	366	$\boldsymbol{0.669 \pm 0.018}$
$\leq$ 35 kg	371	$\textbf{0.675} \pm \textbf{0.018}$
$\leq$ 350 kg	389	$\boldsymbol{0.706 \pm 0.016}$
$\leq$ 3670 kg	391	$0.710 \pm 0.021$

### Single Source

#### Distributed Sources

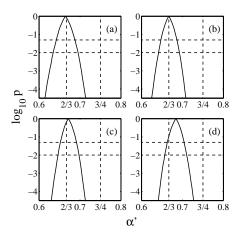
### Detection

Frame 18/78





## Analysis of residuals—p-values—mammals:



- (a) M < 3.2 kg</li>(b) M < 10 kg</li>(c) M < 32 kg</li>(d) all mammals.
- ► For a-d,  $p_{2/3} > 0.05$  and  $p_{3/4} \ll 10^{-4}$ .

### Single Source

#### Distributed Sources

cility location

A reasonable derivation Global redistribution

#### Structure Detection

Hierarchy by division
Hierarchy by shuffling

Hierarchies & Missing Links General structure detection

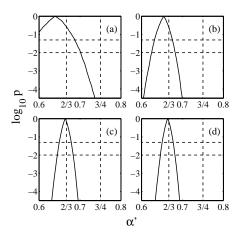
Final words

Poforonoos

Frame 19/78



## Analysis of residuals—p-values—birds:



- (a) M < 0.1 kg
  - (b) M < 1 kg
  - (c) M < 10 kg
  - (d) all birds.
- For a-d,  $p_{2/3} > 0.05$  and  $p_{3/4} \ll 10^{-4}$ .

### Single Source

### Distributed Sources

cility location

A reasonable derivation
Global redistribution

### Structure Detection

Hierarchy by division Hierarchy by shuffling

Hierarchies & Missing Links
General structure detection

Final words

References

Frame 20/78



### Single Source

# Distributed Sources Facility location

Size-density law
A reasonable derivation
Global redistribution networks

### Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links
General structure detection

Final words

References

### Single Source

Distributed
Sources
Facility location
Size-density law
A reasonable derivation

### Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods

\_ . . .

Final words

References

Frame 21/78



### How do we distribute sources?

- Focus on 2-d (results generalize to higher dimensions)
- ► Sources = hospitals, post offices, pubs, ...
- Key problem: How do we cope with uneven population densities?
- Obvious: if density is uniform then sources are best distributed uniformly.
- Which lattice is optimal? The hexagonal lattice Q1: How big should the hexagons be?
- Q2: Given population density is uneven, what do we do?

Single Source

Distributed Sources

Facility location
Size-density law
A reasonable derivation

Global redistributi networks

Structure Detection

Hierarchy by division Hierarchy by shuffling Spectral methods Hierarchies & Missing Links

inal words

References



### How do we distribute sources?

- Focus on 2-d (results generalize to higher dimensions)
- ► Sources = hospitals, post offices, pubs, ...
- Key problem: How do we cope with uneven population densities?
- Obvious: if density is uniform then sources are best distributed uniformly.
- Which lattice is optimal? The hexagonal lattice Q1: How big should the hexagons be?
- Q2: Given population density is uneven, what do we do?

Single Source

Distributed Sources

Facility location Size-density law

A reasonable deriv Global redistribution etworks

Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Link

inal words

References



### How do we distribute sources?

- Focus on 2-d (results generalize to higher dimensions)
- Sources = hospitals, post offices, pubs, ...
- Key problem: How do we cope with uneven population densities?
- Obvious: if density is uniform then sources are best distributed uniformly.
- Which lattice is optimal? The hexagonal lattice Q1: How big should the hexagons be?
- Q2: Given population density is uneven, what do we do?

Single Source

Distributed Sources

Facility location Size-density law

A reasonable deriva Global redistribution networks

Structure Detection

Hierarchy by division Hierarchy by shuffling Spectral methods Hierarchies & Missing Link

inal words

References



### How do we distribute sources?

- Focus on 2-d (results generalize to higher dimensions)
- Sources = hospitals, post offices, pubs, ...
- Key problem: How do we cope with uneven population densities?
- Obvious: if density is uniform then sources are best distributed uniformly.
- Which lattice is optimal? The hexagonal lattice Q1: How big should the hexagons be?
- Q2: Given population density is uneven, what do we do?

Single Source

Distributed Sources

Facility location
Size-density law
A reasonable derivation

Global redistribut networks

Structure Detection

Hierarchy by division Hierarchy by shuffling Spectral methods Hierarchies & Missing Link

inal words

References



### How do we distribute sources?

- Focus on 2-d (results generalize to higher dimensions)
- Sources = hospitals, post offices, pubs, ...
- Key problem: How do we cope with uneven population densities?
- Obvious: if density is uniform then sources are best distributed uniformly.
- Which lattice is optimal? The hexagonal lattice Q1: How big should the hexagons be?
- Q2: Given population density is uneven, what do we do?

Single Source

Distributed Sources

Facility location
Size-density law
A reasonable derivation
Global redistribution

Structure Detection

Hierarchy by division Hierarchy by shuffling Spectral methods Hierarchies & Missing Link

inal words

References



### How do we distribute sources?

- Focus on 2-d (results generalize to higher dimensions)
- Sources = hospitals, post offices, pubs, ...
- Key problem: How do we cope with uneven population densities?
- Obvious: if density is uniform then sources are best distributed uniformly.
- Which lattice is optimal? The hexagonal lattice Q1: How big should the hexagons be?
- Q2: Given population density is uneven, what do we do?

Single Source

Distributed Sources

Facility location
Size-density law
A reasonable derivation

Global redistribution networks

Structure Detection

Hierarchy by division Hierarchy by shuffling Spectral methods Hierarchies & Missing Link

nal words

References



### How do we distribute sources?

- Focus on 2-d (results generalize to higher dimensions)
- Sources = hospitals, post offices, pubs, ...
- Key problem: How do we cope with uneven population densities?
- Obvious: if density is uniform then sources are best distributed uniformly.
- Which lattice is optimal? The hexagonal lattice Q1: How big should the hexagons be?
- Q2: Given population density is uneven, what do we do?

Single Source

Distributed Sources

Facility location
Size-density law
A reasonable derivation

A reasonable deriv Global redistribution networks

Structure Detection

Hierarchy by division Hierarchy by shuffling Spectral methods Hierarchies & Missing Link

nal words

References



### How do we distribute sources?

- Focus on 2-d (results generalize to higher dimensions)
- Sources = hospitals, post offices, pubs, ...
- Key problem: How do we cope with uneven population densities?
- Obvious: if density is uniform then sources are best distributed uniformly.
- Which lattice is optimal? The hexagonal lattice Q1: How big should the hexagons be?
- Q2: Given population density is uneven, what do we do?

Single Source

Distributed Sources

Facility location
Size-density law
A reasonable derivation

Structure Detection

Hierarchy by division Hierarchy by shuffling Spectral methods Hierarchies & Missing Link

nal words

References



### How do we distribute sources?

- Focus on 2-d (results generalize to higher dimensions)
- Sources = hospitals, post offices, pubs, ...
- Key problem: How do we cope with uneven population densities?
- Obvious: if density is uniform then sources are best distributed uniformly.
- Which lattice is optimal? The hexagonal lattice Q1: How big should the hexagons be?
- Q2: Given population density is uneven, what do we do?

Single Source

Distributed Sources

Facility location
Size-density law
A reasonable derivation

Structure Detection

Hierarchy by division Hierarchy by shuffling Spectral methods Hierarchies & Missing Link

nal words

References



## Optimal source allocation

## Solidifying the basic problem

- Given a region with some population distribution ρ, most likely uneven.
- Given resources to build and maintain N facilities.
- Q: How do we locate these N facilities so as to minimize the average distance between an individual's residence and the nearest facility?
- Problem of interested and studied by geographers, sociologists, computer scientists, mathematicians, ...
- ► See work by Stephan [19, 20] and by Gastner and Newman (2006) [8] and work cited by them.

Single Source

Distributed
Sources
Facility location
Size-density law
A reasonable derivation

Structure
Detection
Hierarchy by division
Hierarchy by shuffling
Spectral methods

Spectral methods Hierarchies & Missing Links General structure detection

inal words

leferences



## Optimal source allocation

## Solidifying the basic problem

- Given a region with some population distribution ρ, most likely uneven.
- Given resources to build and maintain N facilities.
- Q: How do we locate these N facilities so as to minimize the average distance between an individual's residence and the nearest facility?
- ► Problem of interested and studied by geographers, sociologists, computer scientists, mathematicians, ...
- ► See work by Stephan [19, 20] and by Gastner and Newman (2006) [8] and work cited by them.

Single Source

Distributed
Sources
Facility location
Size-density law
A reasonable derivation

Structure Detection Hierarchy by division

Hierarchy by shuffling Spectral methods Hierarchies & Missing Links

inal words

References



- Given a region with some population distribution ρ, most likely uneven.
- Given resources to build and maintain N facilities.
- Q: How do we locate these N facilities so as to minimize the average distance between an individual's residence and the nearest facility?
- Problem of interested and studied by geographers, sociologists, computer scientists, mathematicians, ...
- ► See work by Stephan [19, 20] and by Gastner and Newman (2006) [8] and work cited by them.

Single Source

Distributed
Sources
Facility location
Size-density law

Structure Detection

Hierarchy by division Hierarchy by shuffling Spectral methods Hierarchies & Missing Links

inal words

References



### Given a region with some population distribution ρ, most likely uneven.

- Given resources to build and maintain N facilities.
- Q: How do we locate these N facilities so as to minimize the average distance between an individual's residence and the nearest facility?
- Problem of interested and studied by geographers, sociologists, computer scientists, mathematicians, ...
- ► See work by Stephan [19, 20] and by Gastner and Newman (2006) [8] and work cited by them.

Single Source

Distributed
Sources
Facility location
Size-density law

Structure Detection

Hierarchy by division Hierarchy by shuffling Spectral methods Hierarchies & Missing Link

inal words

References



## Solidifying the basic problem

- Given a region with some population distribution ρ, most likely uneven.
- Given resources to build and maintain N facilities.
- Q: How do we locate these N facilities so as to minimize the average distance between an individual's residence and the nearest facility?
- Problem of interested and studied by geographers, sociologists, computer scientists, mathematicians, ...
- ► See work by Stephan [19, 20] and by Gastner and Newman (2006) [8] and work cited by them.

Single Source

Distributed
Sources
Facility location
Size-density law
A reasonable derivation

Structure Detection

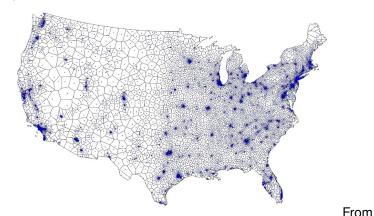
Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Link

Final words

References



## Optimal source allocation



Gastner and Newman (2006) [8]

- Approximately optimal location of 5000 facilities.
- ▶ Based on 2000 Census data.
- ▶ Simulated annealing + Voronoi tessellation.

