

Socio-semantic Networks

Socio-semantic frameworks for techno-social systems

Camille Roth

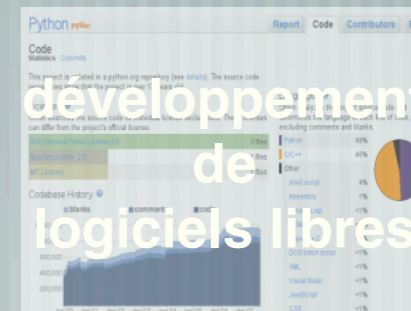
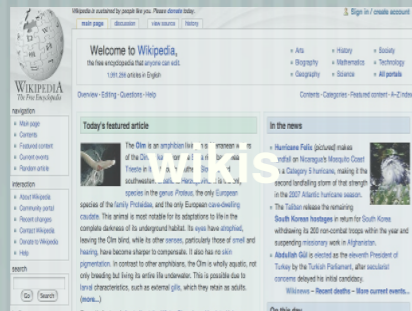
CNRS / EHESS

CAMS (Centre d'Analyse et de Mathématique Sociales)

ISC-PIF (Institut des Systèmes Complexes de Paris Ile-de-France)

Socio-technical systems?

webloggers, communities of scientists, software developers and wiki contributors



more broadly: socio-semantic systems involving agents creating and processing knowledge, exchanging information connecting concepts in a distributed manner...

Social cognition?

Not immediately related to cognitive psychology...

...rather, "information production and processing in a system of a (generally) large number of individuals"

Social cognition?

Not immediately related to cognitive psychology...

...rather, "information production and processing in a system of a (generally) large number of individuals"

- 1. Social and applied epistemology**
- 2. Cultural anthropology**
- 3. Social computing**

Social cognition?

— [social factors influencing individual knowledge

— [organization of cognitive labor (distributed cognition, e.g. scientific communities)

— [notion of collective knowledge

1. **Social and applied epistemology**
2. Cultural anthropology
3. Social computing

Social cognition?

"Culture is acquired information, such as knowledge, beliefs, and values, that is inherited through social learning, and expressed in behaviors and artifacts."
(Mesoudi, Whiten & Laland, 2004)

"(...) explaining the capacity of some representations to propagate until becoming precisely cultural, that is, revealing the reasons of their contagiousity."
(Lenclud, 1998)



1. Social and applied epistemology
2. Cultural anthropology
3. Social computing

3. Social computing

— [“socio-informatics”

essentially from large datasets of *in vivo* human behavior

- from government agencies
(public health, economics, bibliographical records, ...)
- from companies on consumer behavior
(supermarkets, transit networks, cell phone, ...)
- from online services in various contexts
(emails, discussion forums, wikis, blogs, ...)

3. Social computing

“socio-informatics”

essentially from large datasets of *in vivo* human behavior

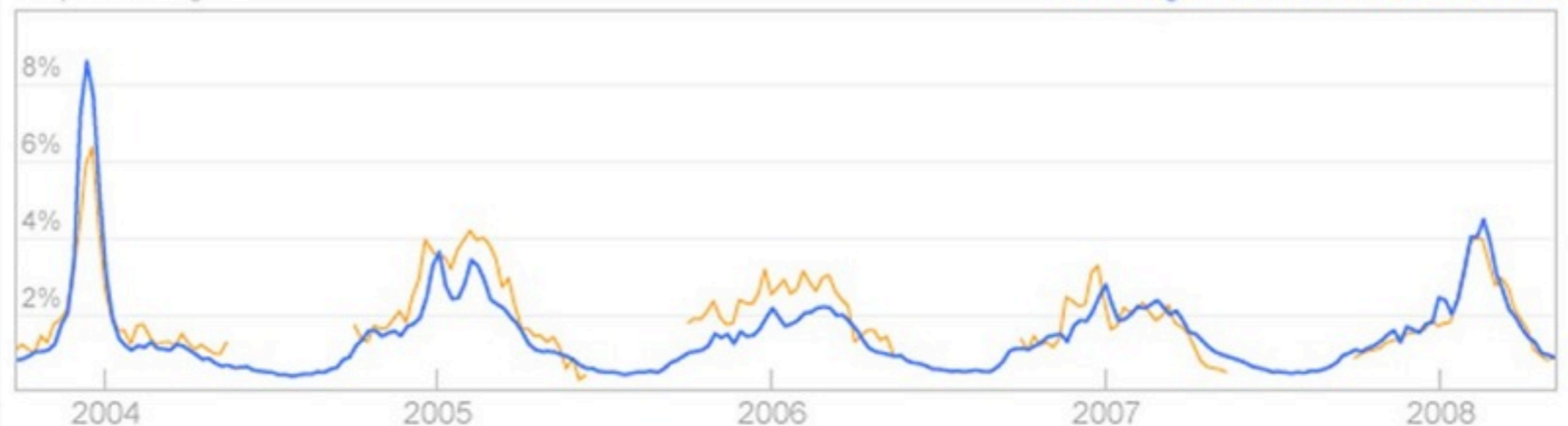
social
sensing

(Google FluTrends, 2008)

Annual U.S. Flu Activity - Mid-Atlantic Region

ILI percentage

● Google Flu Trends ● CDC Data



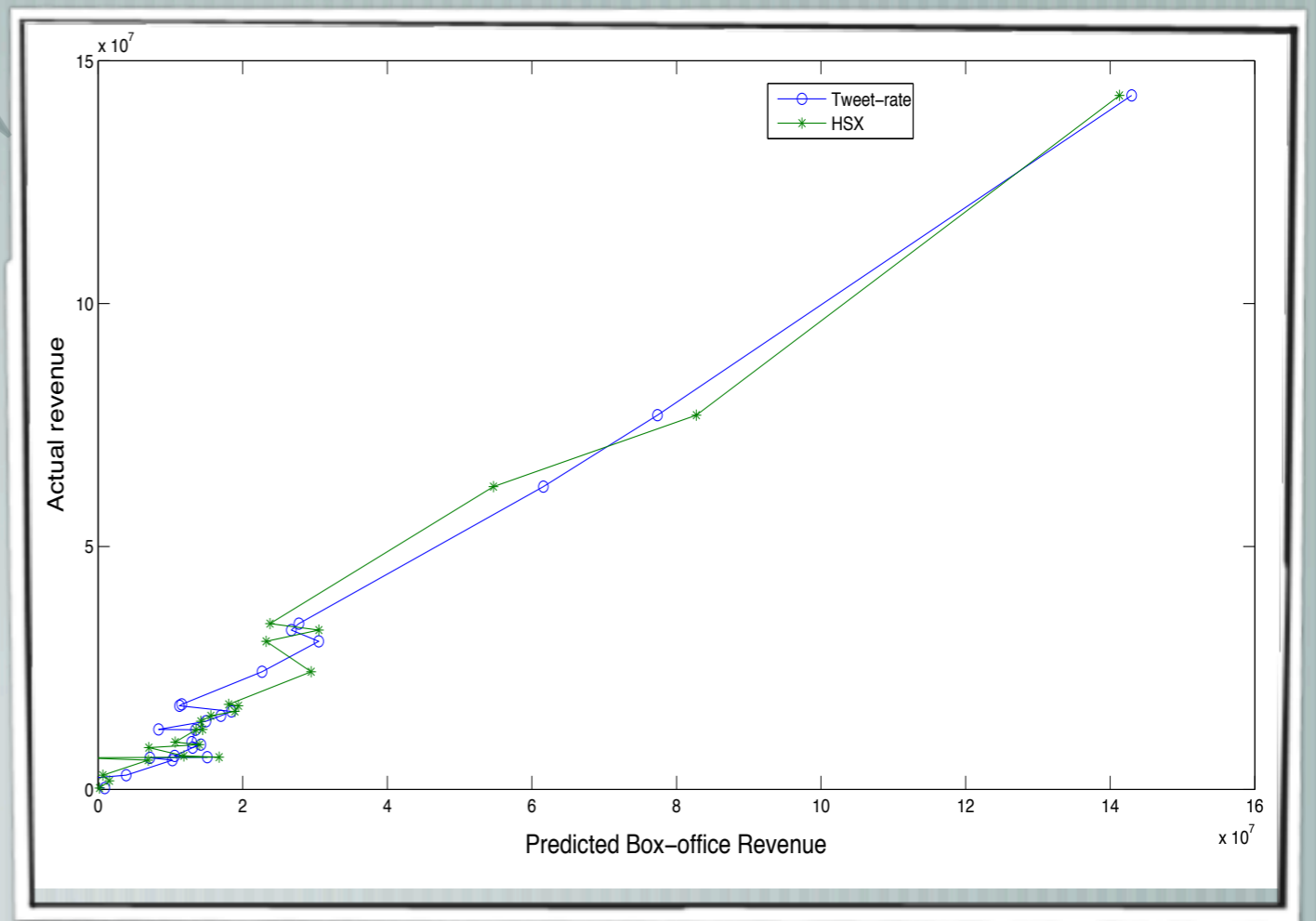
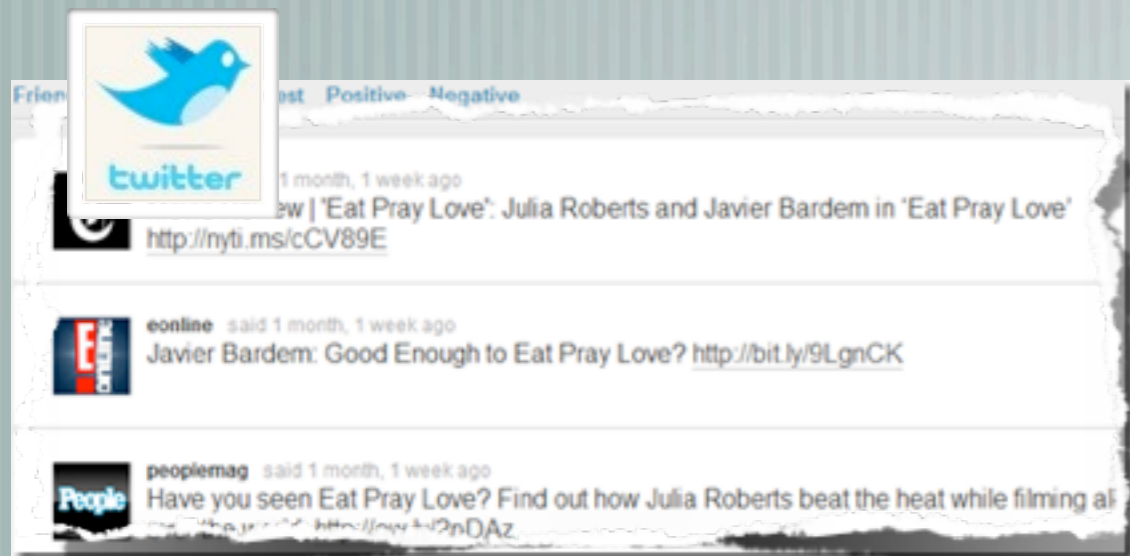
3. Social computing

“socio-informatics”

essentially from large datasets of *in vivo* human behavior

(Asur, Huberman, 2010)

social
sensing



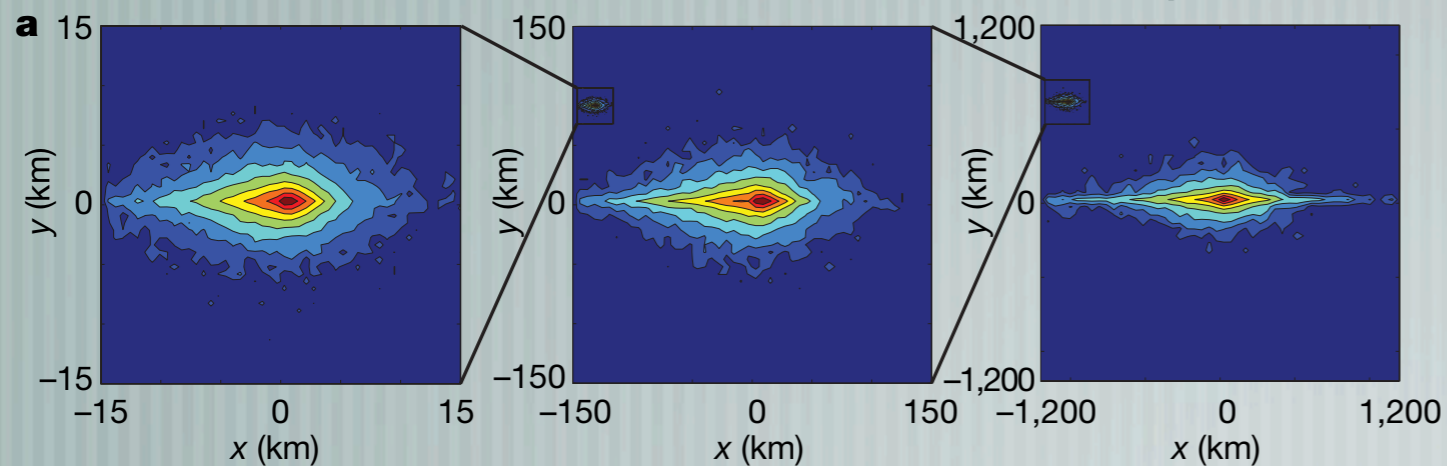
3. Social computing

“socio-informatics”

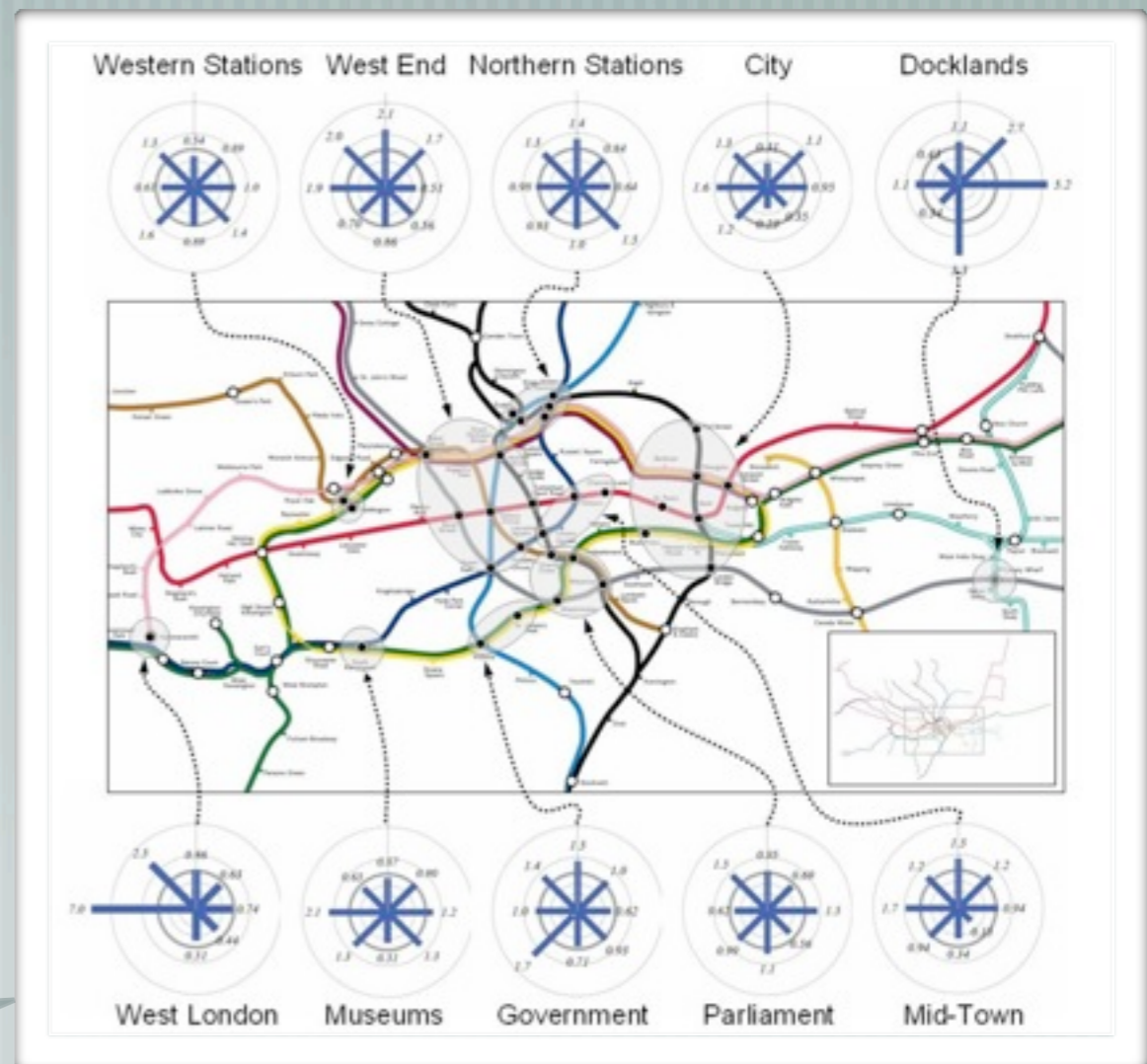
essentially from large datasets of *in vivo* human behavior

description of human dynamics

(Gonzalez, Hidalgo, Barabasi, 2008)



(Roth, Kang, Batty, Barthelemy, 2011)



