

Assessing group cohesion in homophily networks

Benjamin Renoust, benjamin.renoust@labri.fr

Guy Melançon, guy.melancon@labri.fr

Marie-Luce Viaud, mlviaud@ina.fr

CNRS UMR 5800 LaBRI & INRIA Bordeaux Sud-Ouest

Université de Bordeaux I, France

Institut National de l'Audiovisuel, France

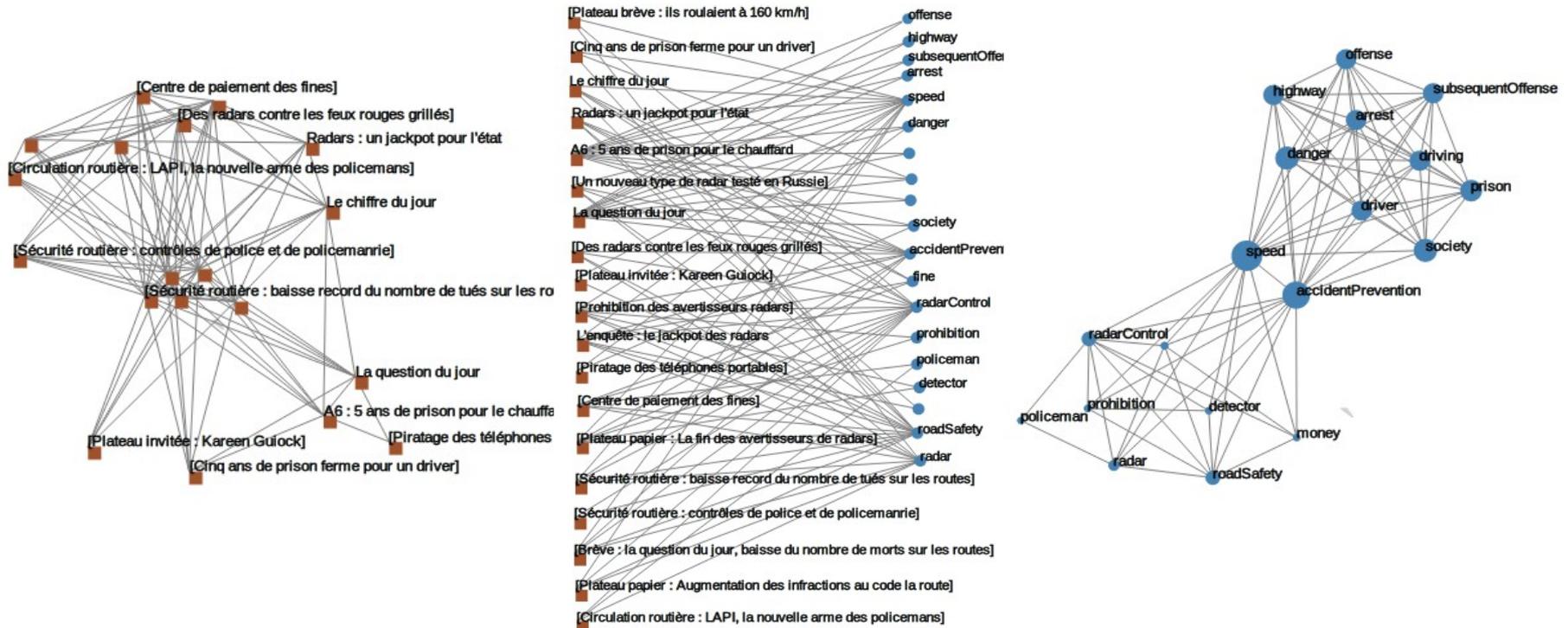
Complex Networks

Documents + Terms = **Networks**

People + Groups = **Networks**

Publications + Authors = **Networks**

... *homophily networks*



How do attributes interact ?

What is a *cohesive* group?