Optimal supply & Structure detection

Single Source

Distributed Sources

Facility location Size-density law

A reasonable der Global redistribut networks

Structure Detection

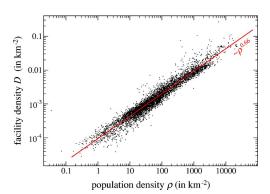
Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

Final words

References

Frame 24/78





## From Gastner and Newman (2006) [8]

- ▶ Optimal facility density D vs. population density  $\rho$ .
- Fit is  $D \propto \rho^{0.66}$  with  $r^2 = 0.94$ .
- ► Looking good for a 2/3 power...

### Single Source

### Distributed Sources

Facility location

Global redistrik networks

## Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

Final words

References

Frame 25/78





### Distributed Sources Facility location

### Structure Detection

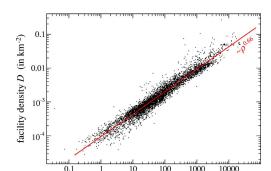
From Gastner and Newman (2006) [8]

0.1

- Optimal facility density D vs. population density ρ.
- Fit is  $D \propto \rho^{0.66}$  with  $r^2 = 0.94$ .
- ▶ Looking good for a 2/3 power...

Frame 25/78





100

population density  $\rho$  (in km<sup>-2</sup>)

1000

0.1

0.01

 $10^{-3}$ 

(in km<sup>-2</sup>)

acility density D



Sources
Facility location
Size-density law
A reasonable derivation

### networks Structure

Detection
Hierarchy by division
Hierarchy by shuffling
Spectral methods

Spectral methods
Hierarchies & Missing Links

Final words

References



0.1

▶ Optimal facility density D vs. population density  $\rho$ .

10

100

population density  $\rho$  (in km<sup>-2</sup>)

1000

10000

- Fit is  $D \propto \rho^{0.66}$  with  $r^2 = 0.94$ .
- ► Looking good for a 2/3 power...

Frame 25/78



### Distributed Sources

Size-density law

A reasonable derivation

Hierarchy by division Hierarchy by shuffling Hierarchies & Missing Links

Single Source

Distributed Sources

Size-density law

Detection

Frame 26/78



## Size-density law:

$$D \propto 
ho^{2/3}$$

▶ In *d* dimensions:

$$D \propto \rho^{d/(d+1)}$$

- ► Why?
- Very different story to branching networks where there is either one source or one sink.
- Now sources & sinks are distributed throughout region...

Single Source

Distributed Sources

Facility location

Size-density law
A reasonable derivation

Global redistribution networks

Structure Detection

Hierarchy by division

Spectral methods
Hierarchies & Missing Links

General structure detection

Final words

References



$$D \propto 
ho^{2/3}$$

▶ In *d* dimensions:

$$D \propto 
ho^{d/(d+1)}$$

- ► Why?
- Very different story to branching networks where there is either one source or one sink.
- Now sources & sinks are distributed throughout region...

Single Source

Distributed Sources

Eacility location

Size-density law
A reasonable derivation

Global redistribution networks

Structure Detection

Hierarchy by division

Spectral methods Hierarchies & Missing Links

- . .

Final words

References



$$D \propto 
ho^{2/3}$$

In d dimensions:

$$D \propto \rho^{d/(d+1)}$$

- ► Why?
- Very different story to branching networks where there is either one source or one sink.
- Now sources & sinks are distributed throughout region...

Single Source

Distributed Sources

Facility location

Size-density law
A reasonable derivatio

Global redistribution networks

Structure Detection

Hierarchy by division

Spectral methods
Hierarchies & Missing Links

Hierarchies & Missing Links General structure detection

Final words

References



## Size-density law:

$$D \propto 
ho^{2/3}$$

In d dimensions:

$$D \propto 
ho^{d/(d+1)}$$

- ► Why?
- Very different story to branching networks where there is either one source or one sink.
- Now sources & sinks are distributed throughout region...

Single Source

### Distributed Sources

Facility loca

Size-density law

A reasonable derivation
Global redistribution

### Structure

Detection

Hierarchy by shuffling
Spectral methods

Hierarchies & Missing Links General structure detection

### Final words

References



## Size-density law:

$$D \propto 
ho^{2/3}$$

In d dimensions:

$$D \propto \rho^{d/(d+1)}$$

- ► Why?
- Very different story to branching networks where there is either one source or one sink.
- Now sources & sinks are distributed throughout region...

Single Source

### Distributed Sources

Facility location

Size-density law

A reasonable derivation

Global redistribution

### Structure Detection

Hierarchy by shuffling
Spectral methods

Hierarchies & Missing Links General structure detection

### Final words

References



Detection

- ➤ One treatment due to Stephan's (1977) [19, 20]: "Territorial Division: The Least-Time Constraint Behind the Formation of Subnational Boundaries" (Science, 1977)
- Zipf-like approach: invokes principle of minimal effort.
- Also known as the Homer principle.

Frame 28/78



Detection

➤ One treatment due to Stephan's (1977) [19, 20]: "Territorial Division: The Least-Time Constraint Behind the Formation of Subnational Boundaries" (Science, 1977)

- Zipf-like approach: invokes principle of minimal effort.
- Also known as the Homer principle.

Frame 28/78



A reasonable derivation Global redistribution networks

Structure Detection

Hierarchy by division Hierarchy by shuffling Spectral methods

Hierarchies & Missing Links
General structure detection

nal words

loforonooo

References

Frame 28/78



- ➤ One treatment due to Stephan's (1977) [19, 20]: "Territorial Division: The Least-Time Constraint Behind the Formation of Subnational Boundaries" (Science, 1977)
- Zipf-like approach: invokes principle of minimal effort.
- Also known as the Homer principle.

### Distributed Sources

A reasonable derivation

Hierarchy by division Hierarchy by shuffling Hierarchies & Missing Links

Single Source

Distributed Sources

A reasonable derivation

Detection

Frame 29/78



## Deriving the optimal source distribution:

- Stronger result obtained by Gusein-Zade (1982). [10]
- Basic idea: Minimize the average distance from a random individual to the nearest facility.
- Assume given a fixed population density ρ defined on a spatial region Ω.
- Formally, we want to find the locations of n sources  $\{\vec{x}_1, \dots, \vec{x}_n\}$  that minimizes the cost function

$$F(\{\vec{x}_1,\ldots,\vec{x}_n\}) = \int_{\Omega} \rho(\vec{x}) \min_i ||\vec{x} - \vec{x}_i|| d\vec{x}.$$

- Also known as the p-median problem.
- Not easy... in fact this one is an NP-hard problem. <sup>[8]</sup>

Single Source

Distributed Sources

Size-density law
A reasonable derivation

Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

inal words

References

Frame 30/78



## Deriving the optimal source distribution:

- Stronger result obtained by Gusein-Zade (1982). [10]
- Basic idea: Minimize the average distance from a random individual to the nearest facility.
- Assume given a fixed population density ρ defined on a spatial region Ω.
- Formally, we want to find the locations of n sources  $\{\vec{x}_1, \dots, \vec{x}_n\}$  that minimizes the cost function

$$F(\{\vec{x}_1,\ldots,\vec{x}_n\}) = \int_{\Omega} \rho(\vec{x}) \min_i ||\vec{x} - \vec{x}_i|| d\vec{x}.$$

- Also known as the p-median problem.
- ▶ Not easy... in fact this one is an NP-hard problem. [8]

### Single Source

## Distributed Sources

Facility location
Size-density law
A reasonable derivation
Global redistribution

### Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

inal words

References





## Size-density law

## Deriving the optimal source distribution:

- Stronger result obtained by Gusein-Zade (1982). [10]
- Basic idea: Minimize the average distance from a random individual to the nearest facility.
- Assume given a fixed population density  $\rho$  defined on a spatial region  $\Omega$ .
- ▶ Formally, we want to find the locations of *n* sources

$$F(\{\vec{x}_1,\ldots,\vec{x}_n\}) = \int_{\Omega} \rho(\vec{x}) \min_i ||\vec{x} - \vec{x}_i|| d\vec{x}.$$

- Also known as the p-median problem.
- Not easy... in fact this one is an NP-hard problem.

### Single Source

### Distributed Sources

A reasonable derivation

### Detection





## Size-density law

## Deriving the optimal source distribution:

- Stronger result obtained by Gusein-Zade (1982). [10]
- Basic idea: Minimize the average distance from a random individual to the nearest facility.
- Assume given a fixed population density  $\rho$  defined on a spatial region Ω.
- Formally, we want to find the locations of n sources  $\{\vec{x}_1, \dots, \vec{x}_n\}$  that minimizes the cost function

$$F(\{\vec{x}_1,\ldots,\vec{x}_n\}) = \int_{\Omega} \rho(\vec{x}) \min_i ||\vec{x} - \vec{x}_i|| d\vec{x}.$$

- Also known as the p-median problem.
- ▶ Not easy... in fact this one is an NP-hard problem. [8]

### Single Source

### Distributed Sources

Facility location
Size-density law
A reasonable derivation

### Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

### inal words

References





## Deriving the optimal source distribution:

- Stronger result obtained by Gusein-Zade (1982). [10]
- Basic idea: Minimize the average distance from a random individual to the nearest facility.
- Assume given a fixed population density  $\rho$  defined on a spatial region Ω.
- Formally, we want to find the locations of n sources  $\{\vec{x}_1, \dots, \vec{x}_n\}$  that minimizes the cost function

$$F(\{\vec{x}_1,\ldots,\vec{x}_n\}) = \int_{\Omega} \rho(\vec{x}) \min_i ||\vec{x} - \vec{x}_i|| d\vec{x}.$$

- Also known as the p-median problem.
- ▶ Not easy... in fact this one is an NP-hard problem. [3]

### Single Source

### Distributed Sources

Facility location
Size-density law
A reasonable derivation

A reasonable derivati Global redistribution networks

### Structure Detection

Hierarchy by division

Hierarchy by shuffling

Spectral methods

Hierarchies & Missing Links

inal words

References

Frame 30/78



## Size-density law

## Deriving the optimal source distribution:

- ▶ Stronger result obtained by Gusein-Zade (1982). [10]
- Basic idea: Minimize the average distance from a random individual to the nearest facility.
- Assume given a fixed population density ρ defined on a spatial region Ω.
- Formally, we want to find the locations of n sources  $\{\vec{x}_1, \dots, \vec{x}_n\}$  that minimizes the cost function

$$F(\{\vec{x}_1,\ldots,\vec{x}_n\}) = \int_{\Omega} \rho(\vec{x}) \min_i ||\vec{x} - \vec{x}_i|| d\vec{x}.$$

- Also known as the p-median problem.
- ▶ Not easy... in fact this one is an NP-hard problem. [8]

Single Source

Distributed Sources

Facility location
Size-density law
A reasonable derivation

Global redistribution networks

Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

inal words

References

Frame 30/78



# Size-density law

# Deriving the optimal source distribution:

- Stronger result obtained by Gusein-Zade (1982). [10]
- Basic idea: Minimize the average distance from a random individual to the nearest facility.
- Assume given a fixed population density  $\rho$  defined on a spatial region Ω.
- Formally, we want to find the locations of n sources  $\{\vec{x}_1, \dots, \vec{x}_n\}$  that minimizes the cost function

$$F(\{\vec{x}_1,\ldots,\vec{x}_n\}) = \int_{\Omega} \rho(\vec{x}) \min_i ||\vec{x} - \vec{x}_i|| d\vec{x}.$$