3. Social computing

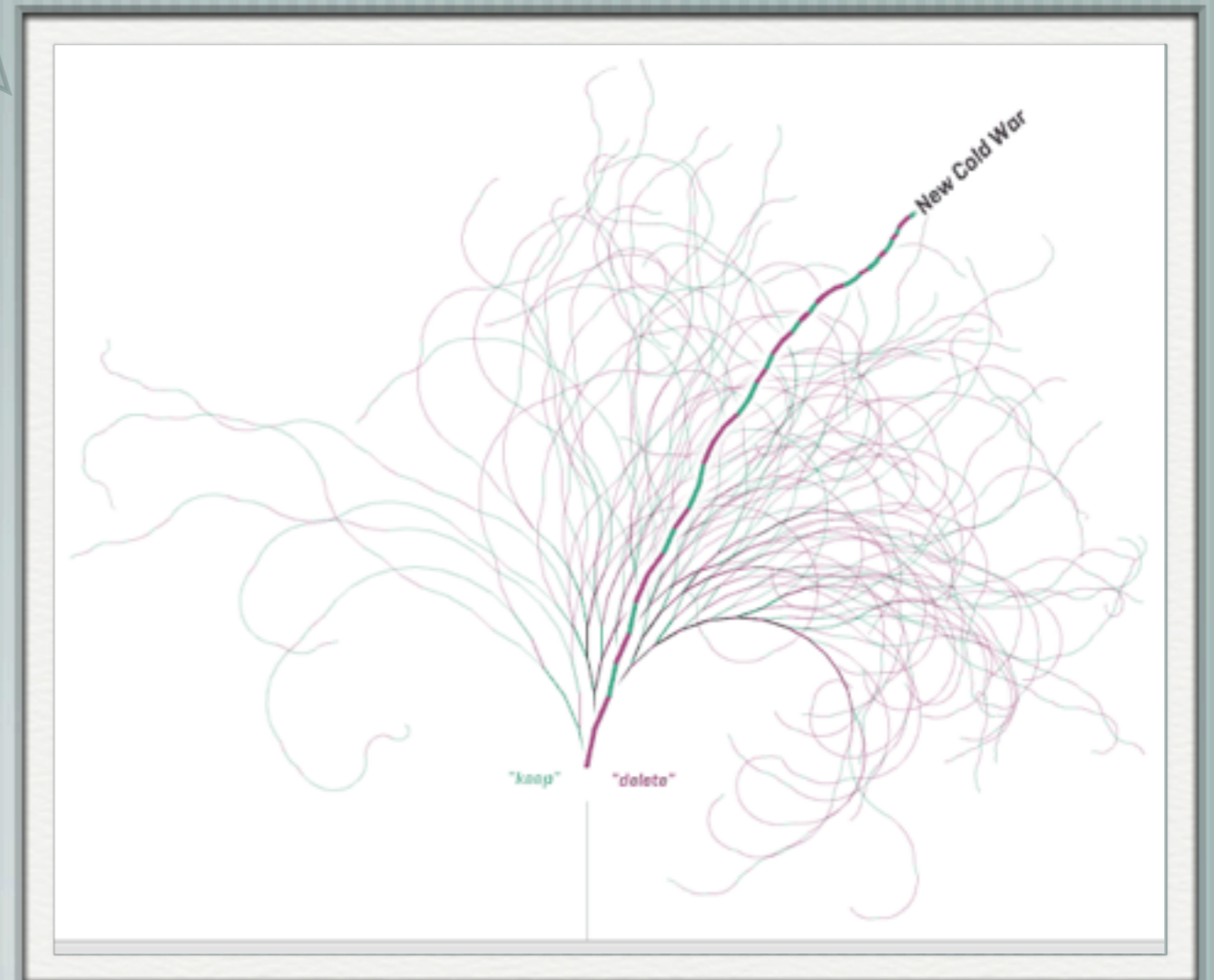
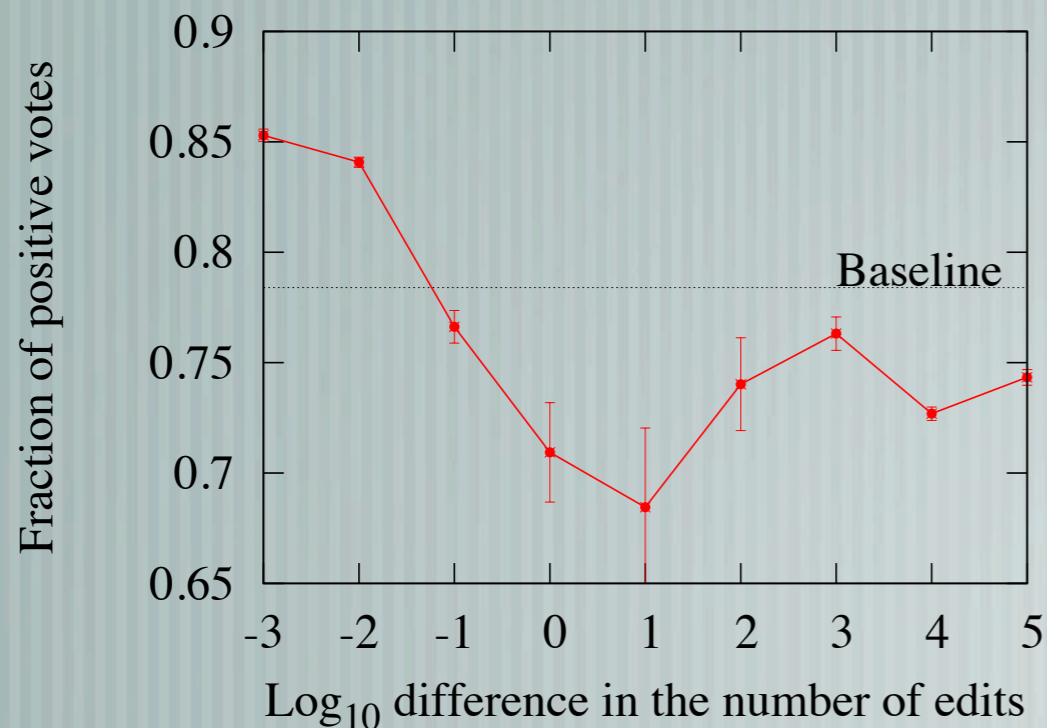
“socio-informatics”

essentially from large datasets of *in vivo* human behavior

description of collective decision processes

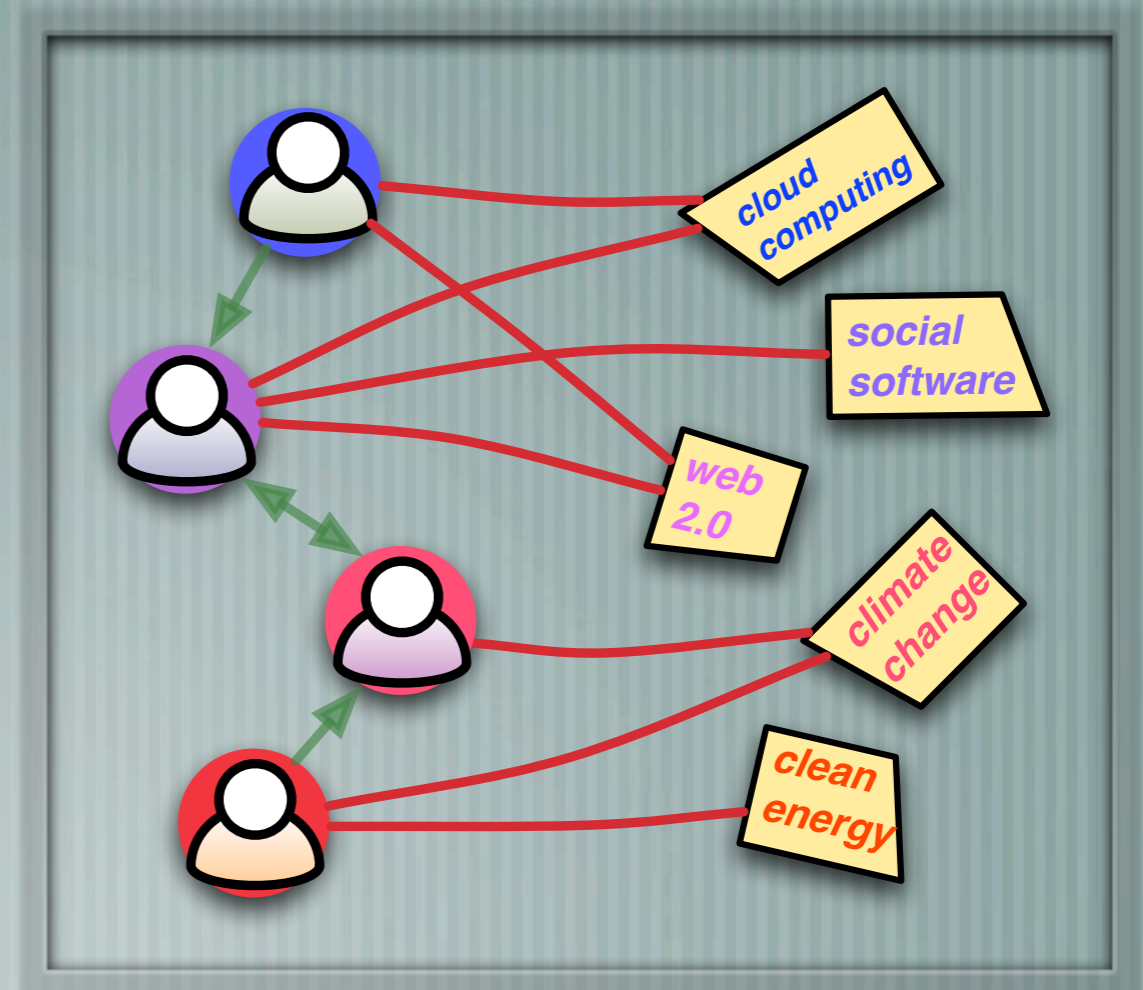
(notabilia.net)

(Leskovec, Huttenlocher, Kleinberg, 2010)



What is needed for an experimental science of cultural dynamics?

- Knowing the shape of social interactions
- Knowing the dynamics of content and being able to describe “cultural items”



Social networks

— [First period of development: 40s-70s

— mathematical sociology and social science

— focused on “small” case-studies,
algebraic definitions



Social networks

— [First period of development: 40s-70s

— mathematical sociology and social science

— focused on “small” case-studies,
algebraic definitions

— [Second period: the new science of networks, end of 90s-now

— large-scale datasets, complex systems standpoint

— notion of “scale-free, small-world” networks,
distinct from random networks

— social networks as a key case: web pages, collaboration, ...



Social networks

First period of development: 40s-70s

— mathematical sociology and social science

— focused on “small” case-studies,
algebraic definitions

Second period: the new science of networks, end of 90s-now

— large-scale datasets, complex systems standpoint

— notion of “scale-free, small-world” networks,
distinct from random networks

— social networks as a key case: web pages, collaboration, ...



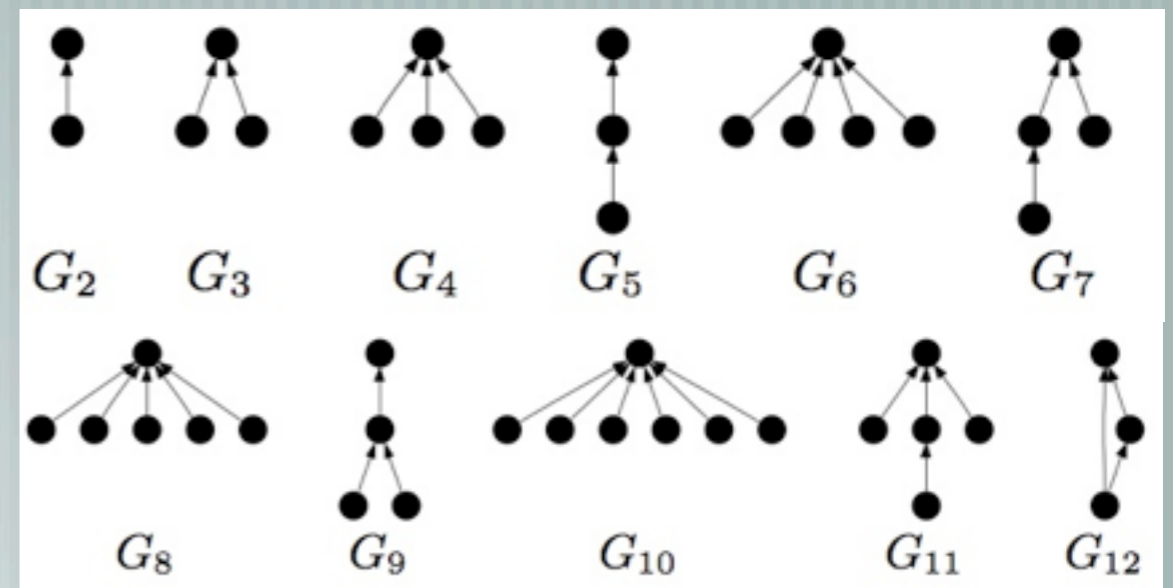
Social networks: blogs

Classical stylized facts:

power-law, topological communities, transitivity, patterns...

Morphogenesis models

random, agent-based models based on posting behavior



(McGlohon, Leskovec, Faloutsos, Hurst, Glance, 2007)

Social networks: blogs

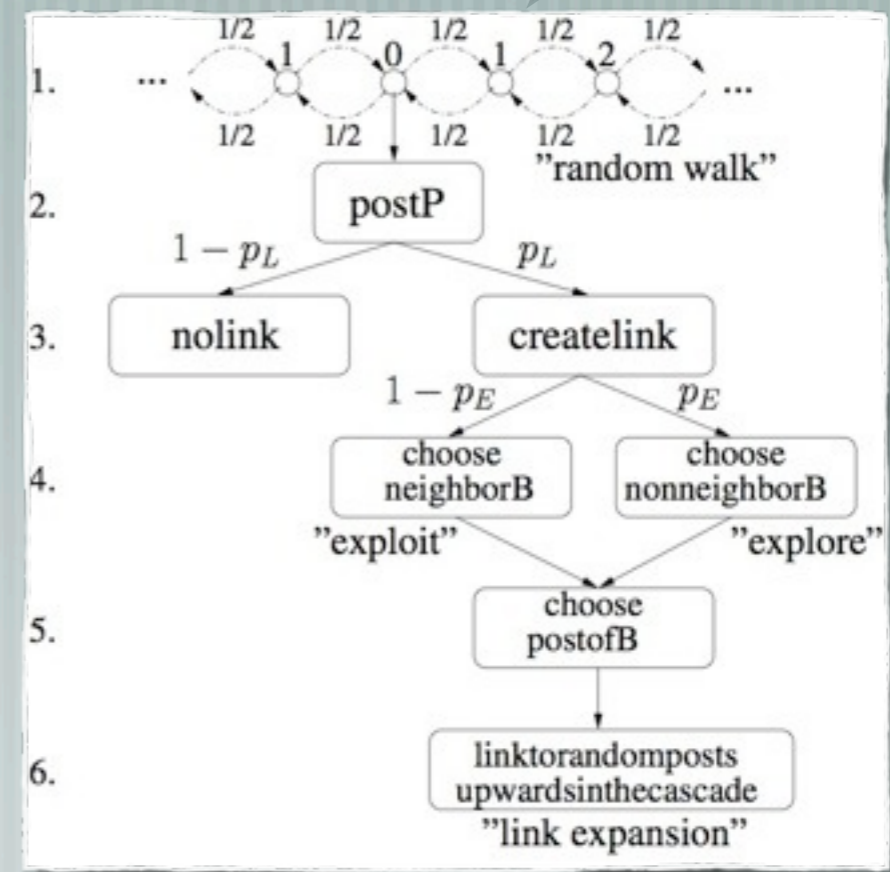
Classical stylized facts:

power-law, topological communities, transitivity, patterns...

Morphogenesis models

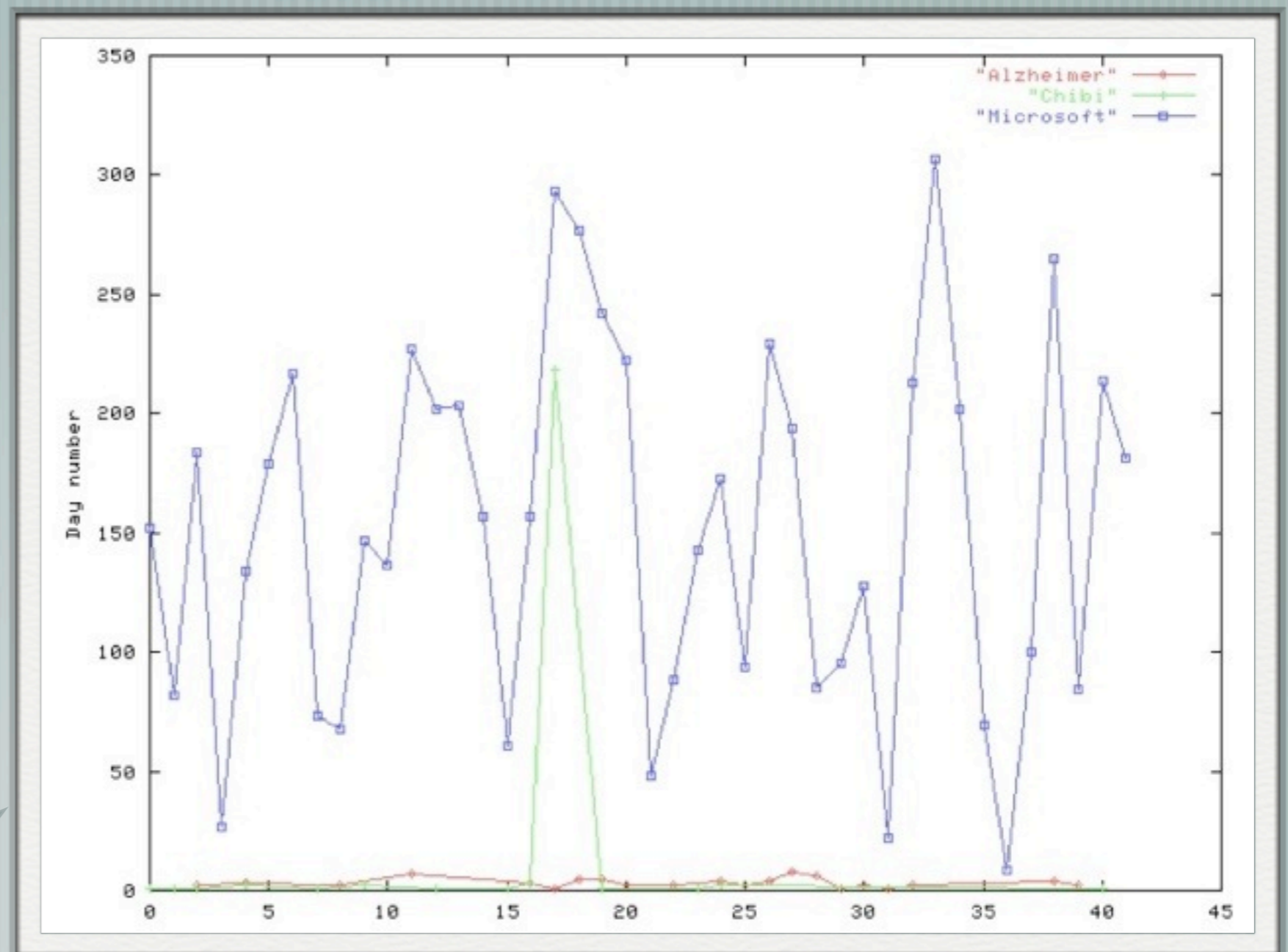
random, agent-based models based on posting behavior