Radars : un jackpot pour l'état

[Des radars contre les feux rouges grillés]

L'enquête : le jackpot des radars

[Centre de paiement des fines]

[Plateau papier : il roulaient à 160 km/h]

[Plateau brève : ils roulaient à 160 km/h]

[Circulation routière : LAPI, la nouvelle arme des policiers]

[Sécurité routière : contrôles de police et de policemannie]

Le chiffre du jour

[Prohibition des avertisseurs radars]

[Brève : la question du jour, baisse du nombre de morts sur les routes]

[Sécurité routière : baisse record du nombre de tués sur les routes]

[Un nouveau type de radar testé en Russie]

[Plateau papier : les avertisseurs de radars]

La question du jour

[Plateau invitée : Kareen Guiock]

[Cinq ans de prison ferme pour un driver]

A6 : 5 ans de prison pour le chauffard

[Piratage]

Radars : un jackpot pour l'état

[Des radars contre les feux rouges grillés]

L'enquête : le jackpot des radars

[Centre de paiement des fines]

[Plateau papier : ils roulaient à 160 km/h]

[Plateau brève : ils roulaient à 160 km/h]

[Circulation routière : LAPI, la nouvelle arme des policiers]

[Sécurité routière : contrôles de police et de policemannie]

Le chiffre du jour

[Prohibition des avertisseurs radars]

[Brève : la question du jour, baisse du nombre de morts sur les routes]

[Sécurité routière : baisse record du nombre de tués sur les routes]

[Un nouveau type de radar testé en Russie]

[Plateau papier : les avertisseurs de radars]

La question du jour

[Plateau invitée : Kareen Guiock]

[Cinq ans de prison ferme pour un driver]

A6 : 5 ans de prison pour le chauffard

[Piratage]

Querying/grouping/clustering of complex networks
Attributes/topic extraction/typology/ontology...

→ Unequal quality

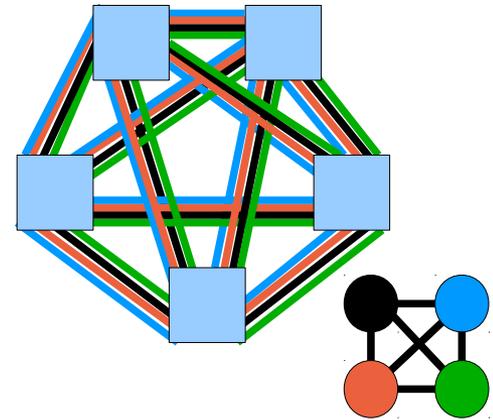
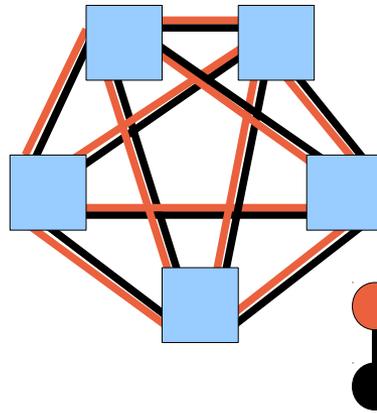
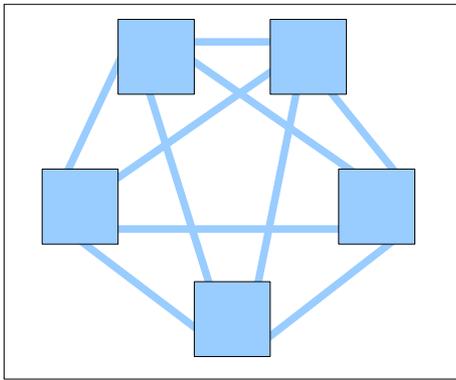
Cohesive group?