- Also known as the p-median problem.
- ▶ Not easy... in fact this one is an NP-hard problem. [8]

Single Source

## Distributed Sources

Facility location Size-density law

A reasonable derivation Global redistribution networks

### Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

inal words

References





▶ By varying  $\{\vec{x}_1,...,\vec{x}_n\}$ , minimize

$$G(A) = c \int_{\Omega} \rho(\vec{x}) A(\vec{x})^{1/2} d\vec{x} - \lambda \left( n - \int_{\Omega} \left[ A(\vec{x}) \right]^{-1} d\vec{x} \right)$$

- ▶ Involves estimating typical distance from  $\vec{x}$  to the nearest source (say i) as  $c_i A(\vec{x})^{1/2}$  where  $c_i$  is a shape factor for the ith Voronoi cell.
- ▶ Sneakiness: set  $c_i = c$ .
- ▶ Compute  $\delta G/\delta A$ , the <u>functional derivative</u> ( $\boxplus$ ).
- ▶ Solve and substitute D = 1/A, we find

$$D(\vec{x}) = \left(\frac{c}{2\lambda}\rho\right)^{2/3}.$$

Single Source

Distributed Sources

Facility location
Size-density law
A reasonable derivation

Structure
Detection
Hierarchy by division
Hierarchy by shuffling

General structure det

-inai words

References



▶ By varying  $\{\vec{x}_1,...,\vec{x}_n\}$ , minimize

$$G(A) = c \int_{\Omega} \rho(\vec{x}) A(\vec{x})^{1/2} d\vec{x} - \lambda \left( n - \int_{\Omega} \left[ A(\vec{x}) \right]^{-1} d\vec{x} \right)$$

- ▶ Involves estimating typical distance from  $\vec{x}$  to the nearest source (say i) as  $c_i A(\vec{x})^{1/2}$  where  $c_i$  is a shape factor for the ith Voronoi cell.
- ▶ Sneakiness: set  $c_i = c$ .
- ▶ Compute  $\delta G/\delta A$ , the <u>functional derivative</u> ( $\boxplus$ ).
- ▶ Solve and substitute D = 1/A, we find

$$D(\vec{x}) = \left(\frac{c}{2\lambda}\rho\right)^{2/3}.$$

## Single Source

## Distributed Sources

Facility location
Size-density law

A reasonable derivation Global redistribution networks

## Structure Detection

Hierarchy by division Hierarchy by shuffling Spectral methods Hierarchies & Missing Links

inal words

References



▶ By varying  $\{\vec{x}_1,...,\vec{x}_n\}$ , minimize

$$G(A) = c \int_{\Omega} \rho(\vec{x}) A(\vec{x})^{1/2} d\vec{x} - \lambda \left( n - \int_{\Omega} \left[ A(\vec{x}) \right]^{-1} d\vec{x} \right)$$

- ▶ Involves estimating typical distance from  $\vec{x}$  to the nearest source (say i) as  $c_i A(\vec{x})^{1/2}$  where  $c_i$  is a shape factor for the ith Voronoi cell.
- ▶ Sneakiness: set  $c_i = c$ .
- ▶ Compute  $\delta G/\delta A$ , the <u>functional derivative</u> ( $\boxplus$ ).
- ▶ Solve and substitute D = 1/A, we find

$$D(\vec{x}) = \left(\frac{c}{2\lambda}\rho\right)^{2/3}.$$

Single Source

Distributed Sources

Facility location Size-density law A reasonable derivation

A reasonable derivation Global redistribution networks

Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Link

inal words

References



▶ By varying  $\{\vec{x}_1,...,\vec{x}_n\}$ , minimize

$$G(A) = c \int_{\Omega} \rho(\vec{x}) A(\vec{x})^{1/2} d\vec{x} - \lambda \left( n - \int_{\Omega} \left[ A(\vec{x}) \right]^{-1} d\vec{x} \right)$$

- ▶ Involves estimating typical distance from  $\vec{x}$  to the nearest source (say i) as  $c_i A(\vec{x})^{1/2}$  where  $c_i$  is a shape factor for the ith Voronoi cell.
- ▶ Sneakiness: set  $c_i = c$ .
- ▶ Compute  $\delta G/\delta A$ , the <u>functional derivative</u> ( $\boxplus$ ).
- ▶ Solve and substitute D = 1/A, we find

$$D(\vec{x}) = \left(\frac{c}{2\lambda}\rho\right)^{2/3}.$$

## Single Source

## Distributed Sources

Facility location Size-density law

A reasonable derivation Global redistribution networks

## Structure Detection

Hierarchy by division Hierarchy by shuffling Spectral methods Hierarchies & Missing Link

inal words

References



▶ By varying  $\{\vec{x}_1,...,\vec{x}_n\}$ , minimize

$$G(A) = c \int_{\Omega} \rho(\vec{x}) A(\vec{x})^{1/2} d\vec{x} - \lambda \left( n - \int_{\Omega} \left[ A(\vec{x}) \right]^{-1} d\vec{x} \right)$$

- ▶ Involves estimating typical distance from  $\vec{x}$  to the nearest source (say i) as  $c_i A(\vec{x})^{1/2}$  where  $c_i$  is a shape factor for the ith Voronoi cell.
- ▶ Sneakiness: set  $c_i = c$ .
- ▶ Compute  $\delta G/\delta A$ , the <u>functional derivative</u> ( $\boxplus$ ).
- ▶ Solve and substitute D = 1/A, we find

$$D(\vec{x}) = \left(\frac{c}{2\lambda}\rho\right)^{2/3}.$$

Single Source

Distributed Sources

Facility location
Size-density law
A reasonable derivation

A reasonable derivation Global redistribution networks

Structure Detection Hierarchy by division

Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Link

inal words

References



## Distributed Sources

A reasonable derivation

## Global redistribution networks

Hierarchy by division Hierarchy by shuffling Hierarchies & Missing Links

## Single Source

Distributed Sources

Global redistribution networks

Detection

Frame 32/78



# One more thing:

- How do we supply these facilities?
- How do we best redistribute mail? People?
- ► How do we get beer to the pubs?
- Gaster and Newman model: cost is a function of basic maintenance and travel time:

$$C_{\text{maint}} + \gamma C_{\text{travel}}$$
.

► Travel time is more complicated: Take 'distance' between nodes to be a composite of shortest path distance ℓ<sub>ij</sub> and number of legs to journey:

$$(1-\delta)\ell_{ij}+\delta(\#\mathsf{hops}).$$

▶ When  $\delta = 1$ , only number of hops matters.

## Single Source

### Distributed Sources

Sources

Size-density law
A reasonable derivation

Global redistribution networks

## Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

General structure de

Final words

References



## One more thing:

- How do we supply these facilities?
- How do we best redistribute mail? People?
- ► How do we get beer to the pubs?
- Gaster and Newman model: cost is a function of

$$C_{\text{maint}} + \gamma C_{\text{travel}}$$
.

Travel time is more complicated: Take 'distance'

$$(1-\delta)\ell_{ij}+\delta(\#\mathsf{hops}).$$

▶ When  $\delta = 1$ , only number of hops matters.

## Single Source

## Distributed

Sources

Global redistribution networks

# Detection



## One more thing:

- How do we supply these facilities?
- How do we best redistribute mail? People?
- How do we get beer to the pubs?
- Gaster and Newman model: cost is a function of

$$C_{\text{maint}} + \gamma C_{\text{travel}}$$

Travel time is more complicated: Take 'distance'

$$(1-\delta)\ell_{ij}+\delta(\#\mathsf{hops}).$$

▶ When  $\delta = 1$ , only number of hops matters.

## Single Source

## Distributed

Sources

Global redistribution networks

# Detection



# One more thing:

- How do we supply these facilities?
- How do we best redistribute mail? People?
- ▶ How do we get beer to the pubs?
- Gaster and Newman model: cost is a function of basic maintenance and travel time:

$$C_{\text{maint}} + \gamma C_{\text{travel}}$$
.

► Travel time is more complicated: Take 'distance' between nodes to be a composite of shortest path distance ℓ<sub>ij</sub> and number of legs to journey:

$$(1-\delta)\ell_{ij}+\delta(\#\mathsf{hops}).$$

▶ When  $\delta = 1$ , only number of hops matters.

## Single Source

## Distributed

Sources

Size-density law
A reasonable derivation

Global redistribution networks

### Structure Detection

Jetection Hierarchy by divisi

Spectral methods
Hierarchies & Missing Links

General structure detect

### Final words

References



## One more thing:

- How do we supply these facilities?
- How do we best redistribute mail? People?
- ▶ How do we get beer to the pubs?
- Gaster and Newman model: cost is a function of basic maintenance and travel time:

$$C_{\text{maint}} + \gamma C_{\text{travel}}$$
.

► Travel time is more complicated: Take 'distance' between nodes to be a composite of shortest path distance ℓ<sub>ij</sub> and number of legs to journey:

$$(1 - \delta)\ell_{ij} + \delta(\#hops).$$

▶ When  $\delta = 1$ , only number of hops matters.

## Single Source

## Distributed Sources

Sources

Size-density law
A reasonable derivation

Global redistribution networks

# Structure Detection

Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

Hierarchies & Missing Links General structure detection

inal words

References



# One more thing:

- How do we supply these facilities?
- How do we best redistribute mail? People?
- ▶ How do we get beer to the pubs?
- Gaster and Newman model: cost is a function of basic maintenance and travel time:

$$C_{\text{maint}} + \gamma C_{\text{travel}}$$
.

► Travel time is more complicated: Take 'distance' between nodes to be a composite of shortest path distance ℓ<sub>ij</sub> and number of legs to journey:

$$(1 - \delta)\ell_{ij} + \delta(\#hops).$$

▶ When  $\delta = 1$ , only number of hops matters.

## Single Source

### Distributed Sources

Facility location Size-density law

A reasonable derivation Global redistribution networks

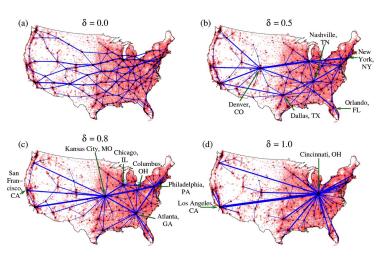
## Structure Detection

Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

Final words

References





From Gastner and Newman (2006) [8]

Optimal supply & Structure detection

Single Source

Distributed Sources

> Facility location Size-density law

Global redistribution networks

Structure Detection

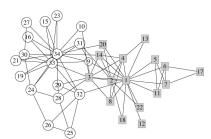
> Herarchy by division Hierarchy by shuffling Spectral methods

Cinal ...auda

References

Frame 34/78





▲ Zachary's karate club [25, 16]

## ► The issue:

how do we elucidate the internal structure of large networks across many scales?

## Single Source

### Distributed Sources

Facility location
Size-density law
A reasonable derivatio

A reasonable derivation Global redistribution networks

## Structure Detection

Hierarchy by shuffling

Spectral methods
Hierarchies & Missing Links
General structure detection

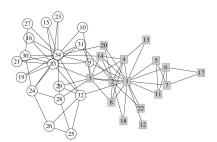
inal words

-ınal words

References







► The issue:

how do we elucidate the internal structure of large networks across many scales?