(Gotz, Leskovec, McGlohon, Faloutsos, 2009)



Dynamics of conversations

Dynamics of discussions



(Gruhl et al., 2004)

Dynamics of conversations

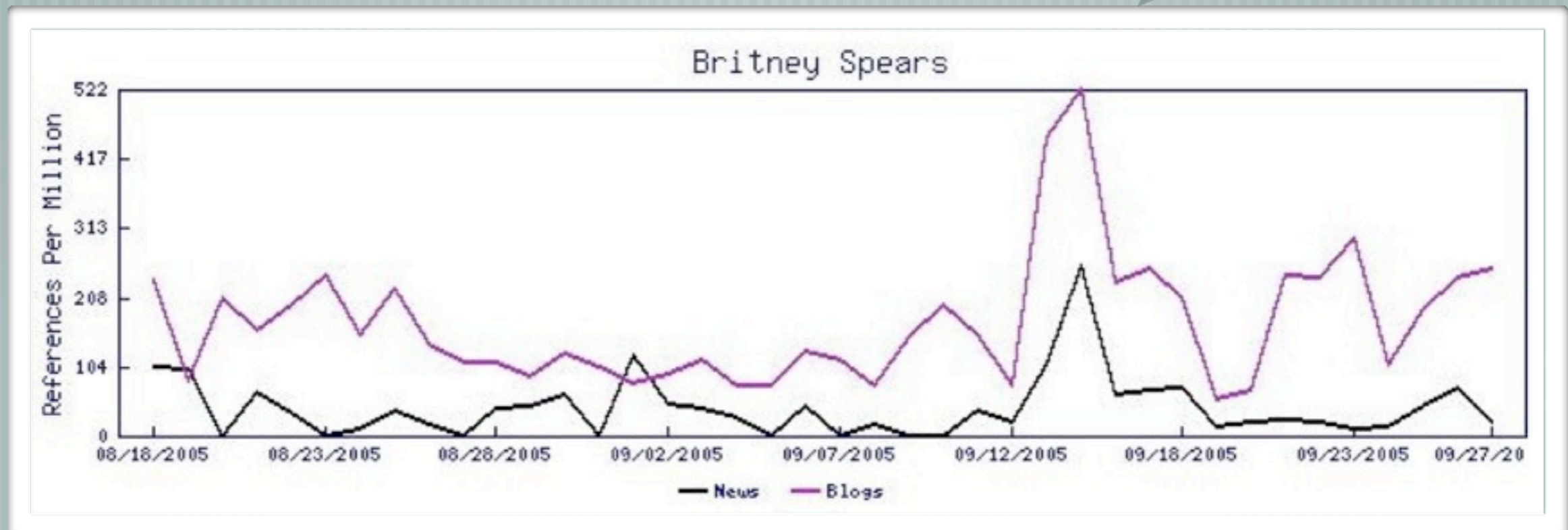
Dynamics of term usage

vs. source type

vs. location

predictive

(Lloyd, Kaulgud, Skiena, 2005)



Dynamics of conversations

Dynamics of term usage

vs. source type

vs. location

predictive



(Lloyd, Kaulgud, Skiena, 2005)

Dynamics of conversations

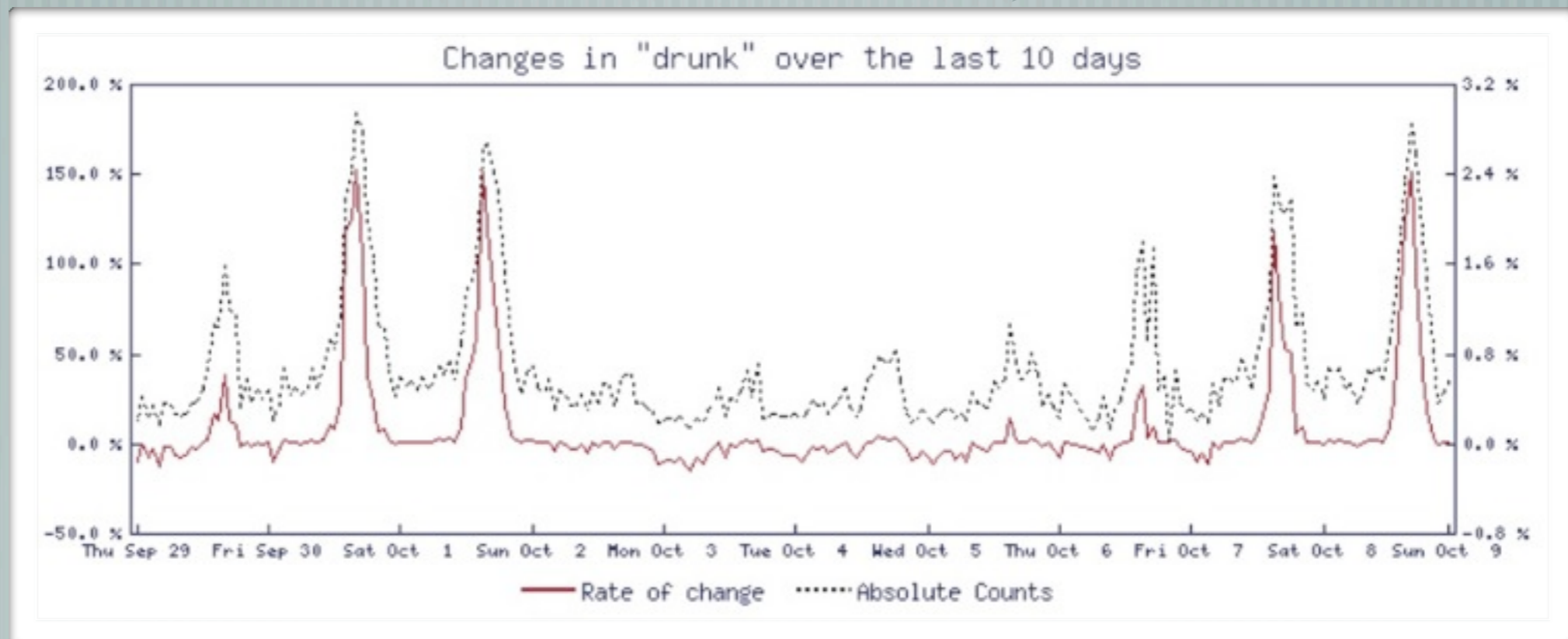
Dynamics of term usage

vs. source type

vs. location

predictive

(Balog et al., 2004; Mishne et al., 2006)



Dynamics of conversations

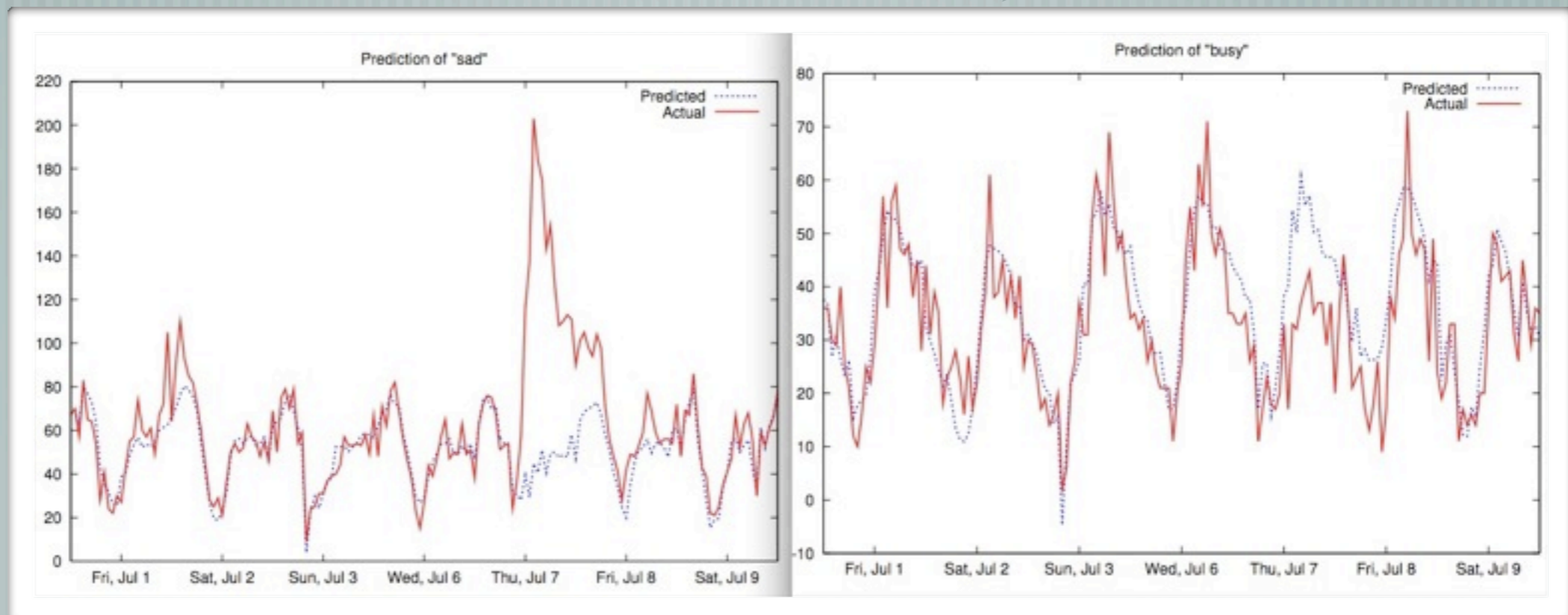
Dynamics of term usage

vs. source type

vs. location

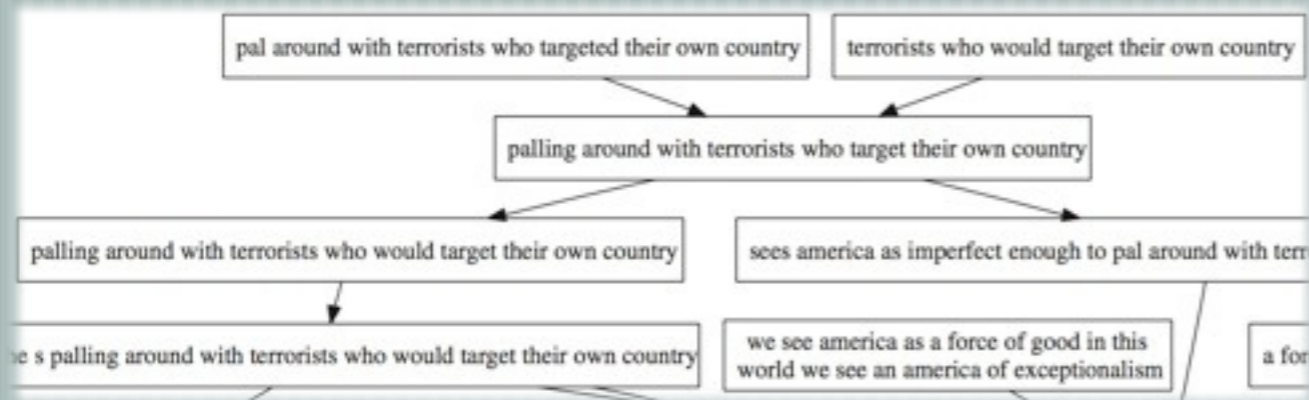
predictive

(Balog et al., 2004; Mishne et al., 2006)

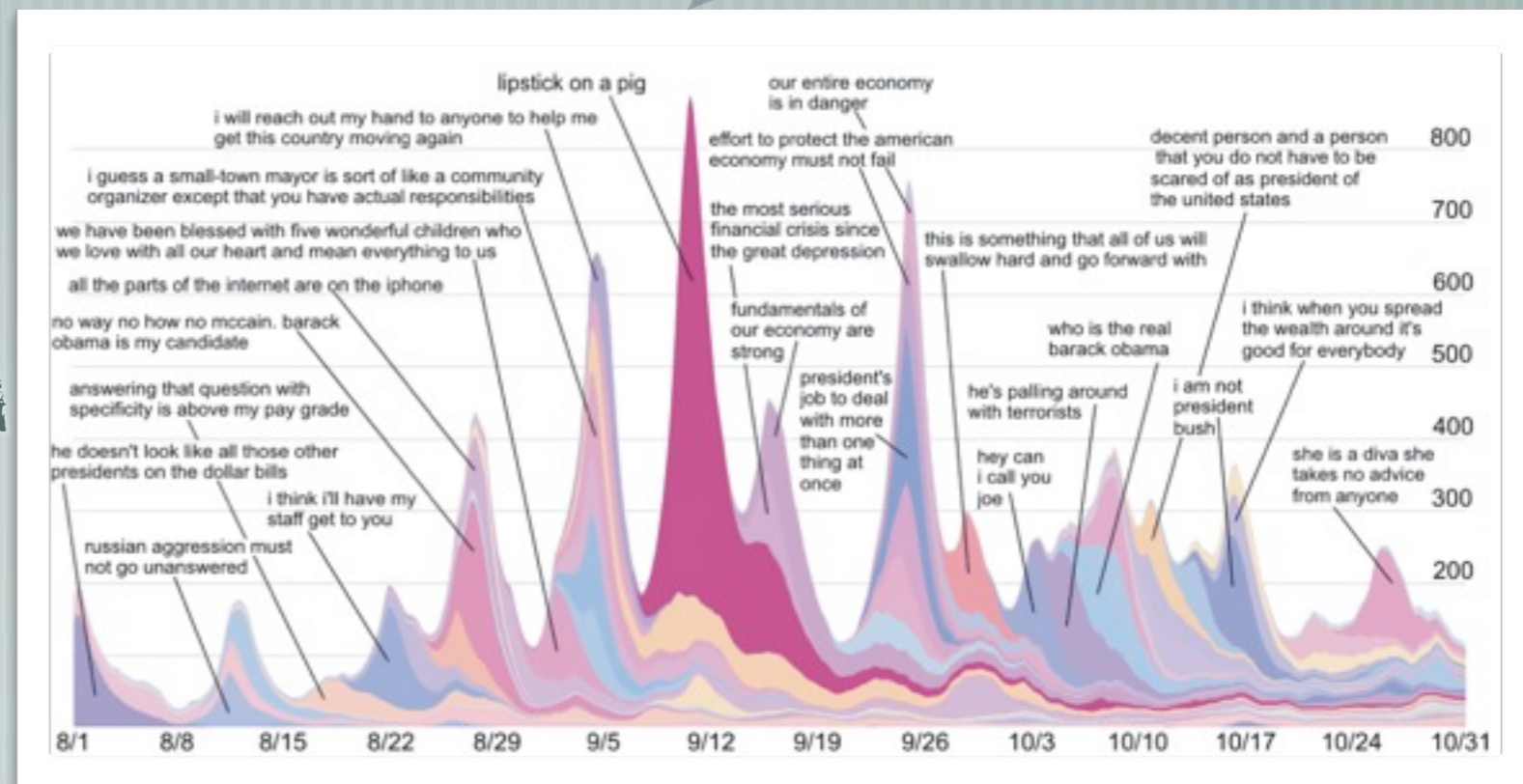


Dynamics of concepts

Dynamics of sentences, called "memes"

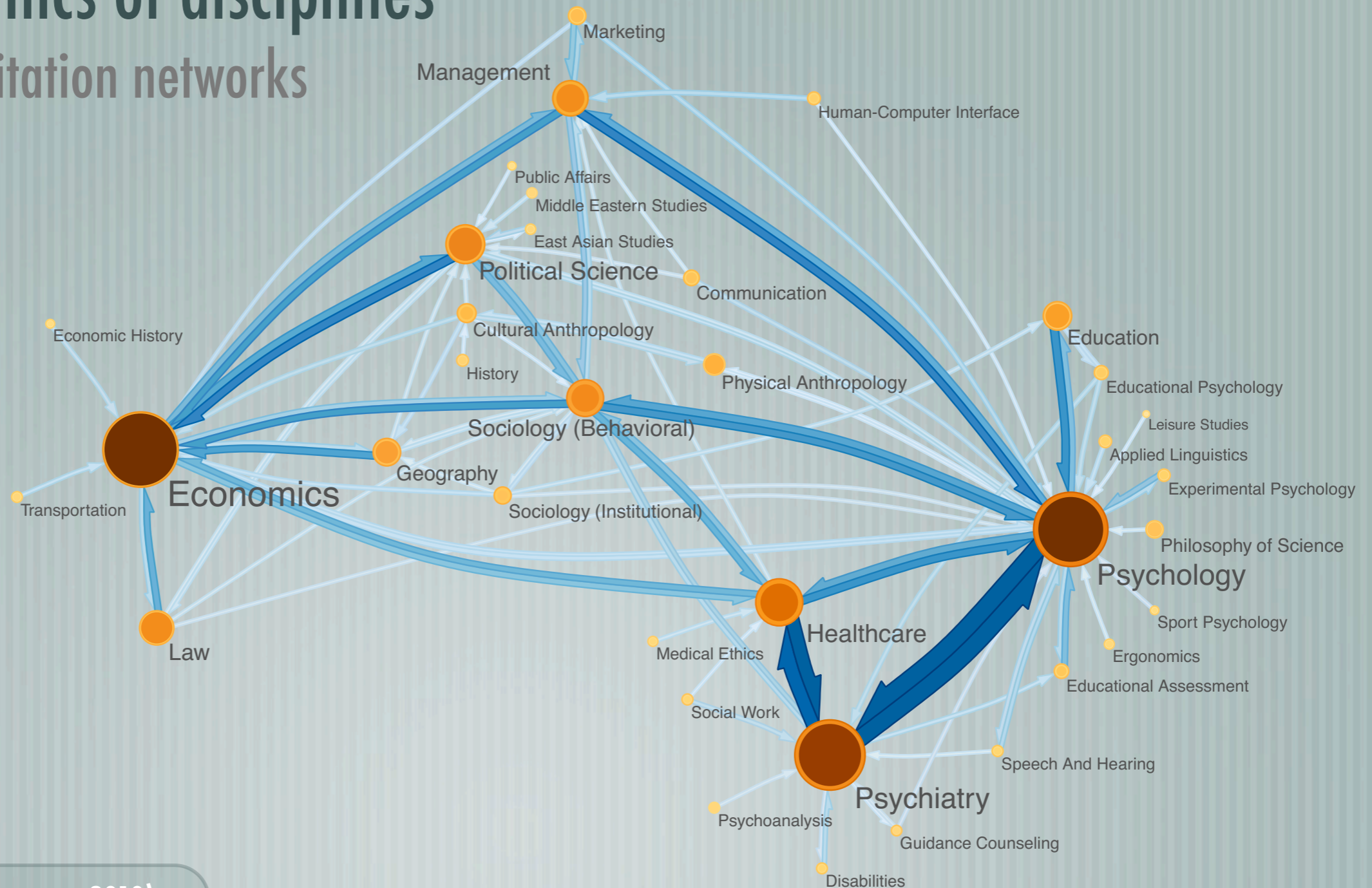


(Leskovec, Backstrom, Kleinberg, 2009)