→ All entities share similar attributes

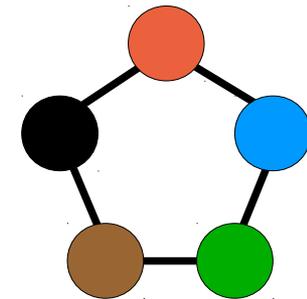
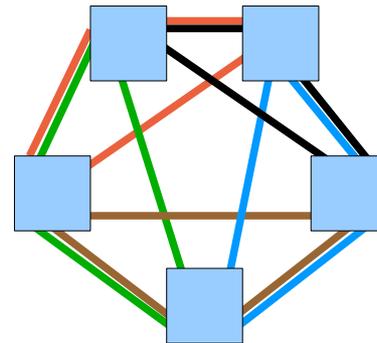
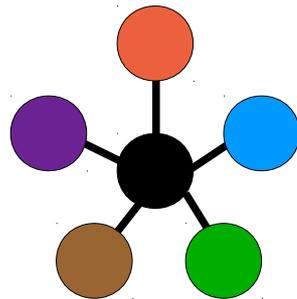
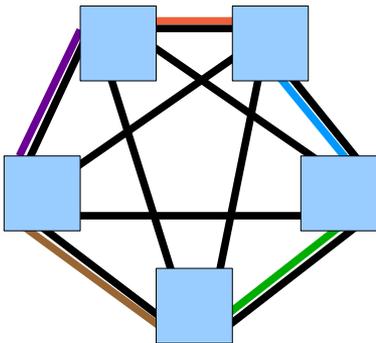
→ All attributes co-occur with one another

Attribute interaction and entanglement

→ How much do attributes interact across entities?



→ How tangled is the attribute interaction throughout the network?



Entanglement Measures

Entanglement: «attribute entanglement is re-inforced through interaction with other tangled attributes* »

$$g_s u = \sum_t c_{t,s} g_t \quad gu = C' g$$

$$C = \begin{bmatrix} c_{s,s} & c_{s,t} & \dots \\ c_{t,s} & c_{t,t} & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

g_s : entanglement index of an attribute s

u : maximum entanglement

$c_{s,t}$: conditional interaction frequency of an attribute t given another attribute s

→ g is thus an eigenvector of the interaction matrix

**Social ambiguity, Burt & Schott (1985)*

Entanglement Measures

$$N = \begin{bmatrix} n_{s,s} & n_{s,t} & \dots \\ n_{t,s} & n_{t,t} & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

$$C = \begin{bmatrix} c_{s,s} & c_{s,t} & \dots \\ c_{t,s} & c_{t,t} & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

N : **(Co)occurrence matrix** - symmetric

$N_{s,s}$: number of occurrence of attributes s over entities **edges**

$N_{s,t}$: number of cooccurrence of attributes s and t

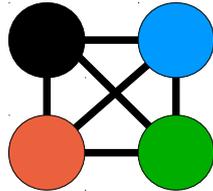
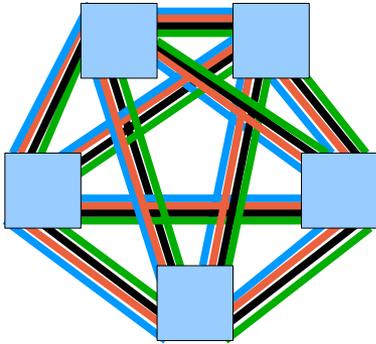
C : **interaction matrix** – not symmetric

$c_{s,s}$: occurrence frequency of an attribute s ($N_{s,t}/|E|$)

$c_{s,t}$: conditional cooccurrence frequency of an attribute t given another attribute s ($N_{s,t}/N_{s,s}$)

Entanglement in a cohesive group

Optimal group



$$C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

All entanglement indices $g_s = 1$

Eigenvalue equals number of attributes $|T|$ (non negative matrix theory)
 $u = 4$