- ▲ Zachary's karate club [25, 16]
  - Possible substructures: hierarchies, cliques, rings, ...
  - ▶ Plus: All combinations of substructures
  - Much focus on hierarchies...

Single Source

Distributed Sources

Facility location

A reasonable derivation Global redistribution networks

Structure Detection

Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

9..............................

Final words

References



► The issue:

how do we elucidate the internal structure of large networks across many scales?

- ▲ Zachary's karate club [25, 16]
  - Possible substructures: hierarchies, cliques, rings, ...
  - Plus: All combinations of substructures.
  - Much focus on hierarchies...

## Single Source

## Distributed Sources

Sources
Facility location

Size-density law

A reasonable derivation
Global redistribution
networks

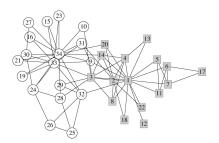
## Structure Detection

Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

inal words

References





## ► The issue:

how do we elucidate the internal structure of large networks across many scales?

- ▲ Zachary's karate club [25, 16]
  - Possible substructures: hierarchies, cliques, rings, ...
  - Plus: All combinations of substructures.
  - Much focus on hierarchies...

## Single Source

## Distributed

Sources

Size-density law A reasonable derivation Global redistribution networks

## Structure Detection

Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

### inal words

References



## Single Source

## Distributed Sources

Facility location
Size-density law

A reasonable derivation

Global redistribution networks

# Structure Detection

Hierarchy by division

Hierarchy by shuffling

Hierarchies & Missing Links

General structure detection

## Final words

References

Single Source

Distributed Sources

Facility location

Size-density law
A reasonable derivation

Global redistrib

Structure Detection

Hierarchy by division

Hierarchy by shuffling

Spectral methods
Hierarchies & Missing Links

denoral of details at

Final words

References

Frame 36/78



## Top down:

- Idea: Identify global structure first and recursively uncover more detailed structure.
- Basic objective: find dominant components that have significantly more links within than without, as compared to randomized version.
- ► Following comes from "Finding and evaluating community structure in networks" by Newman and Girvan (PRE, 2004). [16]
- ► See also
  - "Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality" by Newman (PRE, 2001). [14, 15]
  - "Community structure in social and biological networks" by Girvan and Newman (PNAS, 2002). [9]

Single Source

Distributed Sources

Facility location Size-density law

A reasonable derivatio Global redistribution networks

Detection
Hierarchy by division
Hierarchy by shuffling

dierarchies & Missing Links General structure detection

Final words

References



## Top down:

- Idea: Identify global structure first and recursively uncover more detailed structure.
- Basic objective: find dominant components that have significantly more links within than without, as compared to randomized version.
- ► Following comes from "Finding and evaluating community structure in networks" by Newman and Girvan (PRE, 2004). [16]
- ▶ See also
  - "Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality" by Newman (PRE 2001). [14, 15]
  - "Community structure in social and biological networks" by Girvan and Newman (PNAS, 2002). <sup>[9]</sup>

Single Source

Distributed Sources

Sources Facility location

Size-density law
A reasonable derivation
Global redistribution
networks

Structure Detection

Hierarchy by division Hierarchy by shuffling Spectral methods Hierarchies & Missing Links

inal words

References



## Top down:

- Idea: Identify global structure first and recursively uncover more detailed structure.
- Basic objective: find dominant components that have significantly more links within than without, as compared to randomized version.
- Following comes from "Finding and evaluating community structure in networks" by Newman and Girvan (PRE, 2004). [16]
- See also

Optimal supply & Structure detection

Single Source

Distributed Sources

Structure Detection Hierarchy by division



## Top down:

- Idea: Identify global structure first and recursively uncover more detailed structure.
- Basic objective: find dominant components that have significantly more links within than without, as compared to randomized version.
- Following comes from "Finding and evaluating community structure in networks" by Newman and Girvan (PRE, 2004). [16]
- See also
  - "Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality" by Newman (PRE, 2001). [14, 15]
  - 2. "Community structure in social and biological networks" by Girvan and Newman (PNAS, 2002). [9]

Single Source

Distributed Sources

Size-density law
A reasonable derivation
Global redistribution
networks

Structure
Detection
Hierarchy by division
Hierarchy by shuffling

Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links
General structure detection

inal words

References



## Top down:

- Idea: Identify global structure first and recursively uncover more detailed structure.
- Basic objective: find dominant components that have significantly more links within than without, as compared to randomized version.
- Following comes from "Finding and evaluating community structure in networks" by Newman and Girvan (PRE, 2004). [16]
- See also
  - "Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality" by Newman (PRE, 2001). [14, 15]
  - "Community structure in social and biological networks" by Girvan and Newman (PNAS, 2002). [9]

Single Source

Distributed Sources

Size-density law
A reasonable derivation
Global redistribution

Structure
Detection
Hierarchy by division
Hierarchy by shuffling

Hierarchy by shuffling Spectral methods Hierarchies & Missing Links General structure detection

inal words

References



## Single Source

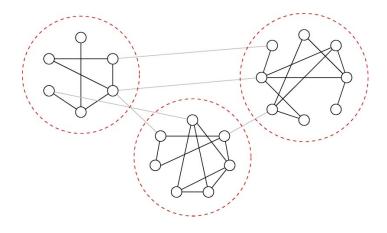
## Distributed Sources

# Detection

Hierarchy by division

Frame 38/78





## Idea:

Edges that connect communities have higher betweenness than edges within communities.

# One class of structure-detection algorithms:

- 1. Compute edge betweenness for whole network.
- 2. Remove edge with highest betweenness
- 3. Recompute edge betweenness
- 4. Repeat steps 2 and 3 until all edges are removed.
- 5 Record when components appear as a function of # edges removed
- 6 Generate dendogram revealing hierarchical structure.

Single Source

Distributed Sources

Size-density law

A reasonable derivation Global redistribution networks

Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

Final words

References



- Compute edge betweenness for whole network.
- 2. Remove edge with highest betweenness.
- 3. Recompute edge betweenness
- 4. Repeat steps 2 and 3 until all edges are removed.
- 5 Record when components appear as a function of # edges removed
- 6 Generate dendogram revealing hierarchical structure.

Single Source

Distributed Sources

Facility location

A reasonable derivation Global redistribution networks

Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

Hierarchies & Missing Links General structure detection

Final words

References



# One class of structure-detection algorithms:

- Compute edge betweenness for whole network.
- Remove edge with highest betweenness.
- 3. Recompute edge betweenness

Single Source

Distributed Sources

Detection

Hierarchy by division





# One class of structure-detection algorithms:

- Compute edge betweenness for whole network.
- Remove edge with highest betweenness.
- 3. Recompute edge betweenness
- 4. Repeat steps 2 and 3 until all edges are removed.

Single Source

Distributed Sources

Detection

Hierarchy by division





# One class of structure-detection algorithms:

- 1. Compute edge betweenness for whole network.
- 2. Remove edge with highest betweenness.
- 3. Recompute edge betweenness
- 4. Repeat steps 2 and 3 until all edges are removed.
- 5 Record when components appear as a function of # edges removed.
- 6 Generate dendogram revealing hierarchical structure.

Single Source

Distributed Sources

Facility location

A reasonable derivation Global redistribution networks

Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods

General structure detection

Final words

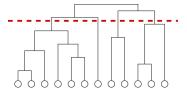
References





# One class of structure-detection algorithms:

- Compute edge betweenness for whole network.
- Remove edge with highest betweenness.
- 3. Recompute edge betweenness
- 4. Repeat steps 2 and 3 until all edges are removed.
- 5 Record when components appear as a function of # edges removed.
- 6 Generate dendogram revealing hierarchical structure.



## Single Source

## Distributed Sources

# Detection

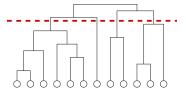
Hierarchy by division





# One class of structure-detection algorithms:

- Compute edge betweenness for whole network.
- 2. Remove edge with highest betweenness.
- 3. Recompute edge betweenness
- 4. Repeat steps 2 and 3 until all edges are removed.
- 5 Record when components appear as a function of # edges removed.
- 6 Generate dendogram revealing hierarchical structure.



Red line indicates appearance of four (4) components at a certain level.

## Single Source

# Distributed Sources

acility location

A reasonable derivation
Global redistribution
networks

## Structure Detection

Hierarchy by division Hierarchy by shuffling Spectral methods Hierarchies & Missing Links

## inal words

References



## Key element:

- Recomputing betweenness.
- Reason: Possible to have a low betweenness in links that connect large communities if other links carry majority of shortest paths.

# When to stop?:

- How do we know which divisions are meaningful?
- Modularity measure: difference in fraction of within component nodes to that expected for randomized version:
  - $Q = 2J_1 P_R = (2J_1 P_R) \cdot 1 = 115 = 115$ where  $e_R$  is the fraction of edges between identified communities i and i

Single Source

Distributed Sources

Sources

Size-density law
A reasonable derivation
Global redistribution
networks

Structure Detection

Hierarchy by division

Spectral methods

Hierarchies & Missing Links General structure detection

Final words

References



Single Source Distributed

Sources

Structure Detection Hierarchy by division

## Key element:

- Recomputing betweenness.
- Reason: Possible to have a low betweenness in links that connect large communities if other links carry majority of shortest paths.





## Key element:

- Recomputing betweenness.
- Reason: Possible to have a low betweenness in links that connect large communities if other links carry majority of shortest paths.

# When to stop?:

- ► How do we know which divisions are meaningful?
- Modularity measure: difference in fraction of within

Single Source

Distributed

Sources

Detection

Hierarchy by division



# Key element:

- Recomputing betweenness.
- Reason: Possible to have a low betweenness in links that connect large communities if other links carry majority of shortest paths.

# When to stop?:

- How do we know which divisions are meaningful?
- Modularity measure: difference in fraction of within component nodes to that expected for randomized version:
  - $Q = \sum_{i} [\theta_{ii} (\sum_{j} \theta_{ij})^{2}] = \text{TrE} ||\text{E}^{2}||_{1},$  where  $\theta_{ij}$  is the fraction of edges between identified communities i and j.

Single Source

Distributed

Sources

Size-density law
A reasonable derivation
Global redistribution

Structure Detection

Hierarchy by division Hierarchy by shuffling Spectral methods

Hierarchies & Missing Links General structure detection

Final words

References



### Key element:

- Recomputing betweenness.
- Reason: Possible to have a low betweenness in links that connect large communities if other links carry majority of shortest paths.

### When to stop?:

- How do we know which divisions are meaningful?
- Modularity measure: difference in fraction of within component nodes to that expected for randomized version:

$$Q = \sum_{i} [e_{ii} - (\sum_{j} e_{ij})^{2}] = \text{Tr}\mathbf{E} - ||\mathbf{E}^{2}||_{1},$$
 where  $e_{ij}$  is the fraction of edges between identified communities  $i$  and  $j$ .

#### Single Source

### Distributed Sources

Facility location
Size-density law

A reasonable derivat Global redistribution networks

#### Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

-- . . . . .

Final words

References

Frame 40/78



### Key element:

- Recomputing betweenness.
- Reason: Possible to have a low betweenness in links that connect large communities if other links carry majority of shortest paths.