Dynamics of concepts

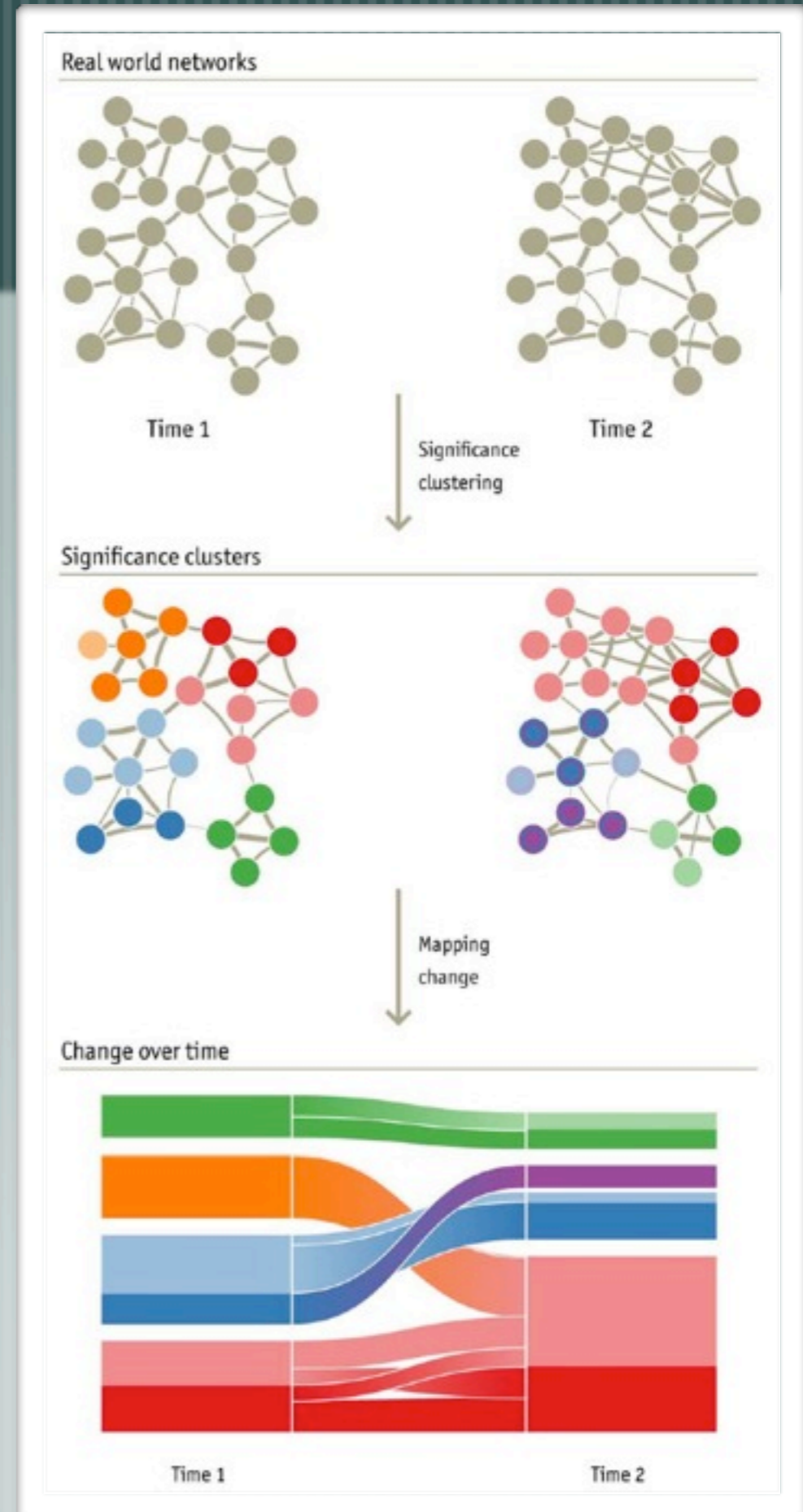
Dynamics of disciplines
using citation networks



(Rosvall & Bergstrom, 2010)

Dynamics of concepts

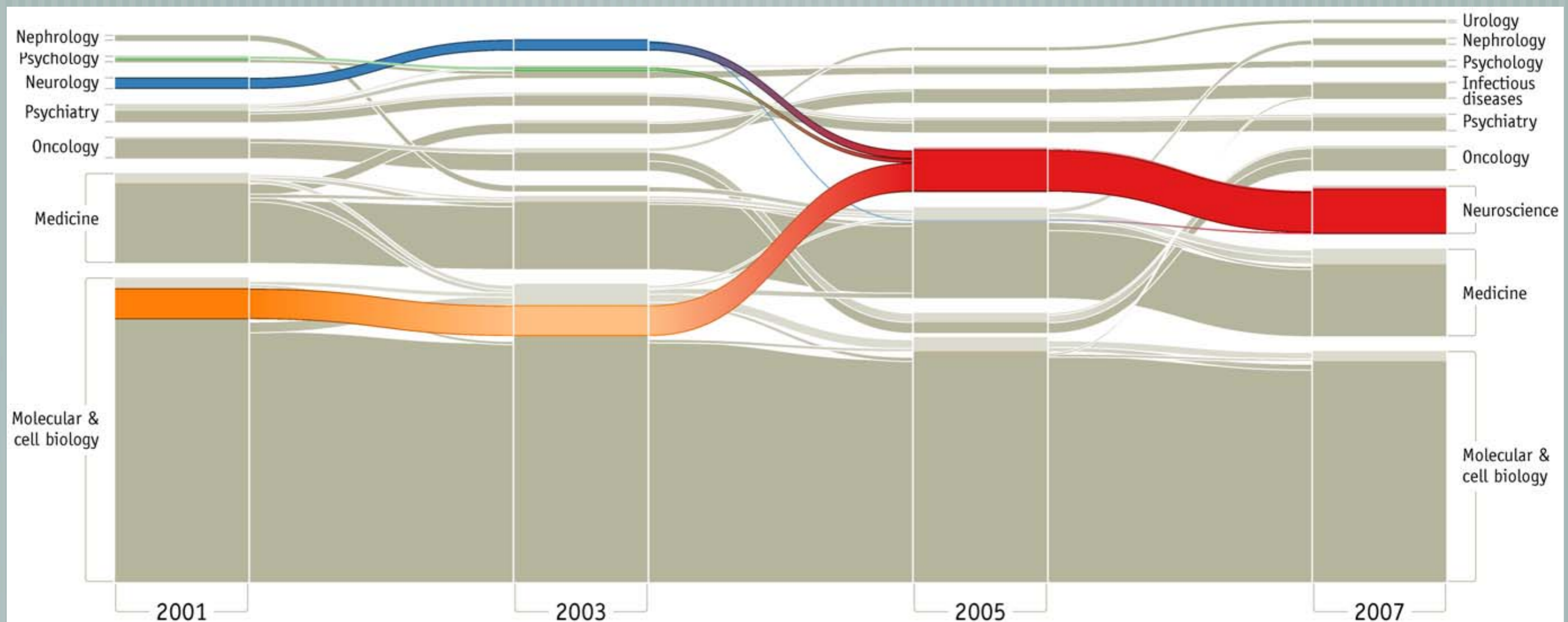
Dynamics of disciplines
using citation networks



(Rosvall & Bergstrom, 2010)

Dynamics of concepts

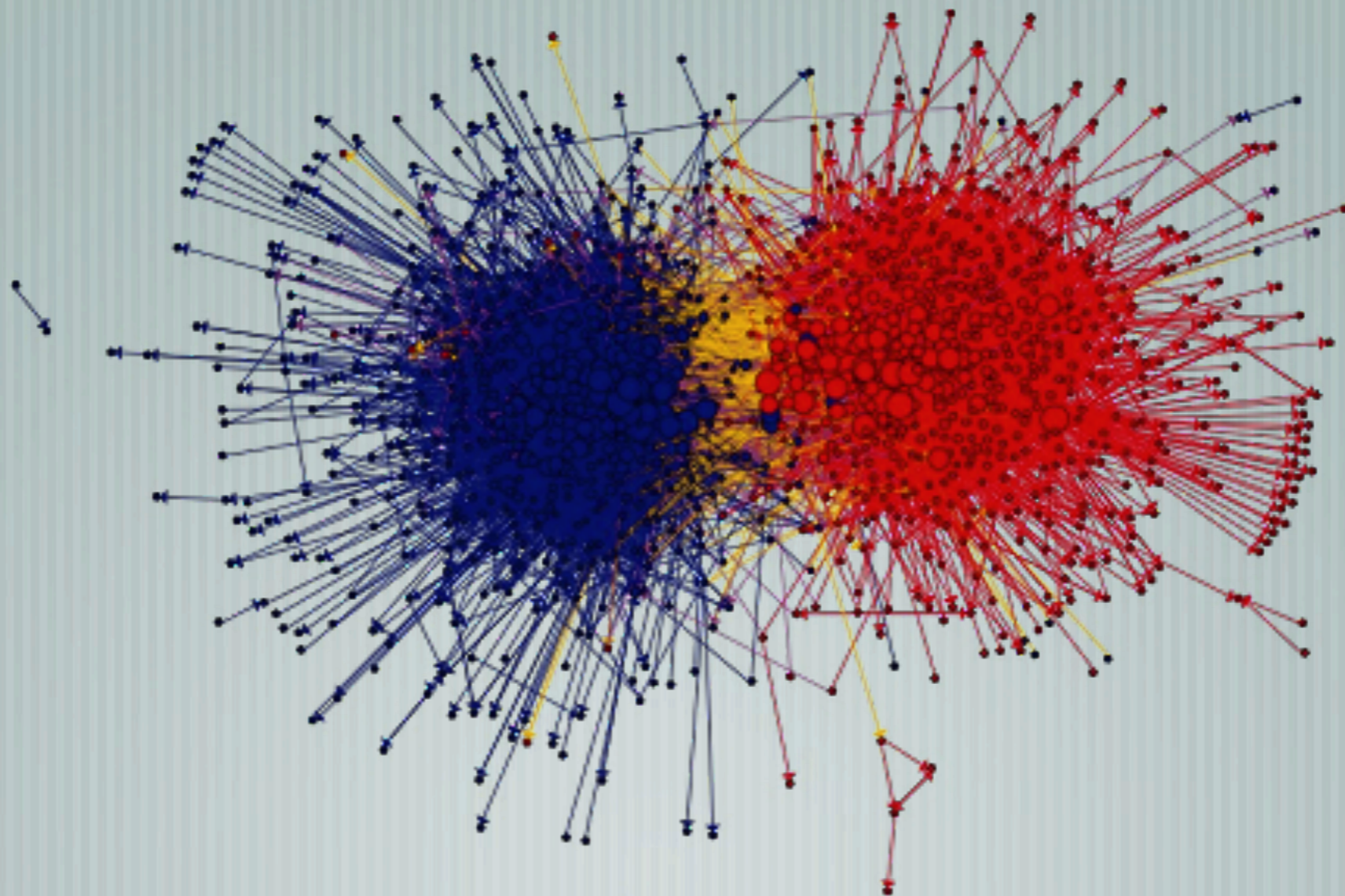
Dynamics of disciplines
using citation networks



(Rosvall & Bergstrom, 2010)

Connecting structure & content

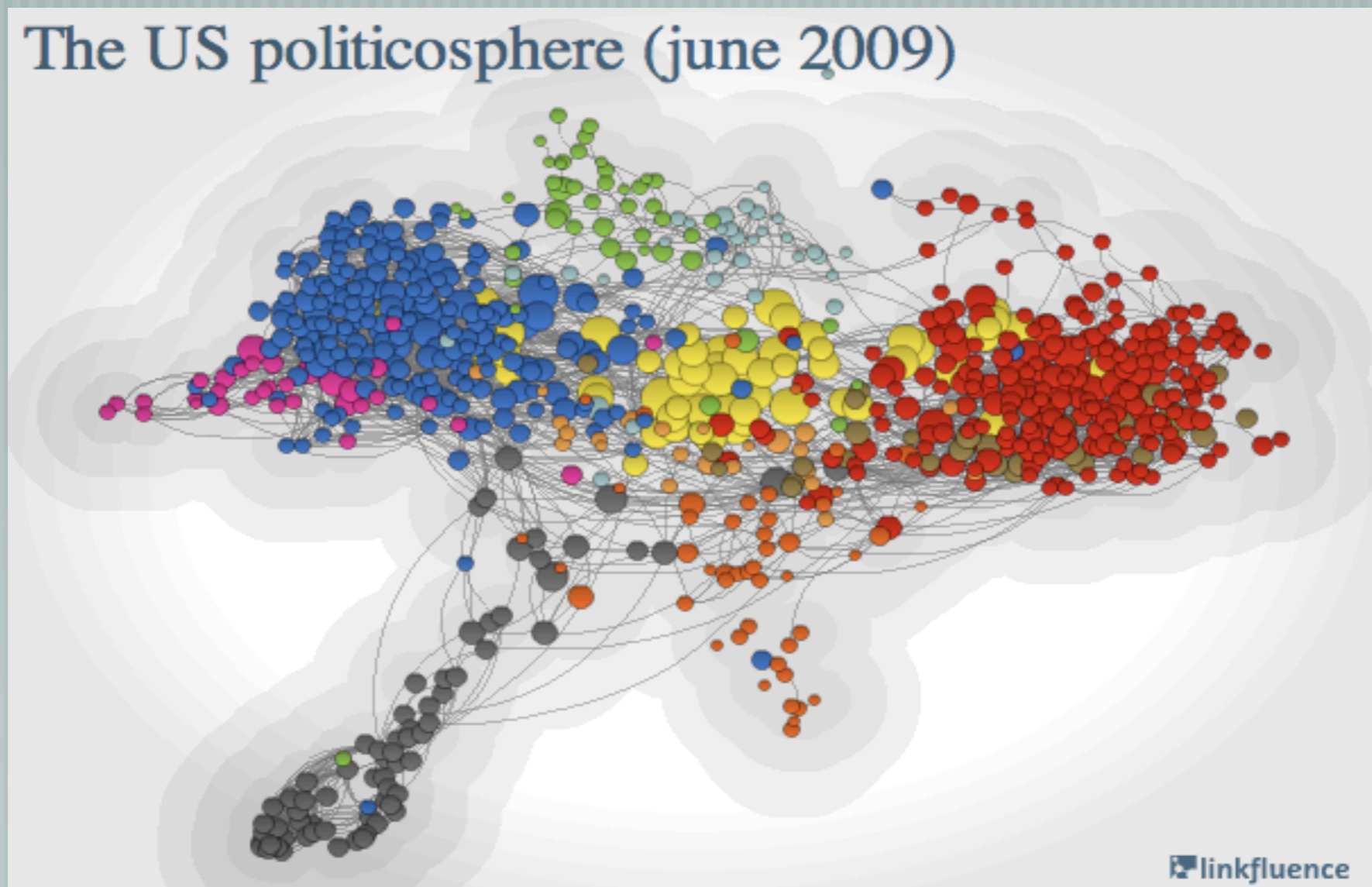
— [Describing the topological structure of opinions



(Adamic & Glance, 2004; Linkfluence, 2009)

Connecting structure & content

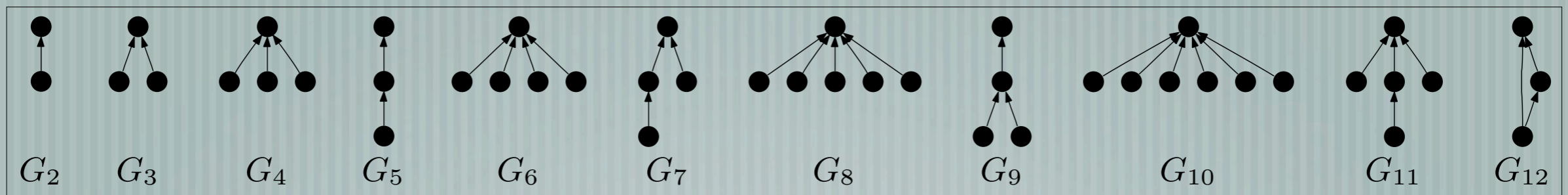
— [Describing the topological structure of opinions



(Adamic & Glance, 2004; Linkfluence, 2009)

Socio-semantic morphology

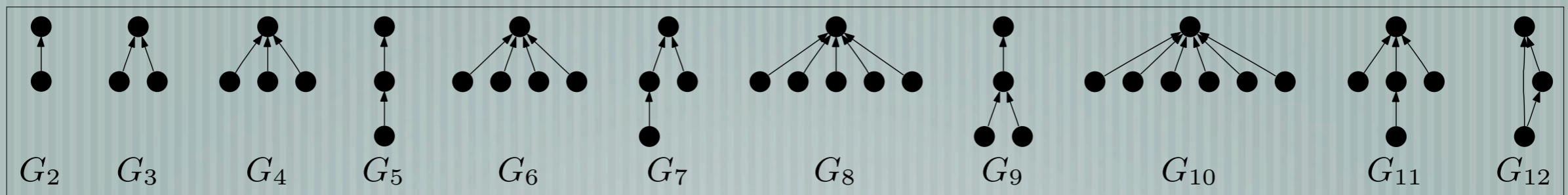
Diffusion cascade shapes in a blog network