Entanglement homogeneity and intensity

Homogeneity: how evenly attributes interaction distribute over the edges

$$\langle g, 1 \rangle$$

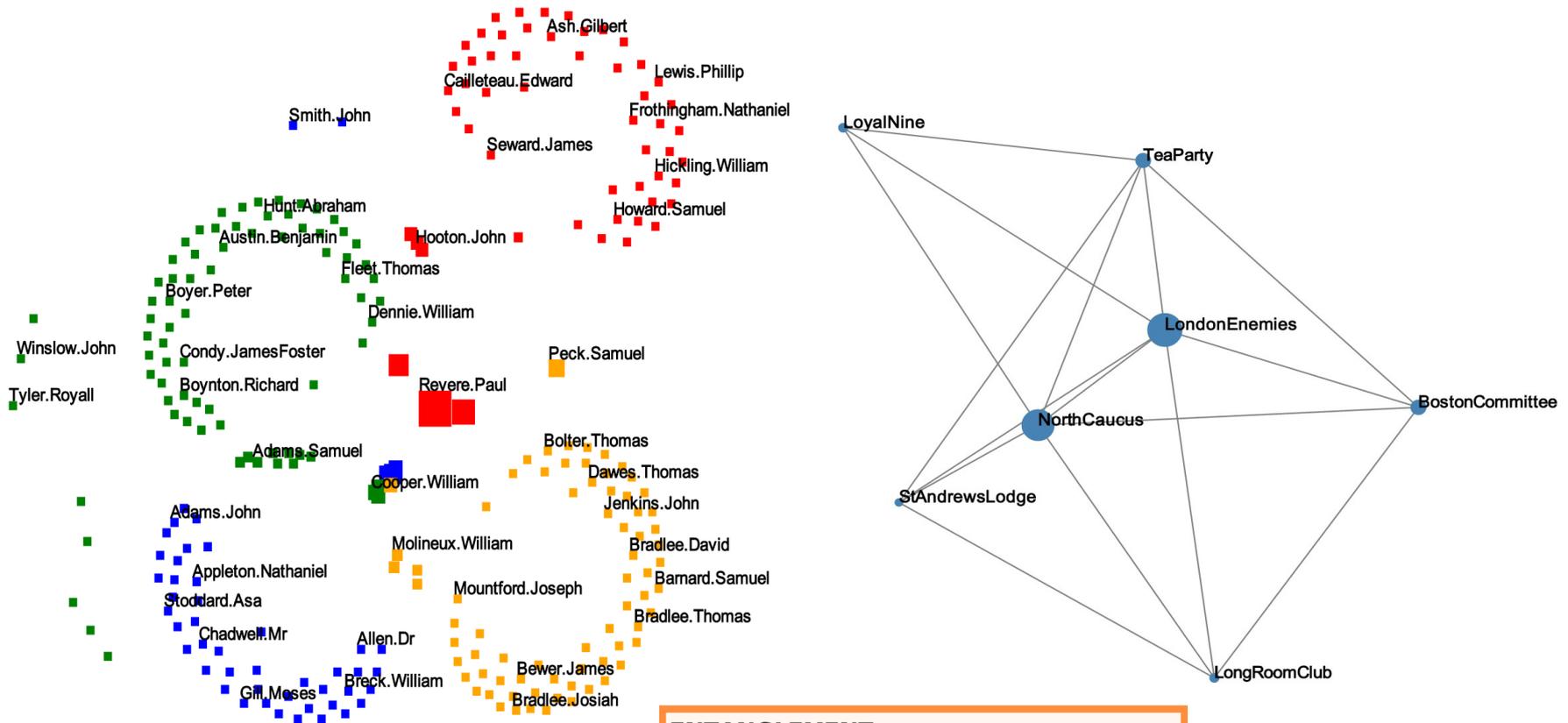
Intensity: how much attributes interact between entities

$$\frac{u}{|T|}$$

Using entanglement:

Paul Revere and the American Revolution

<http://kieranhealy.org/blog/archives/2013/06/09/using-metadata-to-find-paul-revere/>
and Shin-Kap Han (2009)