### When to stop?:

- How do we know which divisions are meaningful?
- Modularity measure: difference in fraction of within component nodes to that expected for randomized version:

$$Q = \sum_{i} [e_{ii} - (\sum_{j} e_{ij})^{2}] = \text{Tr} \mathbf{E} - ||\mathbf{E}^{2}||_{1},$$
 where  $e_{ij}$  is the fraction of edges between identified communities  $i$  and  $j$ .

#### Single Source

### Distributed Sources

Facility location
Size-density law
A reasonable derivation

Global redistribution networks

#### Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

Final words

References

Frame 40/78



Structure Detection

Hierarchy by division

Frame 41/78



### Test case:

- Generate random community-based networks.
- Add edges randomly within and across communities.
- Example:

$$\langle k \rangle_{\rm in} = 6$$
 and  $\langle k \rangle_{\rm out} = 2$ .

### Test case:

- Generate random community-based networks.
- N = 128 with four communities of size 32.
- Add edges randomly within and across communities.
- Example:

Frame 41/78





A reasonable deriva Global redistribution networks

Structure Detection

Hierarchy by division
Hierarchy by shuffling

Spectral methods
Hierarchies & Missing Links
General structure detection

Einal worde

Final words

References



#### Test case:

- Generate random community-based networks.
- $\triangleright$  N = 128 with four communities of size 32.
- Add edges randomly within and across communities.
- Example:

 $\langle k \rangle_{\rm in} = 6$  and  $\langle k \rangle_{\rm out} = 2$ .

Structure Detection

Hierarchy by division

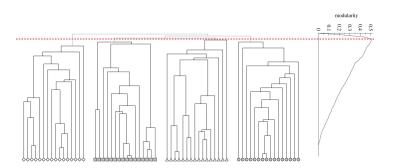
#### Test case:

- Generate random community-based networks.
- N = 128 with four communities of size 32.
- Add edges randomly within and across communities.
- Example:

 $\langle k \rangle_{\rm in} = 6$  and  $\langle k \rangle_{\rm out} = 2$ .

Frame 41/78





- Maximum modularity  $Q \simeq 0.5$  obtained when four communities are uncovered.
- ► Further 'discovery' of internal structure is somewhat meaningless, as any communities arise accidentally.

Single Source

Distributed

Sources Facility location

Size-density law
A reasonable derivation
Global redistribution

Structure Detection

Hierarchy by division

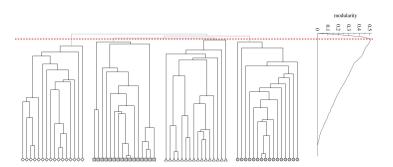
Spectral methods
Hierarchies & Missing Links

inal words

References

Frame 42/78





- Maximum modularity  $Q \simeq 0.5$  obtained when four communities are uncovered.
- Further 'discovery' of internal structure is somewhat meaningless, as any communities arise accidentally.

Single Source

Distributed Sources

Facility location

Size-density law
A reasonable derivation
Global redistribution
networks

Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Link

General structure de

inal words

References

Frame 42/78





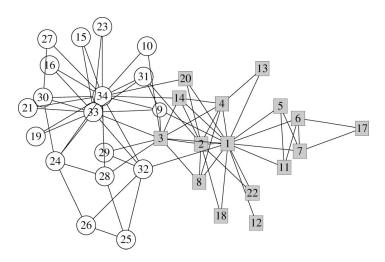
#### Distributed Sources

### Detection

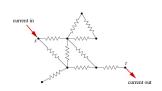
Hierarchy by division

Frame 43/78





► Factions in Zachary's karate club network. [25]



- Unit resistors on each edge.
- For every pair of nodes s (source) and t (sink), set up unit currents in at s and out a t.
- Measure absolute current along each edge  $\ell$ ,  $|I_{\ell,st}|$ .
- Sum  $|I_{\ell,st}|$  over all pairs of nodes to obtain electronic betweenness for edge  $\ell$ .
- ► (Equivalent to random walk betweenness.)
- ► Electronic betweenness for edge between nodes i and j:

$$B_{ij}^{\text{elec}} = a_{ij}|V_i - V_j|.$$

Upshot: specific measure of betweenness not too important. Optimal supply & Structure detection

Single Source

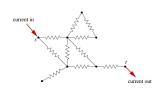
Distributed
Sources
Facility location
Size-density law

Detection
Hierarchy by division
Hierarchy by shuffling
Spectral methods

Final words

References





- Unit resistors on each edge.
- For every pair of nodes s (source) and t (sink), set up unit currents in at s and out at t.
- Measure absolute current along each edge  $\ell$ ,  $|I_{\ell,st}|$ .
- Sum  $|I_{\ell,st}|$  over all pairs of nodes to obtain electronic betweenness for edge  $\ell$ .
- ► (Equivalent to random walk betweenness.)
- ► Electronic betweenness for edge between nodes *i* and *j*:

$$B_{ij}^{\text{elec}} = a_{ij} |V_i - V_j|.$$

Upshot: specific measure of betweenness not too important. Optimal supply & Structure detection

Single Source

#### Distributed Sources

Sources
Facility location
Size-density law
A reasonable derivation

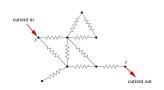
# Structure Detection Hierarchy by division

Spectral methods
Hierarchies & Missing Links

inal words

References





- Unit resistors on each edge.
- For every pair of nodes s (source) and t (sink), set up unit currents in at s and out at t.
- Measure absolute current along each edge  $\ell$ ,  $|I_{\ell,st}|$ .
- Sum  $|I_{\ell,st}|$  over all pairs of nodes to obtain electronic betweenness for edge  $\ell$ .
- ► (Equivalent to random walk betweenness.)
- ► Electronic betweenness for edge between nodes *i* and *j*:

$$B_{ij}^{
m elec}=a_{ij}|V_i-V_j|.$$

Upshot: specific measure of betweenness not too important. Optimal supply & Structure detection

Single Source

#### Distributed Sources

Sources
Facility location
Size-density law
A reasonable derivatio

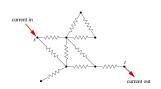
#### Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

inal words

References





- Unit resistors on each edge.
- For every pair of nodes s (source) and t (sink), set up unit currents in at s and out at t.
- Measure absolute current along each edge  $\ell$ ,  $|I_{\ell,st}|$ .
- ▶ Sum  $|I_{\ell,st}|$  over all pairs of nodes to obtain electronic betweenness for edge  $\ell$ .
- ► (Equivalent to random walk betweenness.)
- ► Electronic betweenness for edge between nodes *i* and *j*:

$$B_{ij}^{\text{elec}} = a_{ij} |V_i - V_j|.$$

Upshot: specific measure of betweenness not too important. Optimal supply & Structure detection

Single Source

Distributed Sources

Facility location
Size-density law
A reasonable derivation

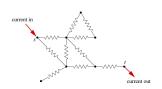
Structure
Detection
Hierarchy by division

Spectral methods
Hierarchies & Missing Links

inal words

References





- Unit resistors on each edge.
- For every pair of nodes s (source) and t (sink), set up unit currents in at s and out at t.
- Measure absolute current along each edge  $\ell$ ,  $|I_{\ell,st}|$ .
- ▶ Sum  $|I_{\ell,st}|$  over all pairs of nodes to obtain electronic betweenness for edge  $\ell$ .
- (Equivalent to random walk betweenness.)
- ► Electronic betweenness for edge between nodes *i* and *j*:

$$B_{ij}^{\text{elec}} = a_{ij} |V_i - V_j|$$

Upshot: specific measure of betweenness not too important. Optimal supply & Structure detection

Single Source

#### Distributed Sources

Sources
Facility location
Size-density law
A reasonable derivatio

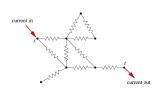
Structure Detection

Hierarchy by division Hierarchy by shuffling Spectral methods Hierarchies & Missing Links

inal words

References





- Unit resistors on each edge.
- For every pair of nodes s (source) and t (sink), set up unit currents in at s and out at t.
- ▶ Measure absolute current along each edge  $\ell$ ,  $|I_{\ell,st}|$ .
- ▶ Sum  $|I_{\ell,st}|$  over all pairs of nodes to obtain electronic betweenness for edge  $\ell$ .
- (Equivalent to random walk betweenness.)
- ► Electronic betweenness for edge between nodes *i* and *j*:

$$B_{ij}^{\text{elec}} = a_{ij}|V_i - V_j|.$$

Upshot: specific measure of betweenness not too important. Optimal supply & Structure detection

Single Source

#### Distributed Sources

Sources
Facility location
Size-density law
A reasonable derivation
Global redistribution

# Detection Hierarchy by division

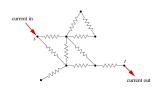
Structure

Hierarchy by shuffling Spectral methods Hierarchies & Missing Links General structure detection

inal words

References





- Unit resistors on each edge.
- For every pair of nodes s (source) and t (sink), set up unit currents in at s and out at t.
- ▶ Measure absolute current along each edge  $\ell$ ,  $|I_{\ell,st}|$ .
- ▶ Sum  $|I_{\ell,st}|$  over all pairs of nodes to obtain electronic betweenness for edge  $\ell$ .
- (Equivalent to random walk betweenness.)
- ► Electronic betweenness for edge between nodes *i* and *j*:

$$B_{ij}^{\text{elec}} = a_{ij}|V_i - V_j|.$$

Upshot: specific measure of betweenness not too important. Optimal supply & Structure detection

Single Source

Distributed Sources

Structure

Facility location
Size-density law
A reasonable derivation
Global redistribution

Detection

Hierarchy by division

Hierarchy by shuffling

Spectral methods

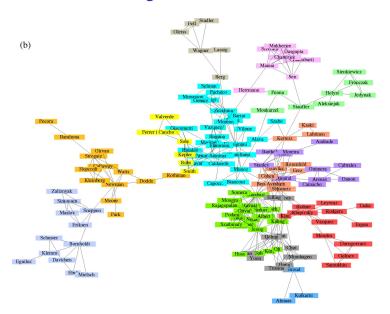
Spectral methods
Hierarchies & Missing Links
General structure detection

inal words

References



## Scientists working on networks



### Optimal supply & Structure detection

#### Single Source

### Distributed

Sources Facility location

A reasonable derivation
Global redistribution

#### Structure Detection

Detection

Hierarchy by division

Hierarchy by shuffling Spectral methods Hierarchies & Missing Links

#### nal words

#### References

Frame 45/78





#### Distributed Sources

Facility location

A reasonable derivation Global redistribution

### Structure

Hierarchy by division

Hierarchy by shuffling

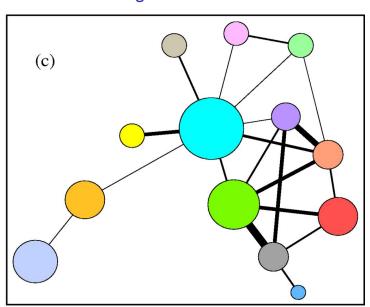
Hierarchies & Missing Links

#### Final word

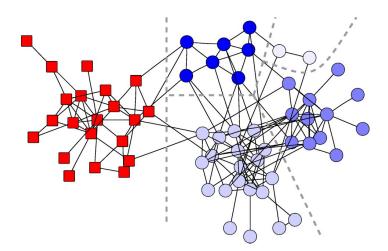
References

Frame 46/78





### Dolphins!



## Optimal supply & Structure detection

#### Single Source

### Distributed

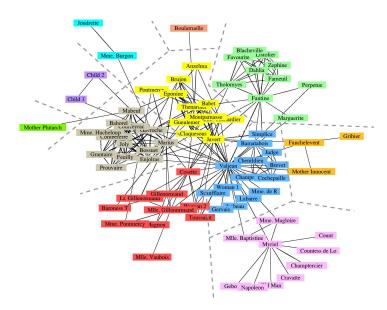
Hierarchy by division

Frame 47/78





### Les Miserables



### Optimal supply & Structure detection

#### Single Source

### Distributed Sources

- Facility locatio
- A reasonable derivation

#### Structure

#### Detection

#### Hierarchy by division

- Hierarchy by shuffling Spectral methods
- Hierarchies & Missing Links

#### inal words

#### References

Frame 48/78



Facility location

A reasonable derivation

#### Structure Detection

Hierarchy by division

Hierarchy by shuffling

Hierarchies & Missing Links

#### Single Source

Distributed Sources

Detection

Hierarchy by shuffling

Frame 49/78



"Extracting the hierarchical organization of complex systems"