(McGlohon,
Leskovec, Faloutsos,
Hurst, Glance, 2007)

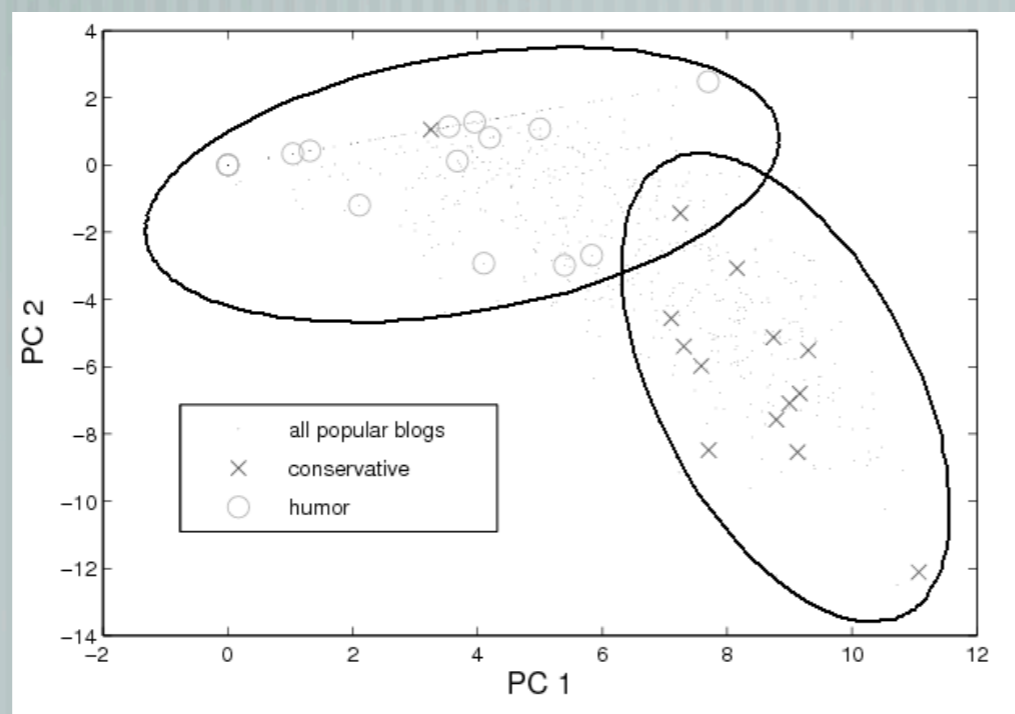
Socio-semantic morphology

Diffusion cascade shapes in a blog network

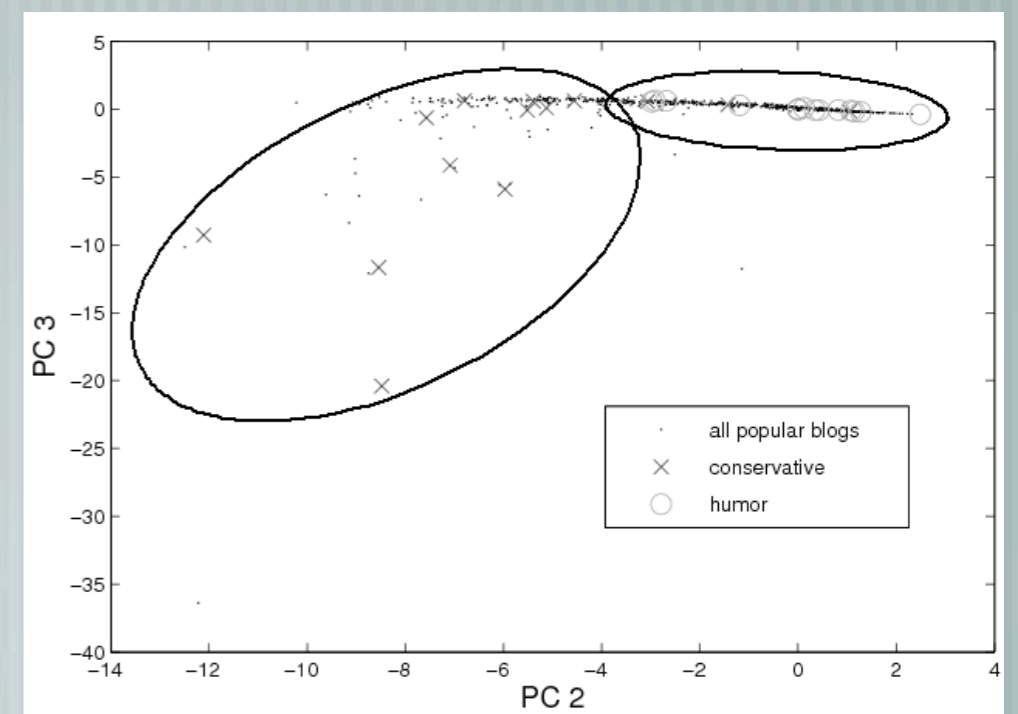


Unsupervised categorization

(McGlohon,
Leskovec, Faloutsos,
Hurst, Glance, 2007)



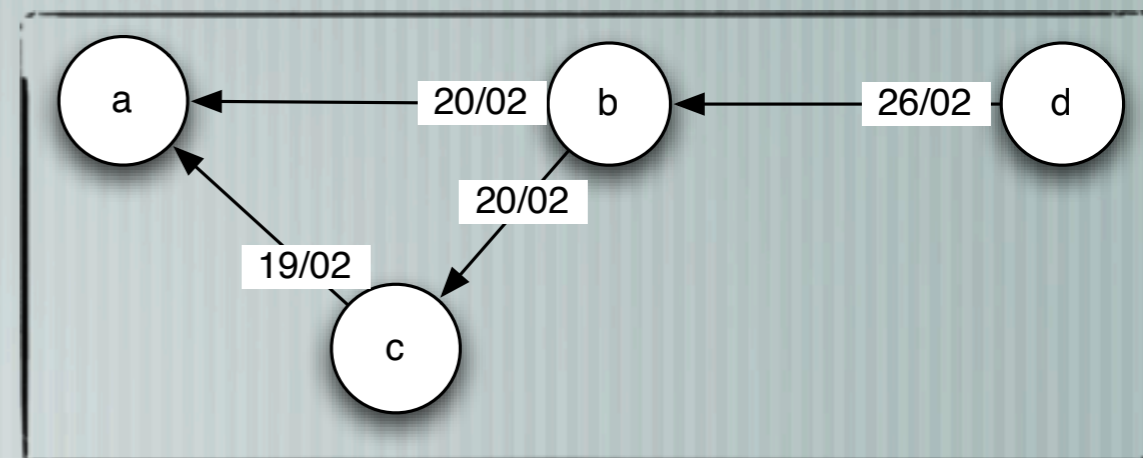
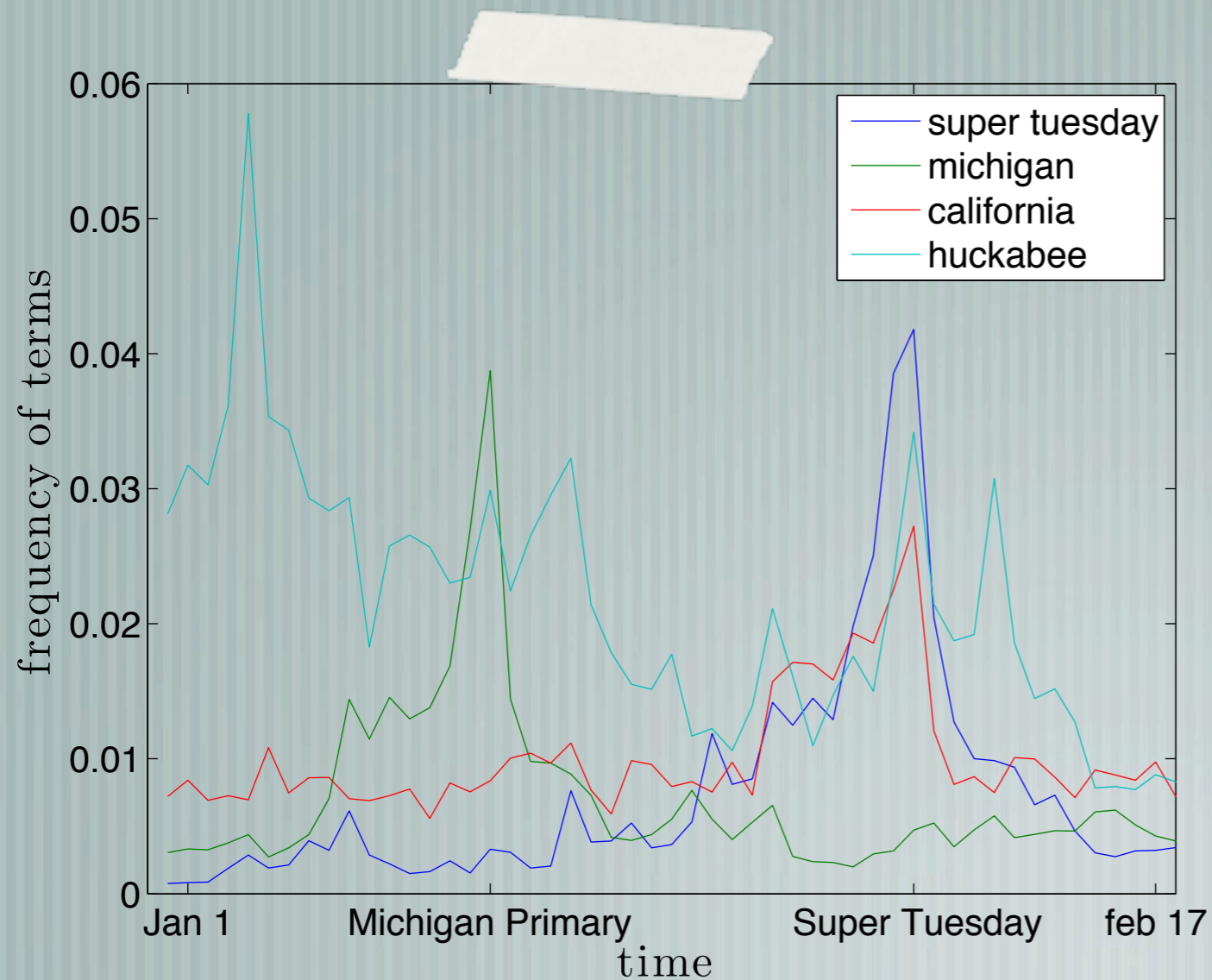
(a) First vs. second PC



(b) Second vs. third PC

Co-evolution in blogspace

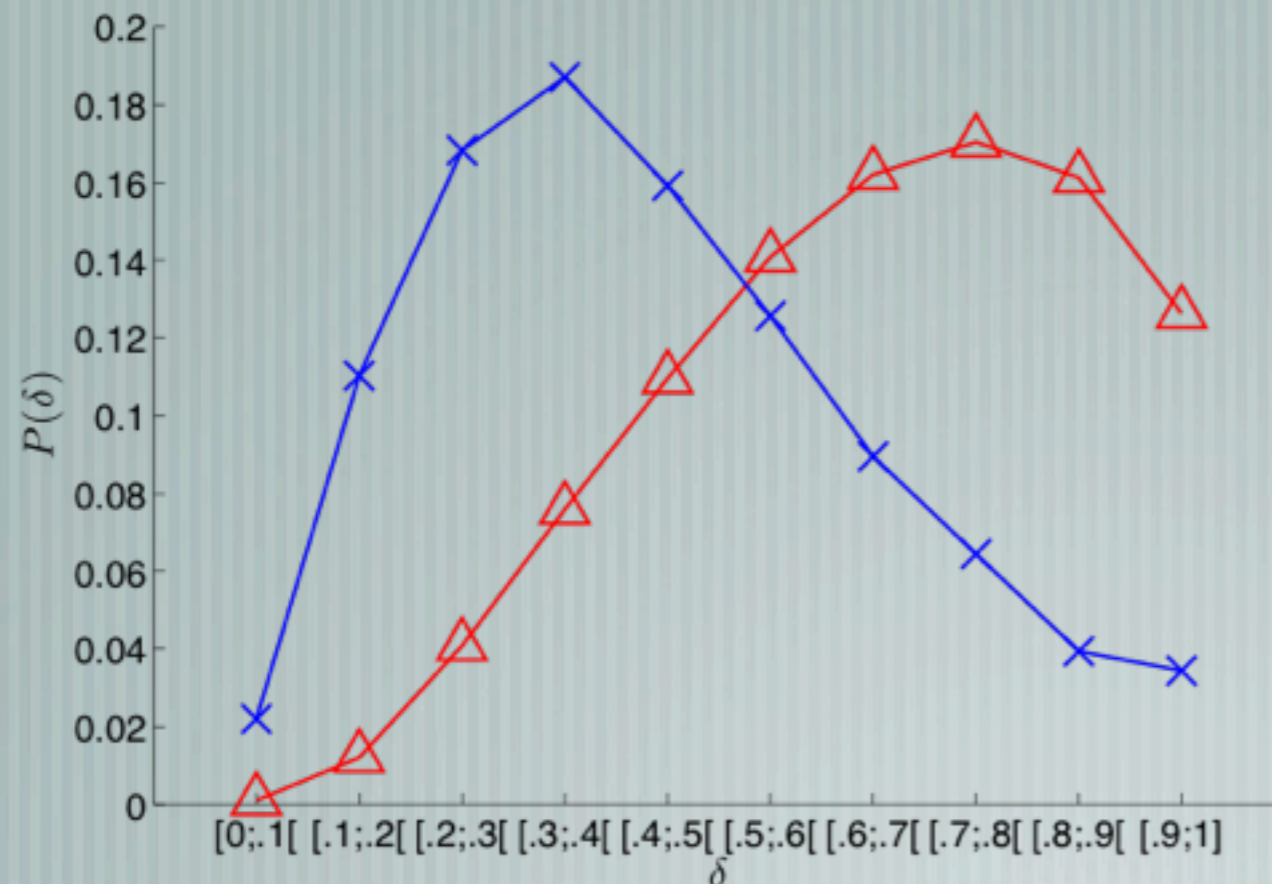
US blogosphere during the presidential primaries in 2008



(Cointet, Roth, 2009; Roth, Cointet, 2010)

Co-evolution in blogspace

US blogosphere during the presidential primaries in 2008



semantic dissimilarity

between blogs i and j :

$$\delta(i, j) = 1 - \frac{\hat{\mathbf{W}}_i \cdot \hat{\mathbf{W}}_j}{\|\hat{\mathbf{W}}_i\| \|\hat{\mathbf{W}}_j\|}$$

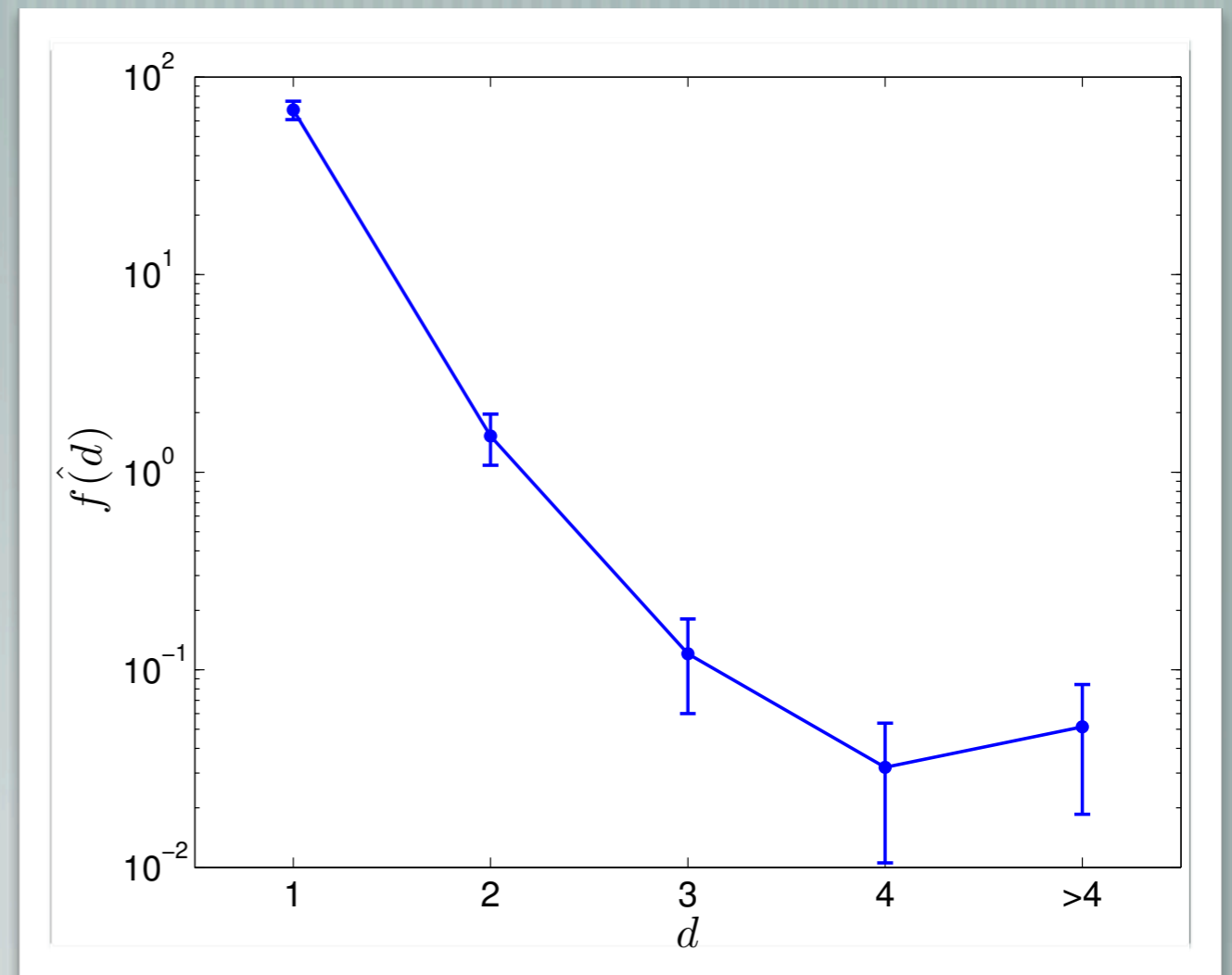
strong homophily

distributions of semantic dissimilarities *triangles*: over all possible blog pairs
crosses: over linked blogs

(Cointet, Roth, 2009; Roth, Cointet, 2010)

Social network vs. content

Interactions are mostly repeated



Social network vs. content

Interactions are mostly repeated

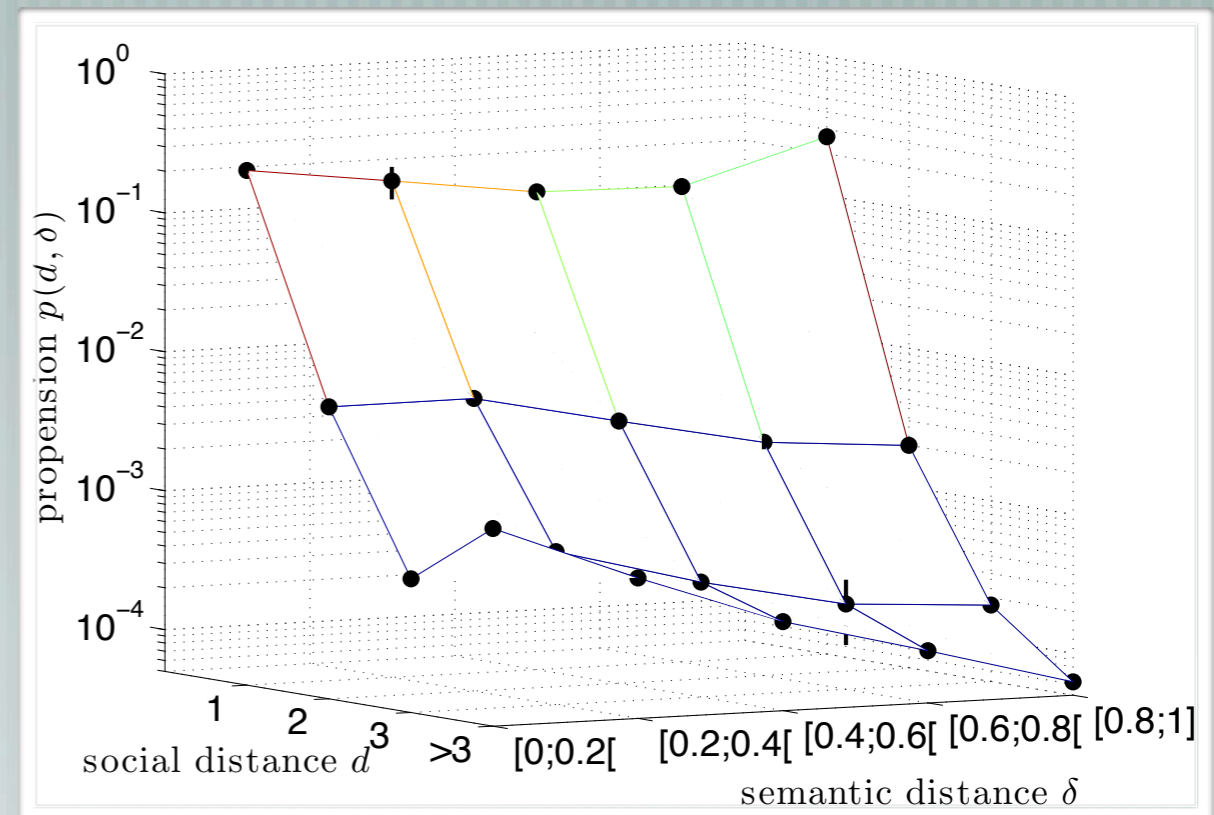
New interactions more likely:

with higher in-degree nodes

with lower topological distance

with higher semantic similarity

“Narrow world”



Social network vs. content

Interactions are mostly repeated

New interactions more likely:

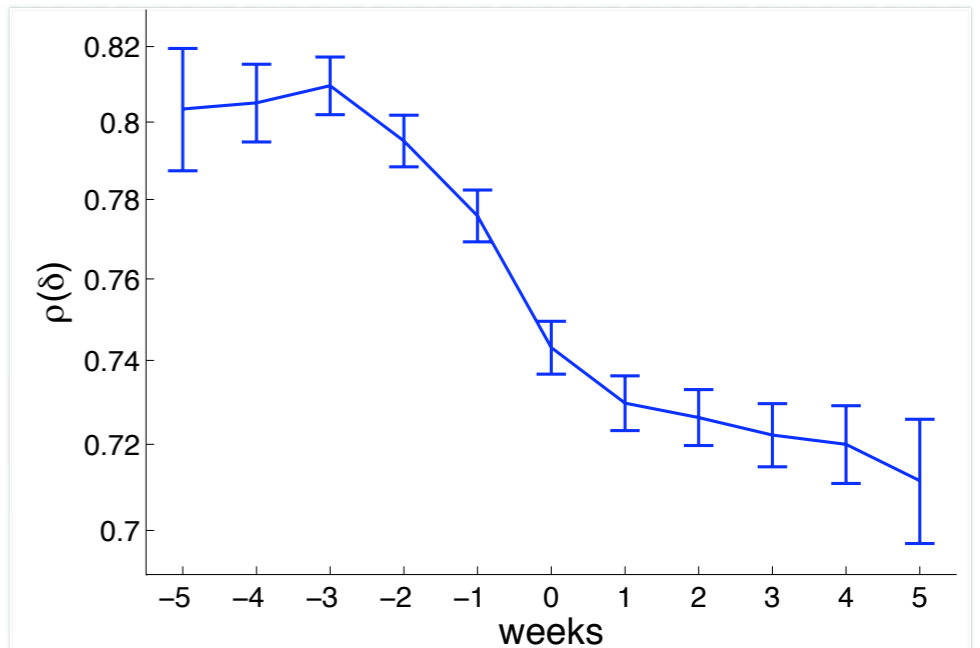
with higher in-degree nodes

with lower topological distance

with higher semantic similarity

“Narrow world”

Contraction



Social network vs. content

Interactions are mostly repeated

New interactions more likely:

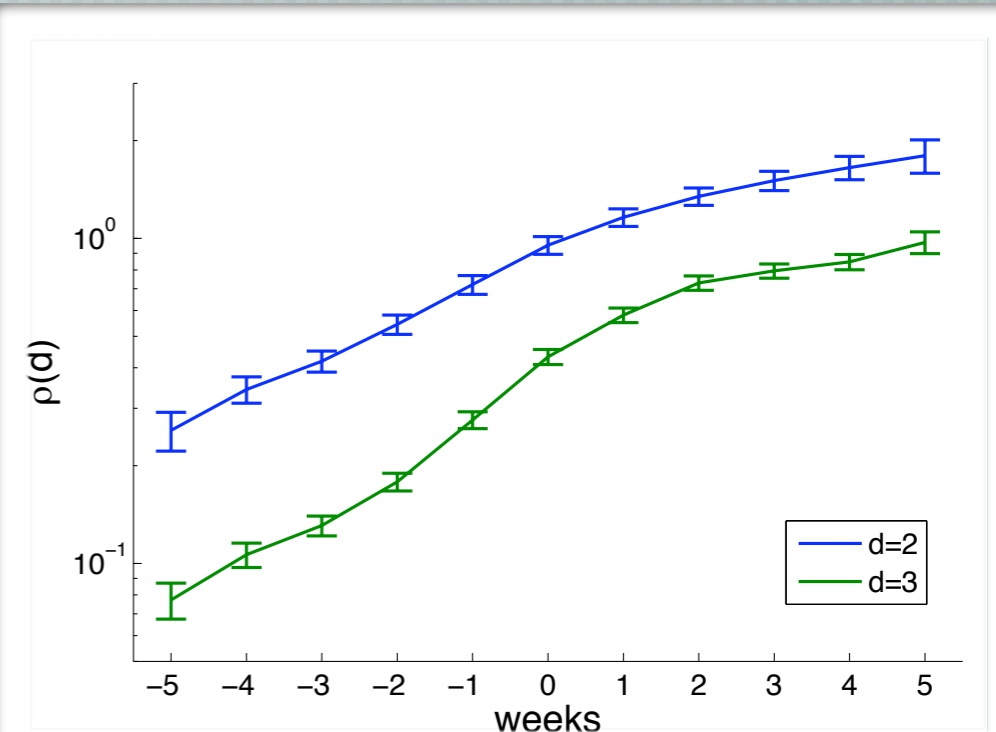
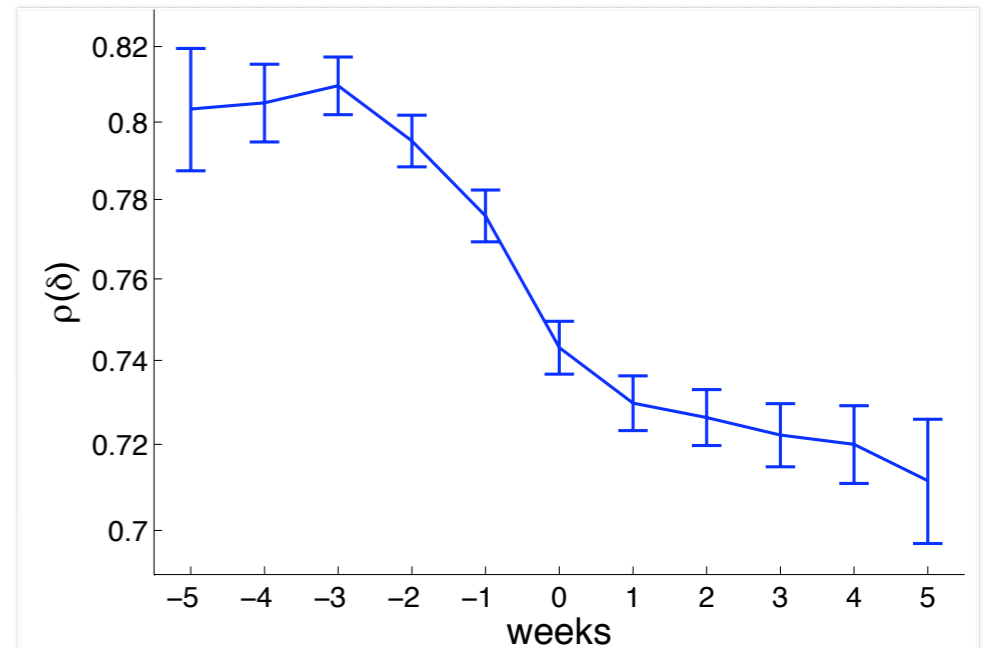
with higher in-degree nodes

with lower topological distance

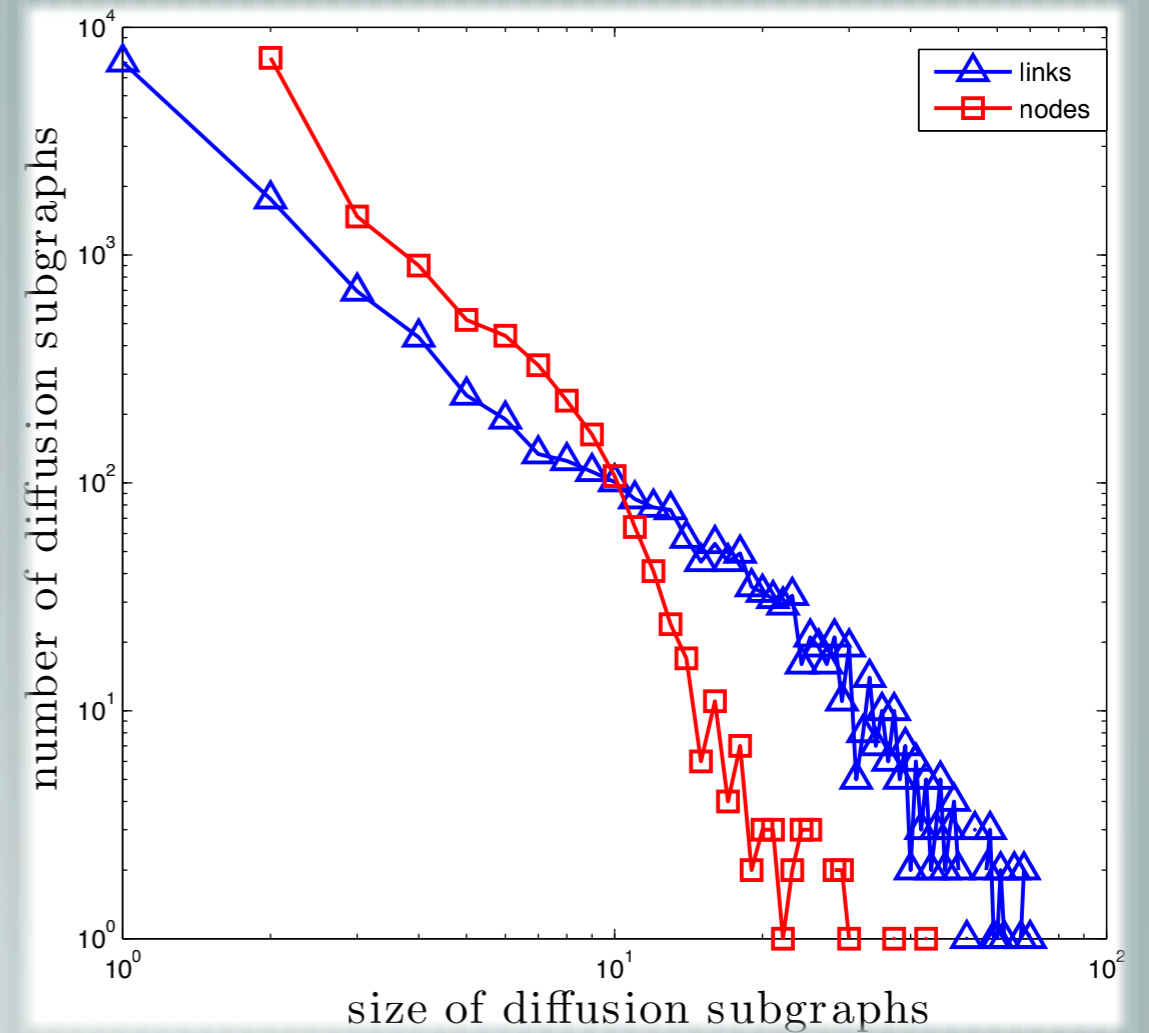
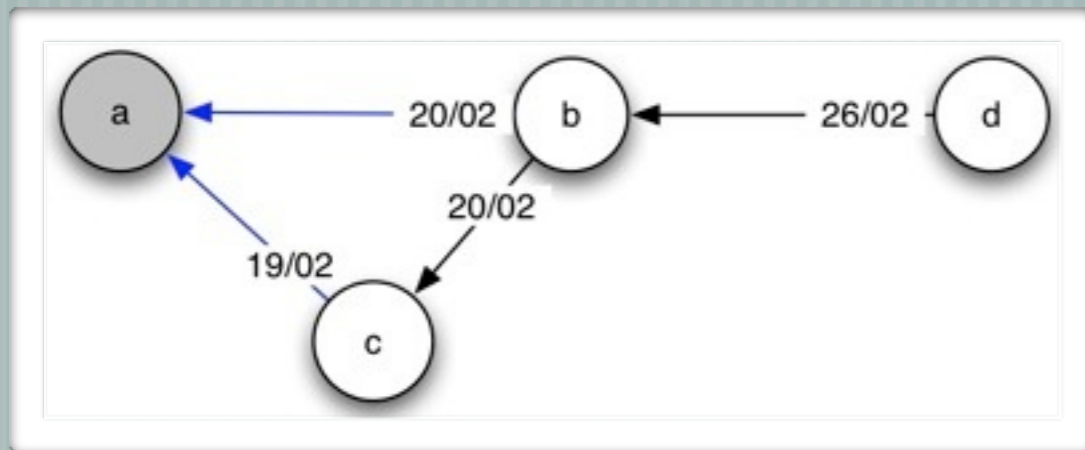
with higher semantic similarity

“Narrow world”