Consider all partitions of networks into m groups

Sales-Pardo et al., PNAS (2007) [17, 18]

As for Newman and Girvan approach, aim is to find partitions with maximum modularity:

$$Q = \sum_{i} [e_{ii} - (\sum_{j} e_{ij})^{2}] = \text{Tr}\mathbf{E} - ||\mathbf{E}^{2}||_{1}.$$

Single Source

### Distributed Sources

Size-density law
A reasonable derivation
Global redistribution
networks

#### Structure Detection

Hierarchy by shuffling Spectral methods Hierarchies & Missing Links

Final words

References

Frame 50/78



systems"

"Extracting the hierarchical organization of complex

Consider all partitions of networks into m groups

► As for Newman and Girvan approach, aim is to find

 $Q = \sum [e_{ii} - (\sum e_{ij})^2] = \text{Tr}\mathbf{E} - ||\mathbf{E}^2||_1.$ 

Sales-Pardo et al., PNAS (2007) [17, 18]

Single Source

Sources

Frame 50/78



Global redistribution networks
Structure

Detection
Hierarchy by division
Hierarchy by shuffling

Spectral methods
Hierarchies & Missing Links

inal words

References

 "Extracting the hierarchical organization of complex systems"
 Sales-Pardo et al., PNAS (2007) [17, 18]

- Consider all partitions of networks into m groups
- As for Newman and Girvan approach, aim is to find partitions with maximum modularity:

$$Q = \sum_i [e_{ii} - (\sum_j e_{ij})^2] = {
m Tr} {f E} - ||{f E}^2||_1.$$

Frame 50/78



- Consider partition network, i.e., the network of all possible partitions.
- Defn: Two partitions are connected if they differ only by the reassignment of a single node.
- ► Look for local maxima in partition network.
- ▶ Construct an affinity matrix with entries  $A_{ij}$ .
- ▶  $A_{ij}$  = **Pr** random walker on modularity network ends up at a partition with i and j in the same group.
- C.f. topological overlap between i and j = # matching neighbors for i and j divided by maximum of k<sub>i</sub> and k<sub>j</sub>.

Single Source

Distributed Sources

Facility location
Size-density law
A reasonable derivation

Structure Detection

Hierarchy by shuffling Spectral methods

Hierarchies & Missing Links General structure detection

Final words

References



- Consider partition network, i.e., the network of all possible partitions.
- Defn: Two partitions are connected if they differ only by the reassignment of a single node.
- ► Look for local maxima in partition network.
- Construct an affinity matrix with entries A<sub>ij</sub>.
- ▶  $A_{ij}$  = **Pr** random walker on modularity network ends up at a partition with i and j in the same group.
- C.f. topological overlap between i and j = # matching neighbors for i and j divided by maximum of k<sub>i</sub> and k<sub>j</sub>.

Single Source

Distributed Sources

Facility location
Size-density law
A reasonable derivation
Global redistribution

Structure Detection

Hierarchy by shuffling Spectral methods Hierarchies & Missing Links

ieneral structure de

Final words

References



- Consider partition network, i.e., the network of all possible partitions.
- Defn: Two partitions are connected if they differ only by the reassignment of a single node.
- Look for local maxima in partition network.
- Construct an affinity matrix with entries A<sub>ij</sub>.
- ▶  $A_{ij}$  = **Pr** random walker on modularity network ends up at a partition with i and j in the same group.
- C.f. topological overlap between i and j = # matching neighbors for i and j divided by maximum of k<sub>i</sub> and k<sub>j</sub>.

Single Source

Distributed Sources

Size-density law
A reasonable derivation
Global redistribution

Structure Detection

Hierarchy by shuffling Spectral methods Hierarchies & Missing Links

Hierarchies & Missing Links General structure detection

Final words

References



- Consider partition network, i.e., the network of all possible partitions.
- Defn: Two partitions are connected if they differ only by the reassignment of a single node.
- Look for local maxima in partition network.
- Construct an affinity matrix with entries A<sub>ij</sub>.
- ▶  $A_{ij}$  = **Pr** random walker on modularity network ends up at a partition with i and j in the same group.
- C.f. topological overlap between i and j = # matching neighbors for i and j divided by maximum of k<sub>i</sub> and k<sub>j</sub>.

Single Source

Distributed Sources

Facility location
Size-density law
A reasonable derivation
Global redistribution
networks

Structure Detection

Hierarchy by shuffling Spectral methods Hierarchies & Missing Links

ieneral structure de

inal words

References



- Consider partition network, i.e., the network of all possible partitions.
- Defn: Two partitions are connected if they differ only by the reassignment of a single node.
- Look for local maxima in partition network.
- Construct an affinity matrix with entries A<sub>ij</sub>.
- ▶  $A_{ij}$  = **Pr** random walker on modularity network ends up at a partition with i and j in the same group.
- C.f. topological overlap between i and j = # matching neighbors for i and j divided by maximum of k<sub>i</sub> and k<sub>j</sub>.

Single Source

Distributed Sources

Structure

Facility location
Size-density law
A reasonable derivation
Global redistribution

Detection
Hierarchy by division
Hierarchy by shuffling

Spectral methods
Hierarchies & Missing Links
General structure detection

inal words

References



- Consider partition network, i.e., the network of all possible partitions.
- Defn: Two partitions are connected if they differ only by the reassignment of a single node.
- Look for local maxima in partition network.
- Construct an affinity matrix with entries A<sub>ij</sub>.
- ▶  $A_{ij}$  = **Pr** random walker on modularity network ends up at a partition with i and j in the same group.
- C.f. topological overlap between i and j = # matching neighbors for i and j divided by maximum of k<sub>i</sub> and k<sub>j</sub>.

Single Source

Distributed Sources

Facility location Size-density law A reasonable derivation Global redistribution

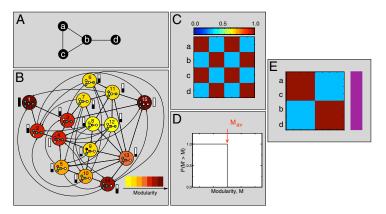
Structure Detection Hierarchy by division

Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

nal words

References





A: Base network; B: Partition network; C: Coclassification matrix; D: Comparison to random networks (all the same!); E: Ordered coclassification matrix; Single Source

Distributed Sources

Facility location

A reasonable derivation Global redistribution networks

Structure Detection

Hierarchy by division
Hierarchy by shuffling

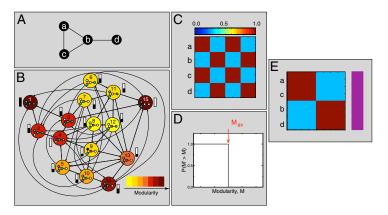
Hierarchies & Missing Links
General structure detection

inal words

References

Frame 52/78





A: Base network; B: Partition network; C: Coclassification matrix; D: Comparison to random networks (all the same!); E: Ordered coclassification matrix; Conclusion: no structure... Single Source

Distributed Sources

Facility location

A reasonable derivation Global redistribution networks

Structure Detection

Hierarchy by division Hierarchy by shuffling

Spectral methods
Hierarchies & Missing Links
General structure detection

Final words

References

Frame 52/78



### Method obtains a distribution of classification hierarchies.

- Note: the hierarchy with the highest modularity score isn't chosen.
- Idea is to weight possible hierarchies according to their basin of attraction's size in the partition network.
- Next step: Given affinities, now need to sort nodes into modules, submodules, and so on.
- Idea: permute nodes to minimize following cost

$$C = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} A_{ij} |i - j|.$$

► Use simulated annealing (slow).

Single Source

Distributed Sources

Size-density law
A reasonable derivation
Global redistribution

Detection
Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Link:

inal words

References





- Method obtains a distribution of classification hierarchies.
- Note: the hierarchy with the highest modularity score isn't chosen.
- Idea is to weight possible hierarchies according to their basin of attraction's size in the partition network.
- Next step: Given affinities, now need to sort nodes into modules, submodules, and so on.
- Idea: permute nodes to minimize following cost

$$C = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} A_{ij} |i - j|.$$

► Use simulated annealing (slow).

#### Single Source

Distributed Sources

Sources
Facility location

A reasonable derivat Global redistribution networks

Detection

Hierarchy by division

Hierarchy by shuffling

Spectral methods

Spectral methods Hierarchies & Missing Links General structure detection

Final words

References





- Method obtains a distribution of classification. hierarchies
- Note: the hierarchy with the highest modularity score isn't chosen.
- Idea is to weight possible hierarchies according to their basin of attraction's size in the partition network.
- Next step: Given affinities, now need to sort nodes
- Idea: permute nodes to minimize following cost

$$C = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} A_{ij} |i - j|.$$

▶ Use simulated annealing (slow).

#### Single Source

### Distributed

Sources

## Detection

Hierarchy by shuffling





- Note: the hierarchy with the highest modularity score isn't chosen.
- Idea is to weight possible hierarchies according to their basin of attraction's size in the partition network.
- Next step: Given affinities, now need to sort nodes into modules, submodules, and so on.
- Idea: permute nodes to minimize following cost

$$C = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} A_{ij} |i - j|.$$

Use simulated annealing (slow).

#### Single Source

Distributed Sources

Size-density law
A reasonable derivation

Structure
Detection
Hierarchy by division
Hierarchy by shuffling

Spectral methods
Hierarchies & Missing Links

inal words

References

Frame 53/78



## Shuffling for structure

- Method obtains a distribution of classification hierarchies.
- Note: the hierarchy with the highest modularity score isn't chosen.
- Idea is to weight possible hierarchies according to their basin of attraction's size in the partition network.
- Next step: Given affinities, now need to sort nodes into modules, submodules, and so on.
- Idea: permute nodes to minimize following cost

$$C = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} A_{ij} |i-j|.$$

▶ Use simulated annealing (slow).

#### Single Source

#### Distributed Sources

Facility location
Size-density law
A reasonable derivation

#### Structure Detection

Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

#### Final words

References





## Shuffling for structure

- Method obtains a distribution of classification hierarchies.
- Note: the hierarchy with the highest modularity score isn't chosen.
- Idea is to weight possible hierarchies according to their basin of attraction's size in the partition network.
- Next step: Given affinities, now need to sort nodes into modules, submodules, and so on.
- Idea: permute nodes to minimize following cost

$$C = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} A_{ij} |i - j|.$$

Use simulated annealing (slow).