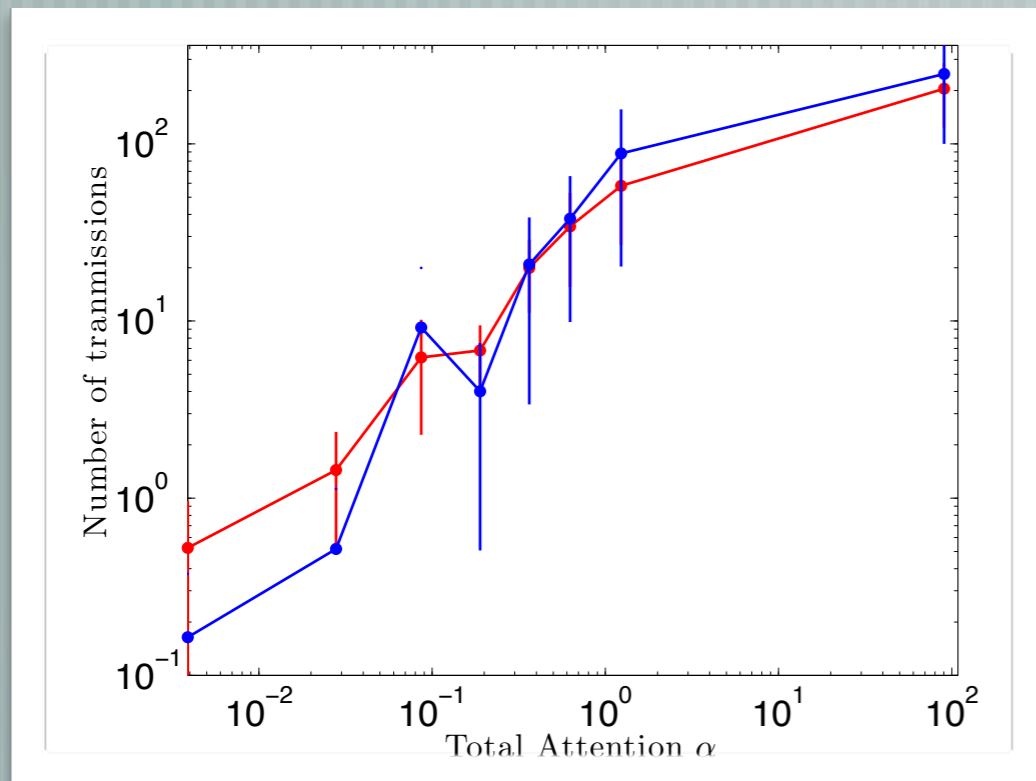
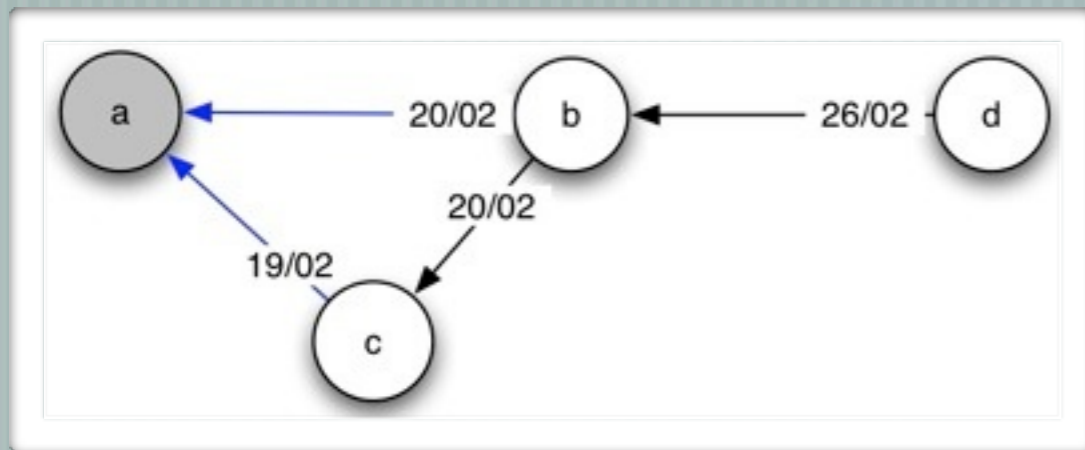
Contraction



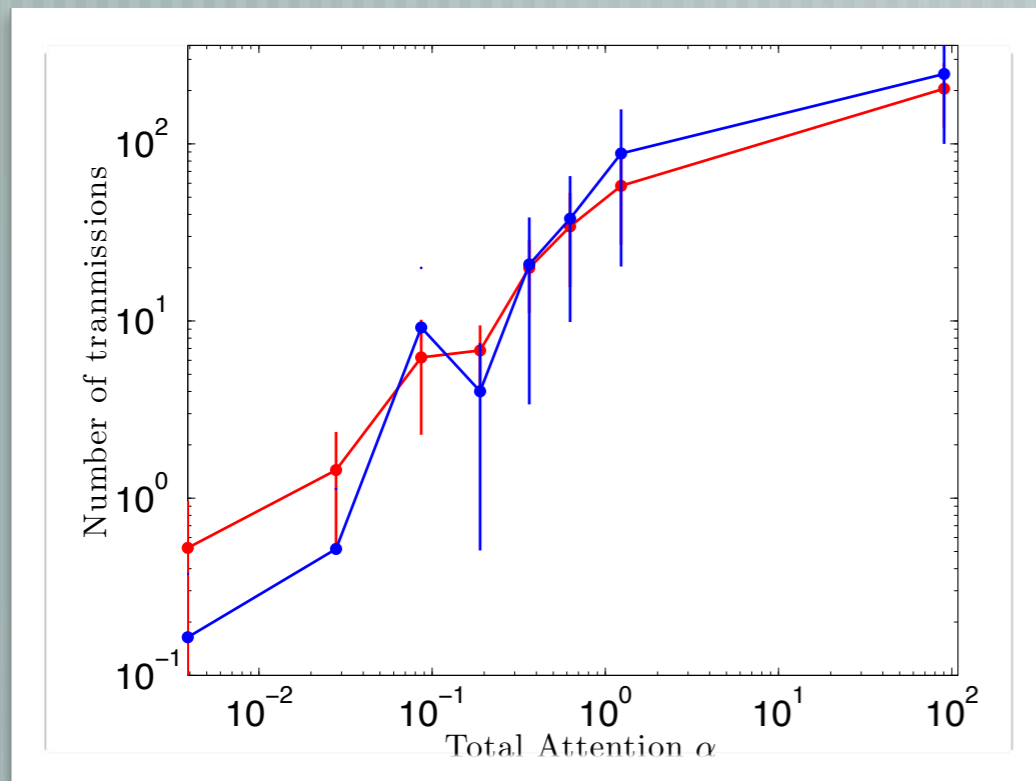
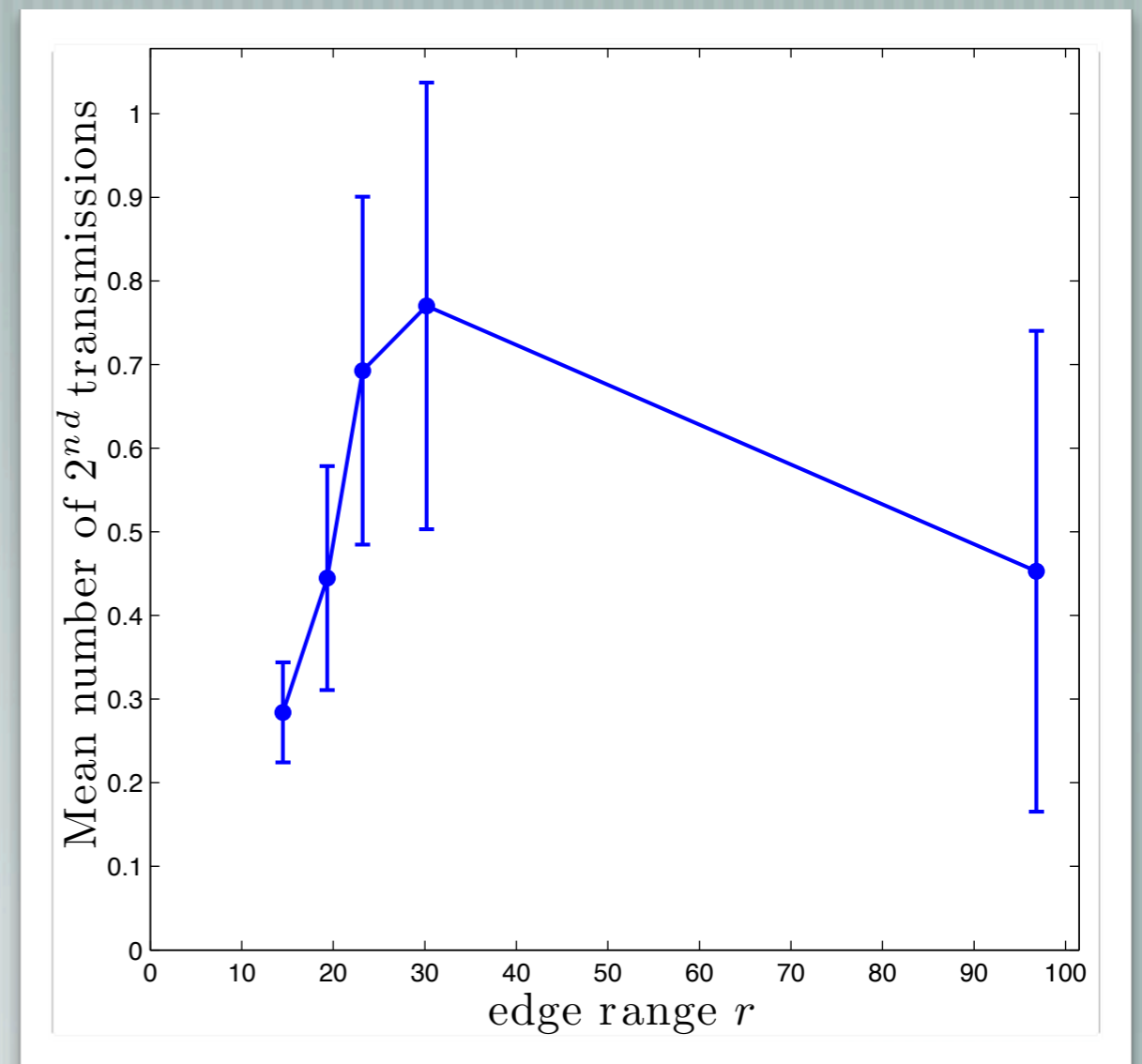
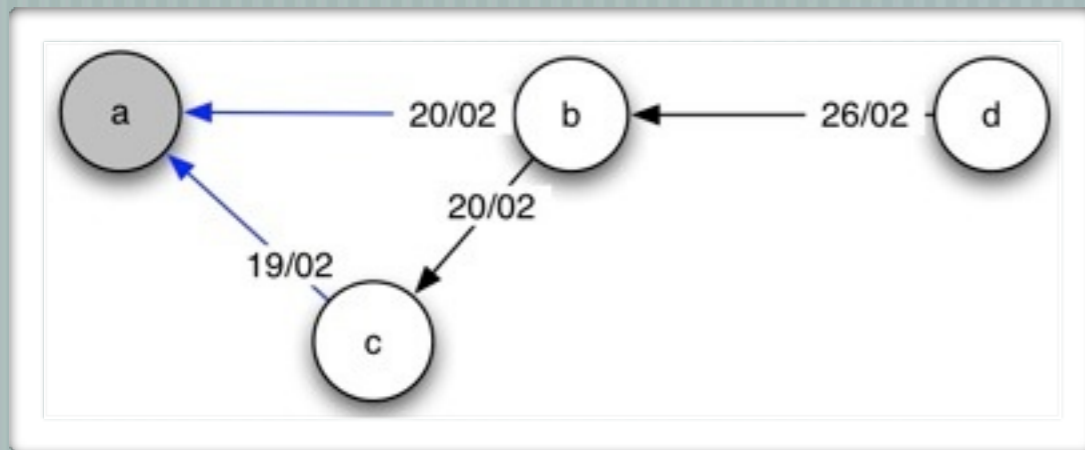
Content vs. social network



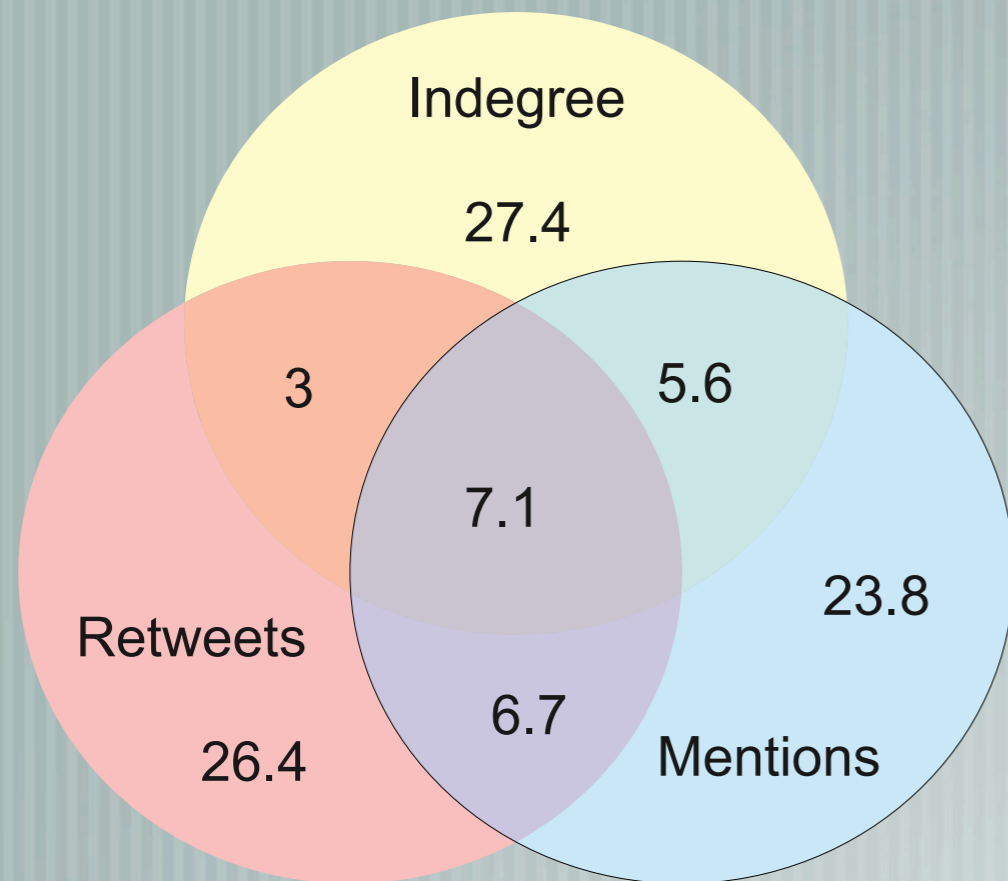
Content vs. social network



Content vs. social network



Socio-semantic influence effects



Venn diagram of the top-100 influentials

(Cha, Haddadi, Benevenuto, Gummadi, 2010)

— [high indegree not correlated with many retweets

— [influence is boosted by focusing on a given topic,

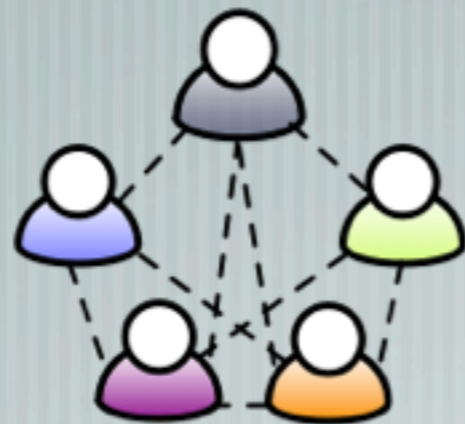
— ...even if most influential users can remain influent on a variety of topics

Socio-semantic hypergraphs

Limits when focusing on the level of individual

influence of characteristics expressable at the mesolevel of the team only,
team formation processes \neq sum of individual rationalities

(Taramasco, Cointet,
Roth, 2010)



VS.

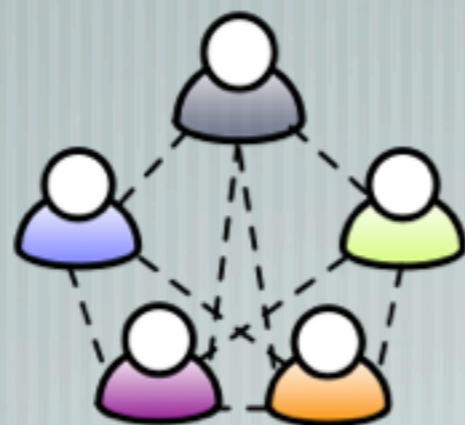


Socio-semantic hypergraphs

Limits when focusing on the level of individual

influence of characteristics expressable at the mesolevel of the team only,
team formation processes \neq sum of individual rationalities

(Taramasco, Cointet,
Roth, 2010)



VS.

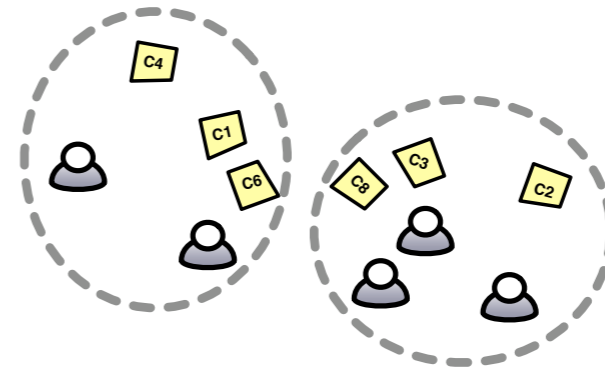


Collaboration also depends on cognitive properties

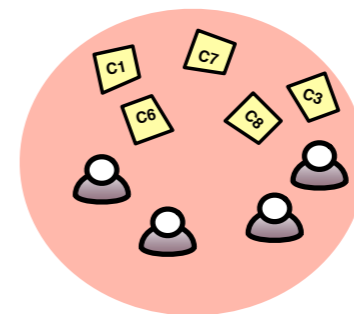
how **teams** are formed, given both social and semantic features?

Socio-semantic hypergraphs

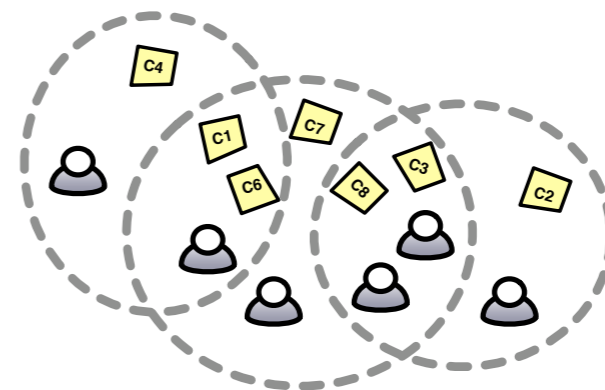
Dynamic hypergraph on socio-semantic teams



\mathcal{E}_{t-1}



$\cup \Delta \mathcal{E}_t$



$= \mathcal{E}_t$

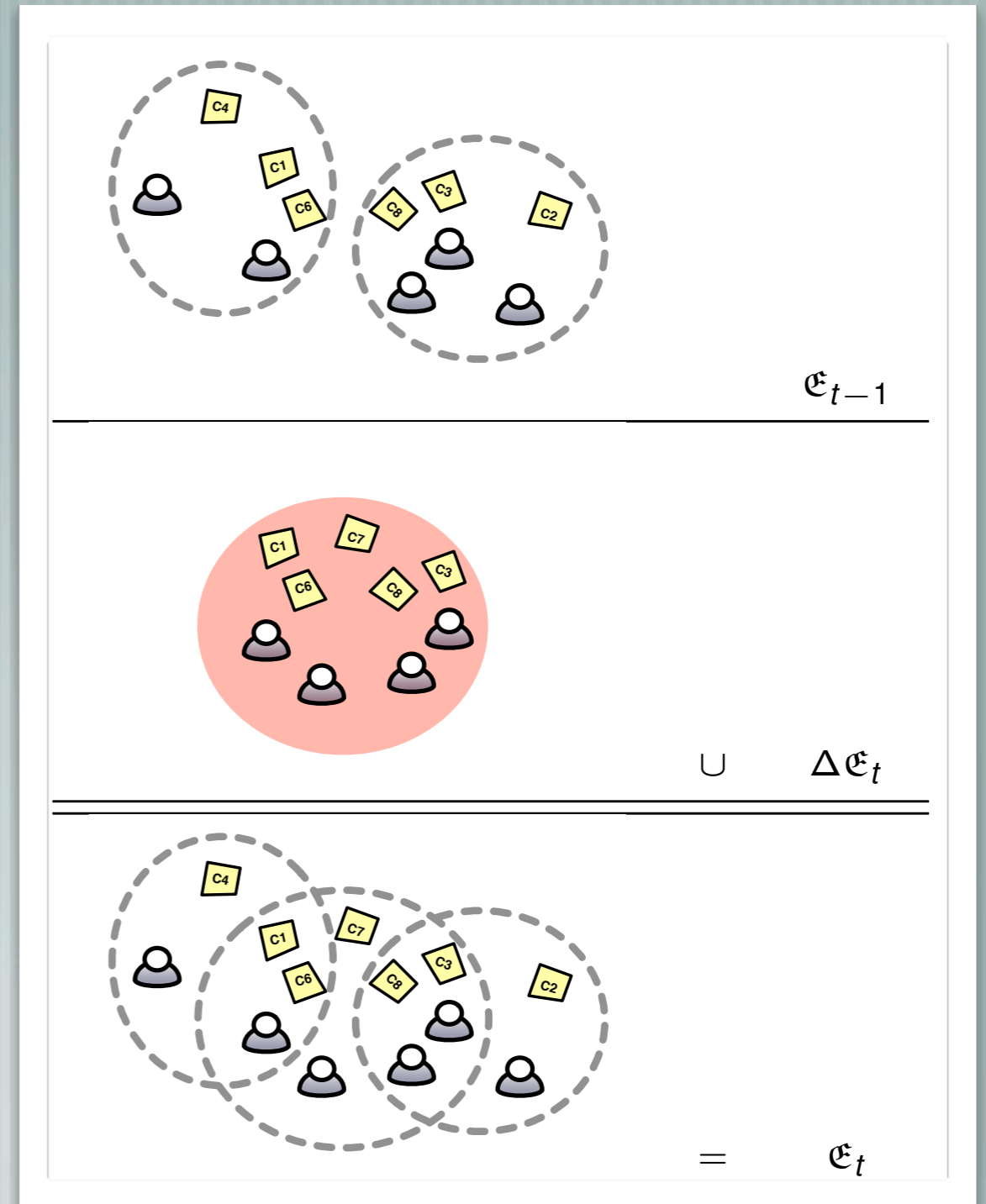
Socio-semantic hypergraphs

Dynamic hypergraph on socio-semantic teams

Defining various hypergraphic indices...

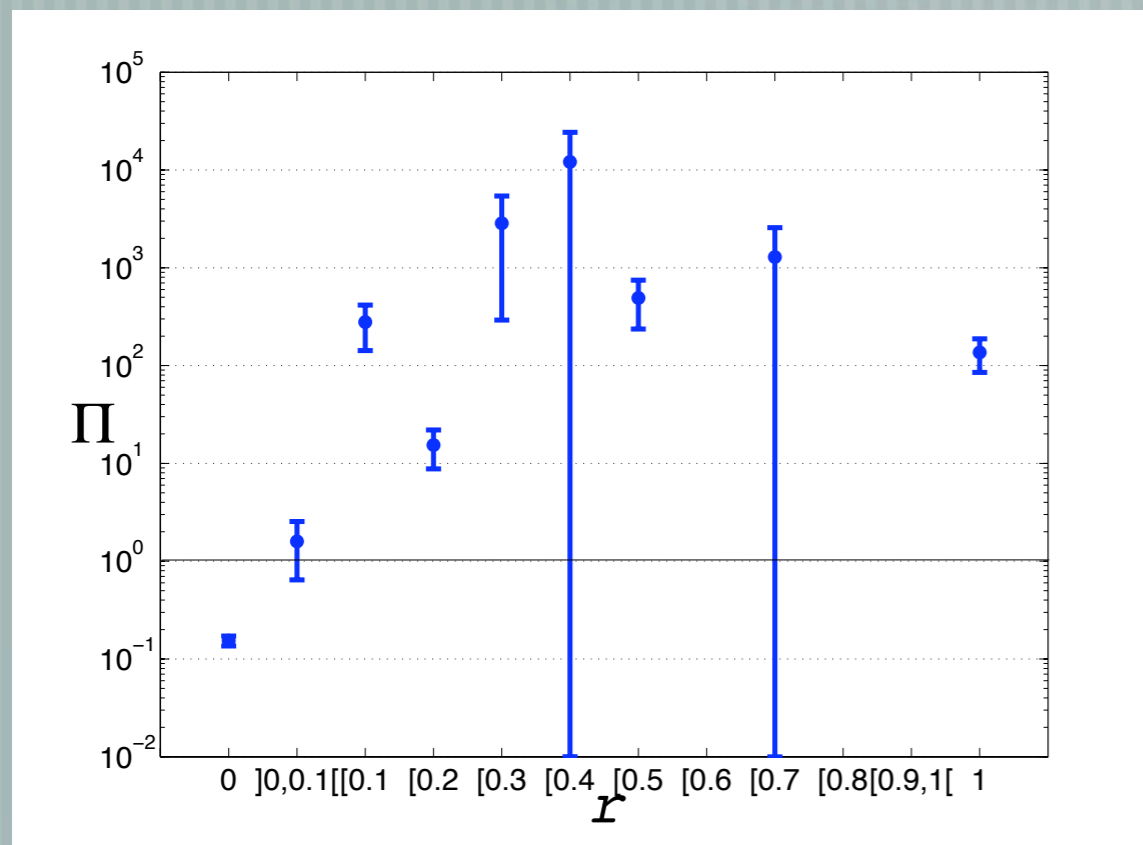
such as expertise ratio, hypergraphic repetition ratio, etc.

...and comparing with an random artificial baseline



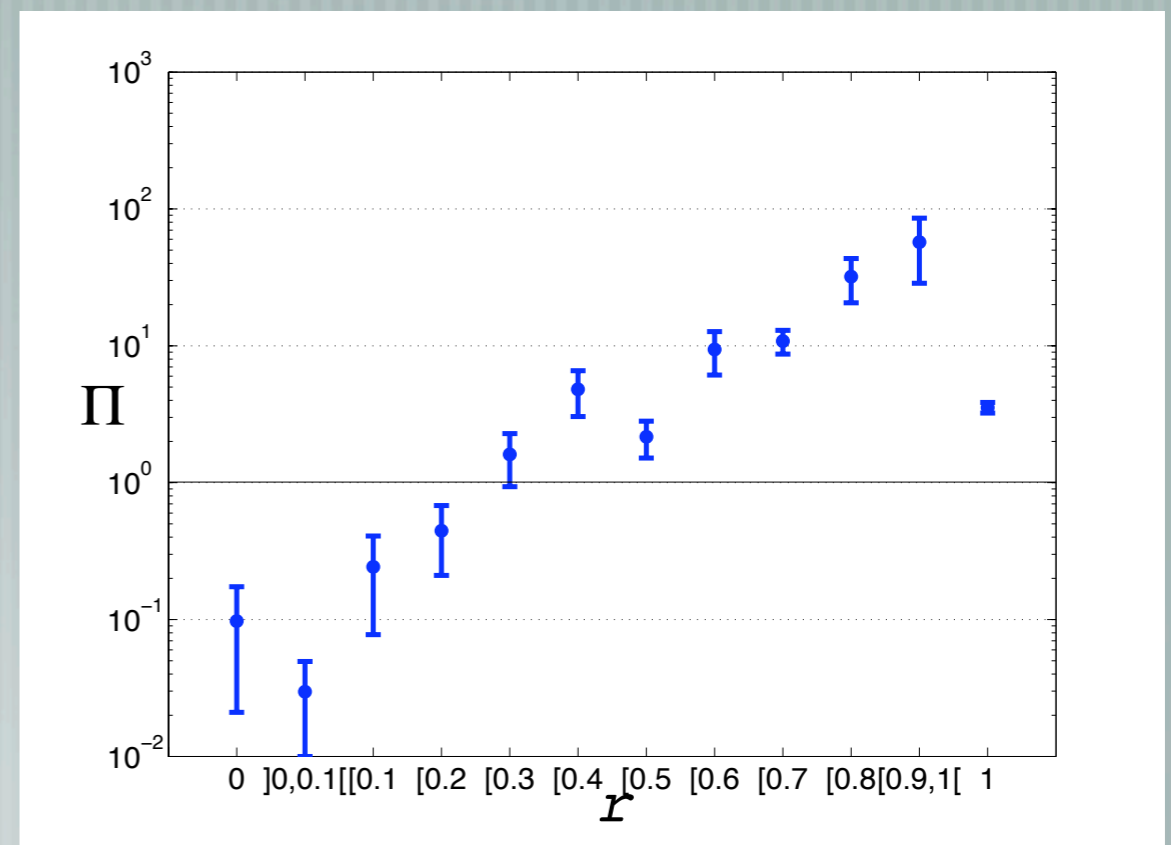
Socio-semantic hypergraphs

Computing the propensity of team formation



social

high proportion of interaction repetitions

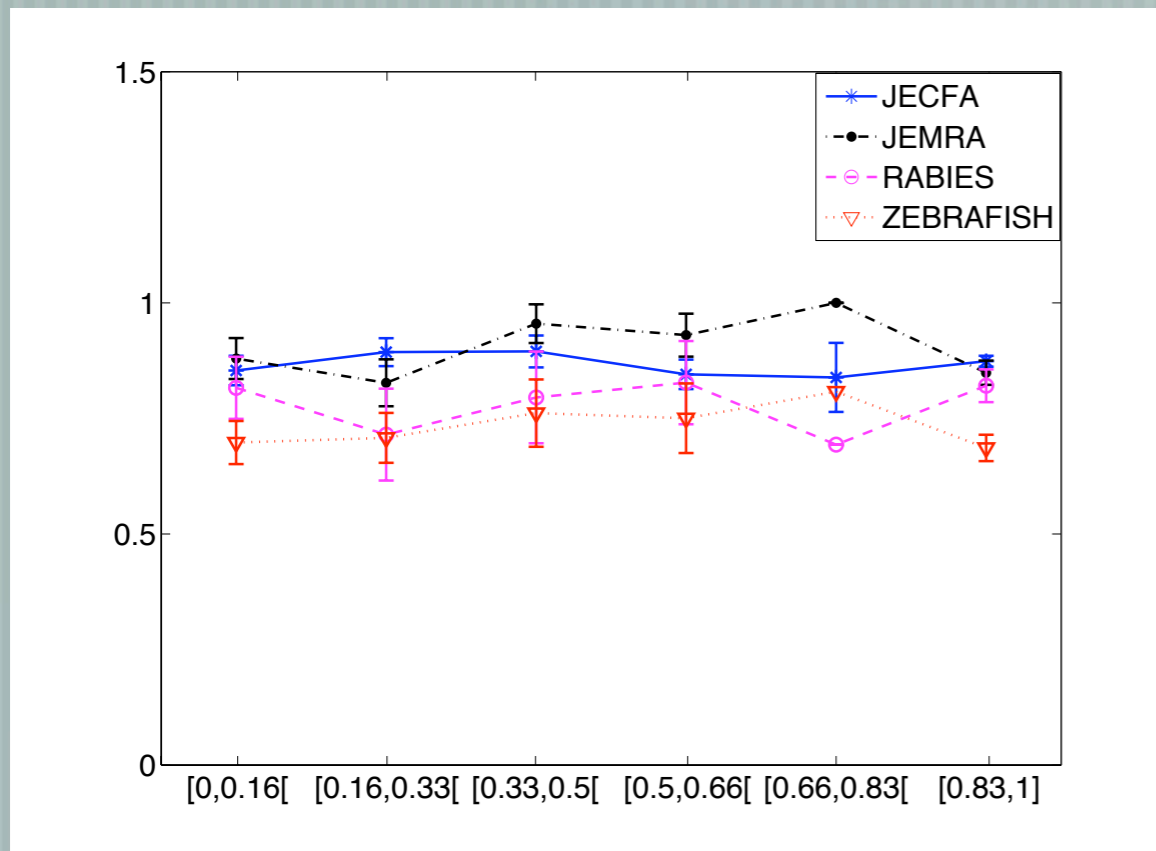


semantic

bias towards gathering groups of concepts which were previously associated

Socio-semantic hypergraphs

Computing the socio-semantic correlation of teams



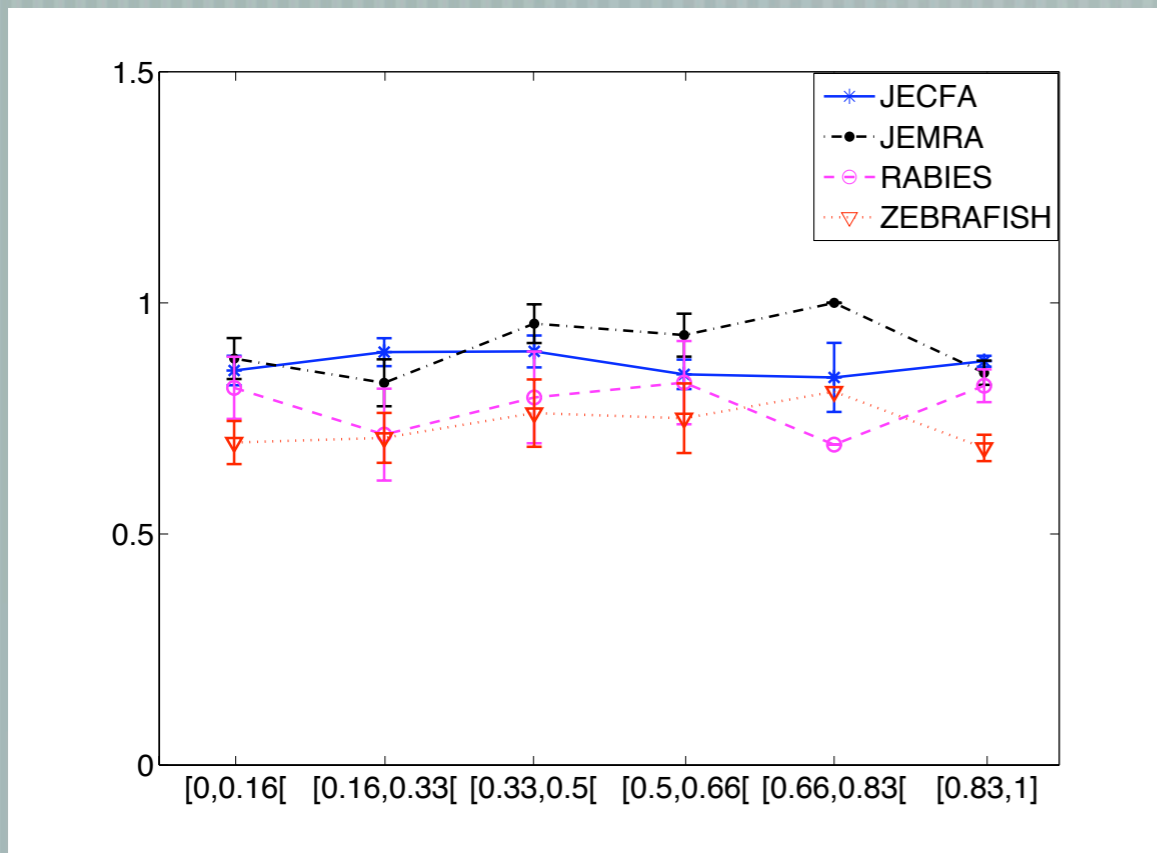
socio-semantic

We observe no correlation

contrarily to intuition, new semantic associations **do not** stem more from original teams than from repeated teams

Socio-semantic hypergraphs

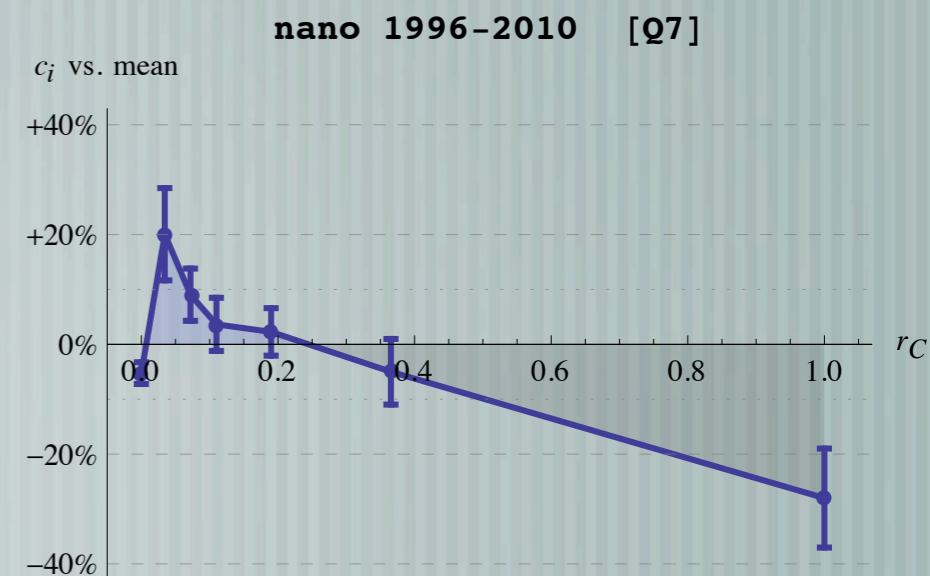
Computing the socio-semantic correlation of teams



socio-semantic

We observe no correlation
contrarily to intuition, new semantic associations **do not** stem more from original teams than from repeated teams

semantic originality vs. "performance"



Towards experimental socio-semantic dynamics

roth@ehess.fr
<http://camille.roth.free.fr>

- [We now have good knowledge of social network processes.
- [We still need to develop a solid framework to describe local cognition processes.