#### Single Source

#### Distributed Sources

Facility location
Size-density law
A reasonable derivation

#### Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods

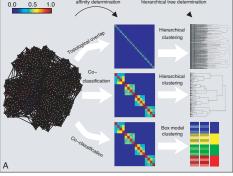
#### Charles and a

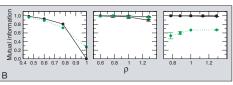
References











- N = 640
- $ightharpoonup \langle k \rangle = 16,$
- 3 tiered hierarchy.

#### Distributed Sources

### Detection

- Hierarchy by shuffling

Frame 54/78

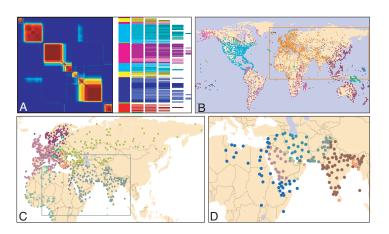


Detection

Hierarchy by shuffling

Frame 55/78





Modules found match up with geopolitical units.

Facility location

A reasonable derivation

### Structure Detection

Hierarchy by division Hierarchy by shuffling

Spectral methods

Hierarchies & Missing Links

Single Source

Distributed Sources

Detection

Spectral methods

Frame 56/78





### Distributed Sources

Facility location
Size-density law
A reasonable derivation
Global redistribution

### Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods

Hierarchies & Missing Links General structure detection

#### inal words

References

- "Detecting communities in large networks" Capocci et al. (2005) [4]
- ► Consider normal matrix  $K^{-1}A$ , random walk matrix  $A^{T}K^{-1}$ , Laplacian K A, and  $AA^{T}$ .
- Basic observation is that eigenvectors associated with secondary eigenvalues reveal evidence of structure.
- ▶ Build on Kleinberg's HITS algorithm. [13]



Global redistributi networks

Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Link

General structure

Finai words

References



- "Detecting communities in large networks" Capocci et al. (2005) [4]
- Consider normal matrix  $K^{-1}A$ , random walk matrix  $A^{T}K^{-1}$ , Laplacian K A, and  $AA^{T}$ .
- Basic observation is that eigenvectors associated with secondary eigenvalues reveal evidence of structure.
- Build on Kleinberg's HITS algorithm. [13]

## Detection

Spectral methods

- "Detecting communities in large networks" Capocci et al. (2005) [4]
- ▶ Consider normal matrix K<sup>-1</sup>A, random walk matrix  $A^{\mathrm{T}}\mathbf{K}^{-1}$ , Laplacian  $\mathbf{K} - \mathbf{A}$ , and  $AA^{\mathrm{T}}$ .
- Basic observation is that eigenvectors associated with secondary eigenvalues reveal evidence of structure.
- ▶ Build on Kleinberg's HITS algorithm. [13]



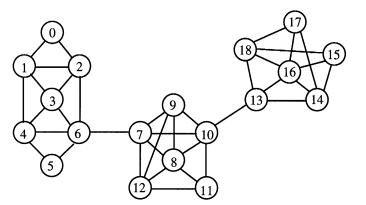
Detection

Spectral methods

- "Detecting communities in large networks" Capocci et al. (2005) [4]
- ▶ Consider normal matrix K<sup>-1</sup>A, random walk matrix  $A^{\mathrm{T}}\mathbf{K}^{-1}$ , Laplacian  $\mathbf{K} - \mathbf{A}$ , and  $AA^{\mathrm{T}}$ .
- Basic observation is that eigenvectors associated with secondary eigenvalues reveal evidence of structure.
- ▶ Build on Kleinberg's HITS algorithm. [13]



### Example network:



#### Single Source

#### Distributed Sources

## Detection

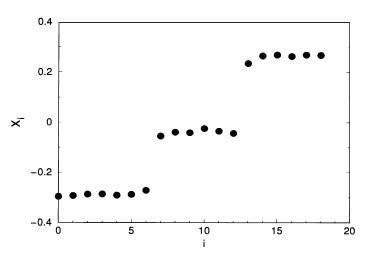
Spectral methods

Frame 58/78





Second eigenvector's components:



#### Single Source

#### Distributed Sources

Facility location
Size-density law
A reasonable derivation

#### Structure Detection

Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

Einal word

References

Frame 59/78



### Distributed Sources

Facility location

Size-density law

A reasonable derivation

Global redistribution networks

### Structure Detection

Hierarchy by division

Hierarchy by shuffling

Hierarchies & Missing Links

General structure detection

Final words

References

Single Source

Distributed Sources

Facility location

A reasonable derivation

Global redistrib

Structure

Detection

Hierarchy by shuffling Spectral methods

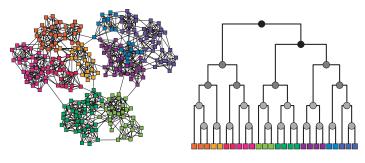
Hierarchies & Missing Links General structure detection

Final words

References

Frame 60/78





- Idea: Shades indicate probability that nodes in left and right subtrees of dendogram are connected.
- ► Handle: Hierarchical random graph models.
- ▶ Plan: Infer consensus dendogram for a given real network.
- Obtain probability that links are missing (big problem...).

#### Single Source

#### Distributed Sources

Facility location

A reasonable derivation Global redistribution networks

#### Structure Detection

Hierarchy by divisio

Spectral methods
Hierarchies & Missing Links

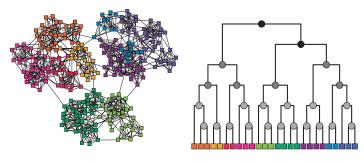
General structure detection

#### ınaı words

References







- Idea: Shades indicate probability that nodes in left and right subtrees of dendogram are connected.
- Handle: Hierarchical random graph models.
- ▶ Plan: Infer consensus dendogram for a given real
- Obtain probability that links are missing (big

Single Source

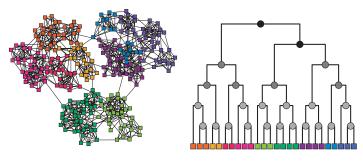
#### Distributed Sources

### Detection

Hierarchies & Missing Links







- Idea: Shades indicate probability that nodes in left and right subtrees of dendogram are connected.
- Handle: Hierarchical random graph models.
- ▶ Plan: Infer consensus dendogram for a given real network.
- Obtain probability that links are missing (big

Optimal supply & Structure detection

Single Source

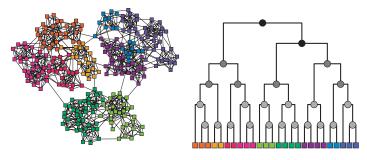
#### Distributed Sources

### Detection

Hierarchies & Missing Links







- Idea: Shades indicate probability that nodes in left and right subtrees of dendogram are connected.
- Handle: Hierarchical random graph models.
- ▶ Plan: Infer consensus dendogram for a given real network.
- Obtain probability that links are missing (big problem...).

Optimal supply & Structure detection

Single Source

## Distributed Sources

Facility location

A reasonable derivation Global redistribution networks

#### Structure Detection

Detection Hierarchy by divisi

Spectral methods
Hierarchies & Missing Links

General structure detection

#### inal words

References



- Model also predicts reasonably well
  - average degree,
  - 2. clustering,
  - 3. and average shortest path length.

#### Table 1 | Comparison of original and resampled networks

Network	$\langle k \rangle_{\rm real}$	$\langle k \rangle_{\rm samp}$	$C_{\text{real}}$	$C_{samp}$	$d_{\rm real}$	$d_{samp}$
T. pallidum	4.8	3.7(1)	0.0625		3.690	3.940(6)
Terrorists	4.9	5.1(2)	0.361	0.352(1)	2.575	2.794(7)
Grassland	3.0	2.9(1)	0.174	0.168(1)	3.29	3.69(2)

Statistics are shown for the three example networks studied and for new networks generated by resampling from our hierarchical model. The generated networks closely match the average degree  $\langle k \rangle$ , clustering coefficient C and average vertex–vertex distance d in each case, suggesting that they capture much of the structure of the real networks. Parenthetical values indicate standard errors on the final digits.

#### Single Source

### Distributed Sources

Facility location

A reasonable derivation Global redistribution

### Structure

Detection

Hierarchy by shuffling Spectral methods

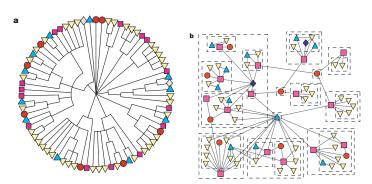
Hierarchies & Missing Links
General structure detection

#### inal word

References

Frame 62/78





- Consensus dendogram for grassland species.
- Copes with disassortative and assortative communities.

#### Distributed Sources

- Facility location
- A reasonable derivation Global redistribution networks

#### Structure Detection

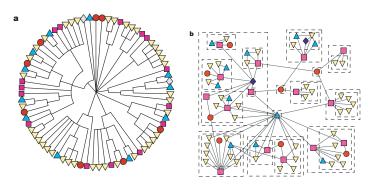
- Detection
- Hierarchy by shuffling Spectral methods
- Hierarchies & Missing Links General structure detection

#### Final words

#### References

Frame 63/78





- Consensus dendogram for grassland species.
- Copes with disassortative and assortative communities.

#### Distributed Sources

- Facility location
- A reasonable derivation Global redistribution
- Structure

### Detection

- Hierarchy by shuffling Spectral methods
- Hierarchies & Missing Links
- Einal worde

#### Finai words

#### References

Frame 63/78



### Distributed Sources

Facility location
Size-density law
A reasonable derivation

Global redistribution networks

### Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links
General structure detection

Final words

References

Single Source

Distributed Sources

Facility location Size-density law

A reasonable derivation Global redistribution networks

Structure Detection

Hierarchy by division
Hierarchy by shuffling

Hierarchies & Missing Links
General structure detection

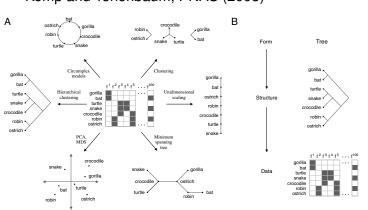
Final words

References

Frame 64/78



# ► "The discovery of structural form" Kemp and Tenenbaum, PNAS (2008) [12]



#### Single Source

#### Distributed Sources

acility location

A reasonable derivation Global redistribution networks

#### Structure Detection

Hierarchy by division
Hierarchy by shuffling

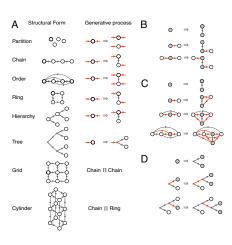
Hierarchies & Missing Links
General structure detection

Final words

References

Frame 65/78





- Top down description of form.
- Node replacement graph grammar: parent node becomes two
- ▶ B-D: Growing chains, orders, and trees.

#### Single Source

#### Distributed Sources

Facility location
Size-density law
A reasonable derivation
Global redistribution

#### Structure Detection

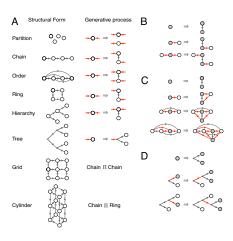
Hierarchy by shuffling Spectral methods Hierarchies & Missing Links General structure detection

#### Final words

References

Frame 66/78





- Top down description of form.
- Node replacement graph grammar: parent node becomes two child nodes.
- B-D: Growing chains, orders, and trees.

#### Single Source

#### Distributed Sources

Facility location
Size-density law
A reasonable derivation

#### Structure Detection

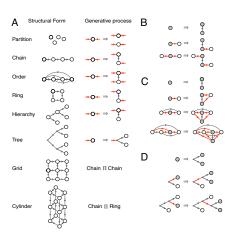
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links
General structure detection

#### Final words

References

Frame 66/78





- Top down description of form.
- Node replacement graph grammar: parent node becomes two child nodes.
- B-D: Growing chains, orders, and trees.

#### Single Source

#### Distributed Sources

Facility location
Size-density law
A reasonable derivation

#### Structure Detection

Hierarchy by division

Hierarchy by shuffling

Spectral methods

Hierarchies & Missing Links

General structure detection

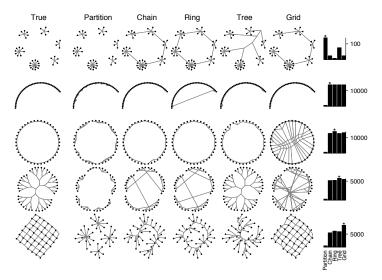
#### Final words

References

Frame 66/78



Performance for test networks.



Single Source

#### Distributed Sources

Facility location

A reasonable derivation Global redistribution networks

#### Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods

Hierarchies & Missing Links General structure detection

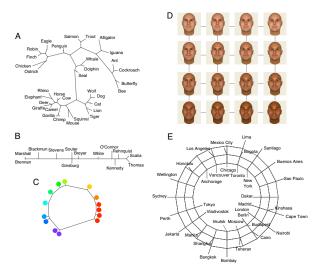
#### inal word

References

Frame 67/78



## Example learned structures:



▶ Biological features; Supreme Court votes; perceived color differences; face differences; & distances between cities. Optimal supply & Structure detection

Single Source

Distributed Sources

Facility location

A reasonable derivation Global redistribution networks

Structure Detection

Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links
General structure detection

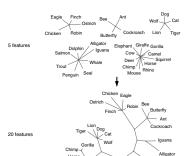
Cinal words

. ......

References

Frame 68/78





Rhino Giraffe Squirrel Dolphin

Eagle Penguin Trout

Iguana

Cockroach

Butterfly

Salmon

Penguin

Whale Dolphin

~ Cat

Salmon Trout

Seal

Horse

Cow -

Robin

Finch

Ostrich

Flenhant\*

Giraffe Camel Chimp Mouse

110 features

Effect of adding features on detected form.

Single Source

Distributed Sources

Detection

General structure detection

Frame 69/78





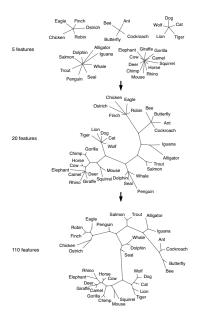
### Detection

General structure detection

Distributed Sources

Frame 69/78





Effect of adding features on detected form.

Straight partition simple tree complex tree

A reasonable derivation Global redistribution networks

#### Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods

Hierarchies & Missing Link General structure detection

#### Final words

References

riciciences

### Science in three steps:

- 1. Find interesting/meaningful/important phenomena involving spectacular amounts of data.
- Describe what you see.
- 3. Explain it.

### A plea/warning

Beware your assumptions—don't use tools/models because they're there, or because everyone else does...

Frame 70/78



## Distributed Sources

Size-density law

A reasonable derivation Global redistribution networks

#### Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods

Hierarchies & Missing Link General structure detection

#### Final words

References

### Science in three steps:

- 1. Find interesting/meaningful/important phenomena involving spectacular amounts of data.
- 2. Describe what you see.
- 3. Explain it.

### A plea/warning

Beware your assumptions—don't use tools/models because they're there, or because everyone else does...

Frame 70/78



Detection

Final words

### Science in three steps:

- 1. Find interesting/meaningful/important phenomena involving spectacular amounts of data.
- Describe what you see.
- 3. Explain it.

Frame 70/78





Structure Detection

Final words

### Frame 70/78



### Science in three steps:

- 1. Find interesting/meaningful/important phenomena involving spectacular amounts of data.
- Describe what you see.
- 3. Explain it.

### A plea/warning

Beware your assumptions—don't use tools/models because they're there, or because everyone else does...

## Distributed Sources

Facility location Size-density law

A reasonable deriva Global redistribution networks

#### Structure Detection

Detection

Hierarchy by divi

Spectral methods

Hierarchies & Missing Link

Hierarchies & Missing Links General structure detection

#### Final words

References

Frame 71/78



### A real theory of everything:

- 1. Is not just about the small stuff..
- 2. It's about the increase of complexity

## Distributed Sources

Size-density law
A reasonable derivation

Global redis networks

#### Structure Detection

Hierarchy by division Hierarchy by shuffling

Hierarchies & Missing Links

<u>-</u>. . .

#### Final words

References

#### Frame 71/78



### A real theory of everything:

- 1. Is not just about the small stuff...
- It's about the increase of complexity

#### Distributed Sources

### Structure

Detection

### Final words

#### Frame 71/78



### A real theory of everything:

- 1. Is not just about the small stuff...
- 2. It's about the increase of complexity

#### Distributed Sources

Detection

#### Final words

### A real theory of everything:

- 1. Is not just about the small stuff...
- 2. It's about the increase of complexity

Symmetry breaking/ VS. Accidents of history

Universality

Frame 71/78



## Distributed Sources

Facility location Size-density law

A reasonable derivation Global redistribution networks

### Structure

Detection

Hierarchy by shuffling
Spectral methods

General structure detecti

#### Final words

References

References

### A real theory of everything:

- 1. Is not just about the small stuff...
- 2. It's about the increase of complexity

Symmetry breaking/ Accidents of history vs. Universality

How probable is a certain level of complexity?

Frame 71/78



Detection

References

Frame 72/78



[1] J. R. Banavar, F. Colaiori, A. Flammini, A. Maritan, and A. Rinaldo.

Topology of the fittest transportation network.

Phys. Rev. Lett., 84:4745–4748, 2000. pdf (⊞)

- [2] J. R. Banavar, A. Maritan, and A. Rinaldo. Size and form in efficient transportation networks. Nature, 399:130–132, 1999. pdf (⊞)
- [3] S. Bohn and M. O. Magnasco. Structure, scaling, and phase transition in the optimal transport network.

Phys. Rev. Lett., 98:088702, 2007. pdf (⊞)

[4] A. Capocci, V. Servedio, G. Caldarelli, and F. Colaiori.

Detecting communities in large networks.

Physica A: Statistical Mechanics and its Applications, 352:669–676, 2005. pdf (⊞)

[5] A. Clauset, C. Moore, and M. E. J. Newman. Hierarchical structure and the prediction of missing links in networks.

Nature, 453:98-101, 2008. pdf (⊞)

[6] P. S. Dodds and D. H. Rothman. Unified view of scaling laws for river networks. Physical Review E, 59(5):4865–4877, 1999. pdf (⊞) Single Source

Distributed Sources

Facility location
Size-density law
A reasonable derivatio

A reasonable derivation Global redistribution networks

Structure Detection

Hierarchy by division Hierarchy by shuffling Spectral methods Hierarchies & Missing Link

inal words

References

Frame 73/78



and deviations.

[7] P. S. Dodds and D. H. Rothman. Geometry of river networks. I. Scaling, fluctuations,

*Physical Review E*, 63(1):016115, 2001. pdf (⊞)

- [8] M. T. Gastner and M. E. J. Newman.
  Optimal design of spatial distribution networks.

  Phys. Rev. E, 74:Article # 016117, 2006. pdf (H)
- [9] M. Girvan and M. E. J. Newman. Community structure in social and biological networks.

Proc. Natl. Acad. Sci., 99:7821-7826, 2002. pdf (⊞)

[10] S. M. Gusein-Zade. Geogr. Anal., 14:246–, 1982. Single Source

Distributed Sources

Size-density law
A reasonable derivation

A reasonable derivat Global redistribution networks

Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

References

Frame 74/78



[11] A. A. Heusner.

Size and power in mammals.

Journal of Experimental Biology, 160:25–54, 1991.

[12] C. Kemp and J. B. Tenenbaum.

The discovery of structural form.

Proc. Natl. Acad. Sci., 105:10687-10692, 2008.  $pdf(\boxplus)$ 

[13] J. M. Kleinberg.

Authoritative sources in a hyperlinked environment.

Proc. 9th ACM-SIAM Symposium on Discrete Algorithms, 1998. pdf  $(\boxplus)$ 

[14] M. E. J. Newman.

Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality.

*Phys. Rev. E*, 64(1):016132, 2001. pdf (⊞)

Single Source

Distributed Sources

Structure Detection

References

Frame 75/78



### References V

[15] M. E. J. Newman.

Erratum: Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality [Phys. Rev. E 64, 016132 (2001)].

Phys. Rev. E, 73:039906(E), 2006.  $\underline{pdf}$  ( $\boxplus$ )

[16] M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks.

*Phys. Rev. E*, 69(2):026113, 2004. <u>pdf</u> (⊞)

[17] M. Sales-Pardo, R. Guimerà, A. A. Moreira, and L. A. N. Amaral.

Extracting the hierarchical organization of complex systems.

*Proc. Natl. Acad. Sci.*, 104:15224–15229, 2007. pdf (⊞)

#### Single Source

Distributed Sources

Facility location Size-density law A reasonable derivatio Global redistribution

Structure Detection

Hierarchy by division
Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links
General structure detection

inal words

References

Frame 76/78



[18] M. Sales-Pardo, R. Guimerà, A. A. Moreira, and I A N Amaral

Extracting the hierarchical organization of complex systems: Correction.

Proc. Natl. Acad. Sci., 104:18874, 2007. pdf (⊞)

[19] G. E. Stephan.

Territorial division: The least-time constraint behind the formation of subnational boundaries.

Science, 196:523–524, 1977. pdf (⊞)

[20] G. E. Stephan. Territorial subdivision. Social Forces, 63:145–159, 1984. pdf (⊞)

[21] D. L. Turcotte, J. D. Pelletier, and W. I. Newman. Networks with side branching in biology. Journal of Theoretical Biology, 193:577–592, 1998.

Single Source

Distributed Sources

Structure Detection

References

Frame 77/78



[22] G. B. West, J. H. Brown, and B. J. Enquist.

A general model for the origin of allometric scaling laws in biology.

Science, 276:122-126, 1997. pdf (⊞)

[23] Q. Xia.

The formation of a tree leaf.

Submitted. pdf (⊞)

[24] Q. Xia.

Optimal paths related to transport problems.

Communications in Contemporary Mathematics, 5:251–279, 2003. pdf (⊞)

[25] W. W. Zachary.

An information flow model for conflict and fission in small groups.

J. Anthropol. Res., 33:452-473, 1977.

Single Source

Distributed Sources

> Facility location Size-density law

A reasonable derivation
Global redistribution
networks

Structure Detection

Hierarchy by shuffling
Spectral methods
Hierarchies & Missing Links

Final words

References

Frame 78/78