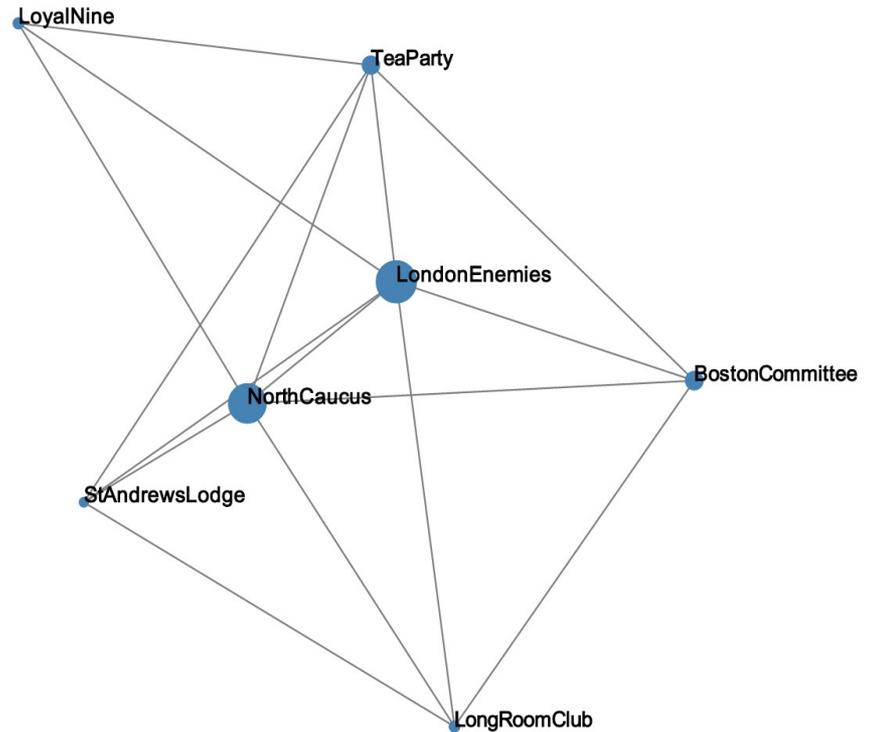
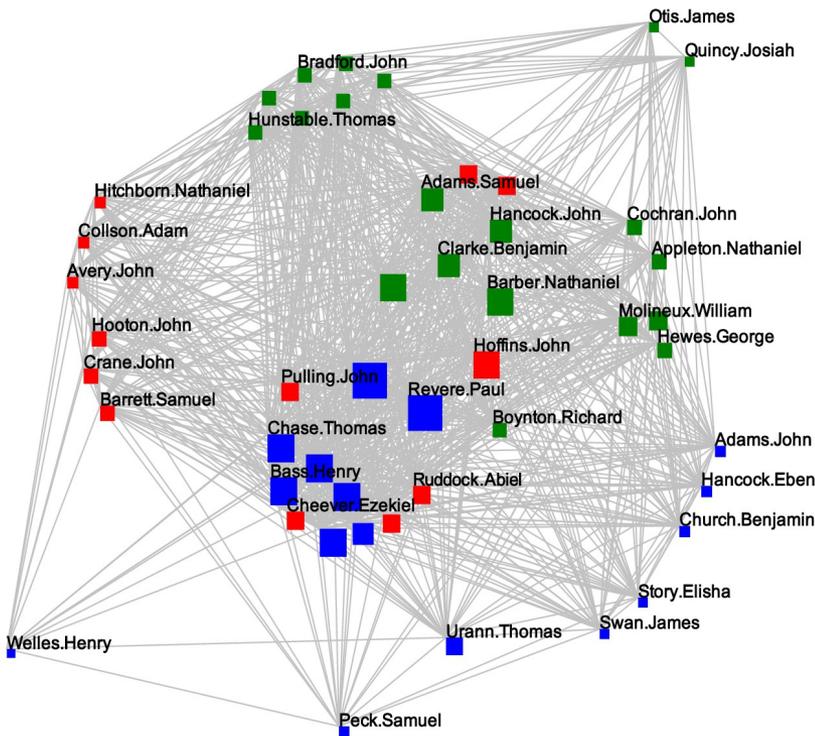
ENTANGLEMENT:

Intensity: 0.06065

Homogeneity: 0.69331

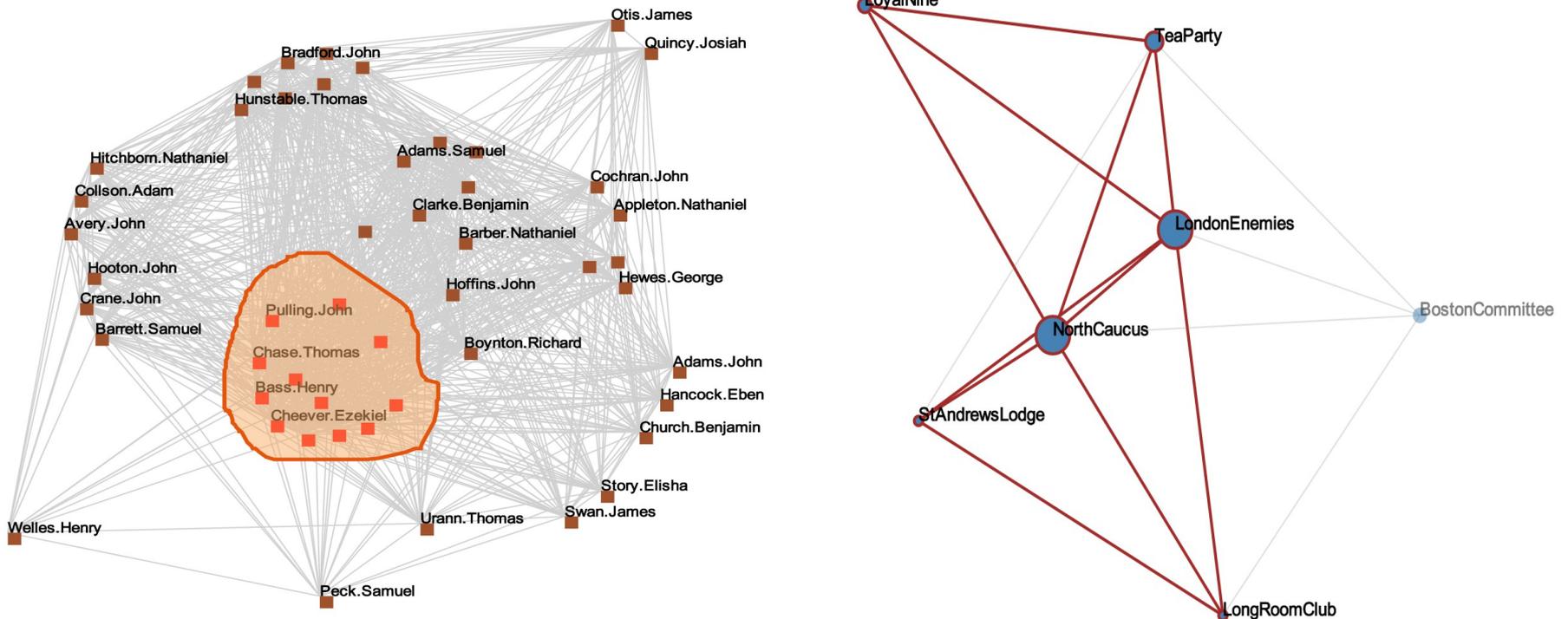
Using entanglement:

Paul Revere and the American Revolution



ENTANGLEMENT:
Intensity: 0.18242
Homogeneity: 0.7091

Using entanglement: Query from a group of people

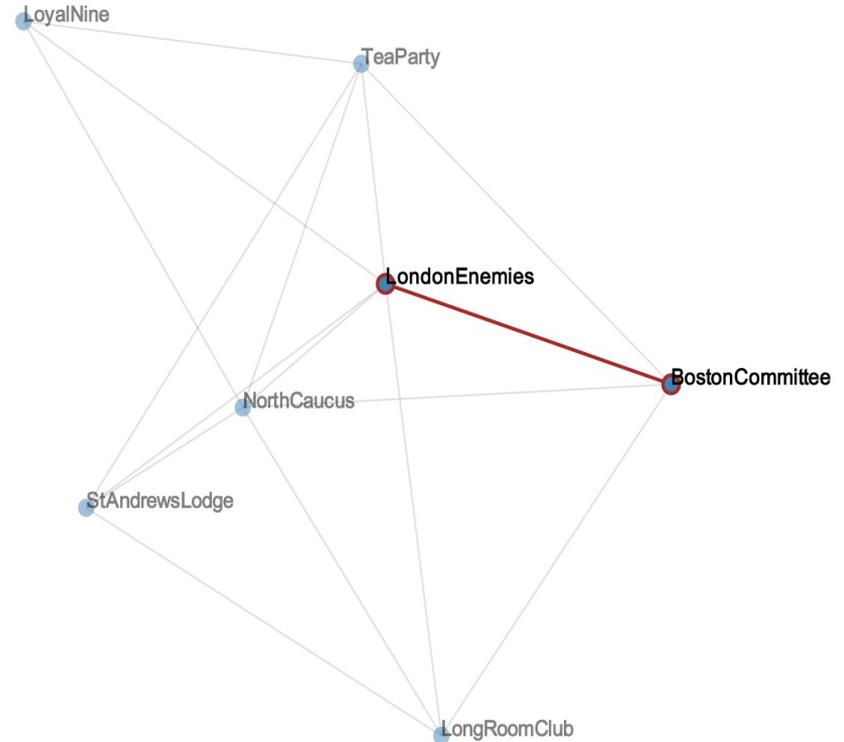
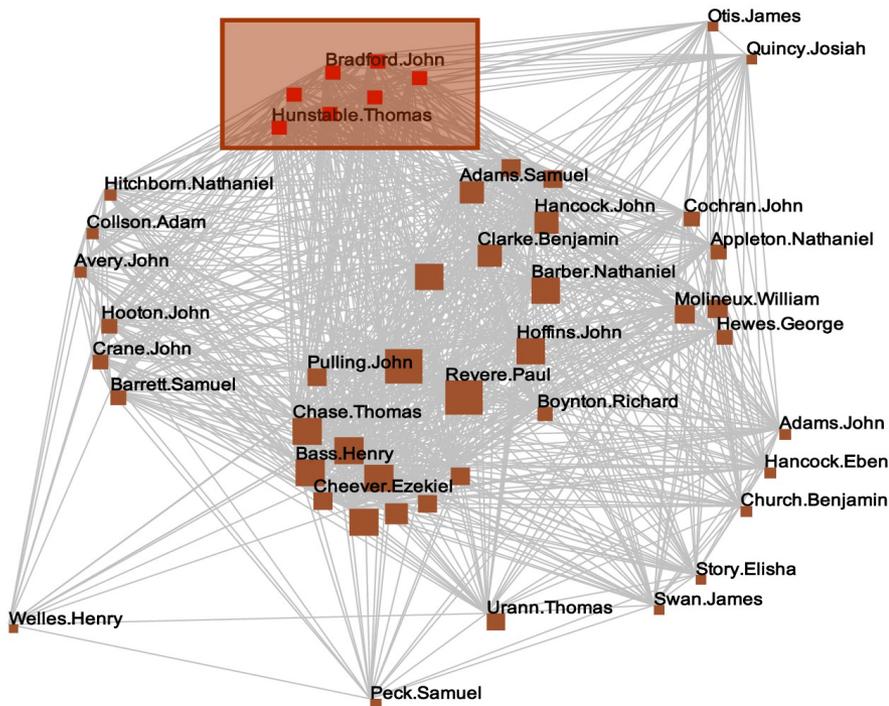


ENTANGLEMENT:

Intensity: 0.43296

Homogeneity: 0.71967

Using entanglement: Query from a group of people

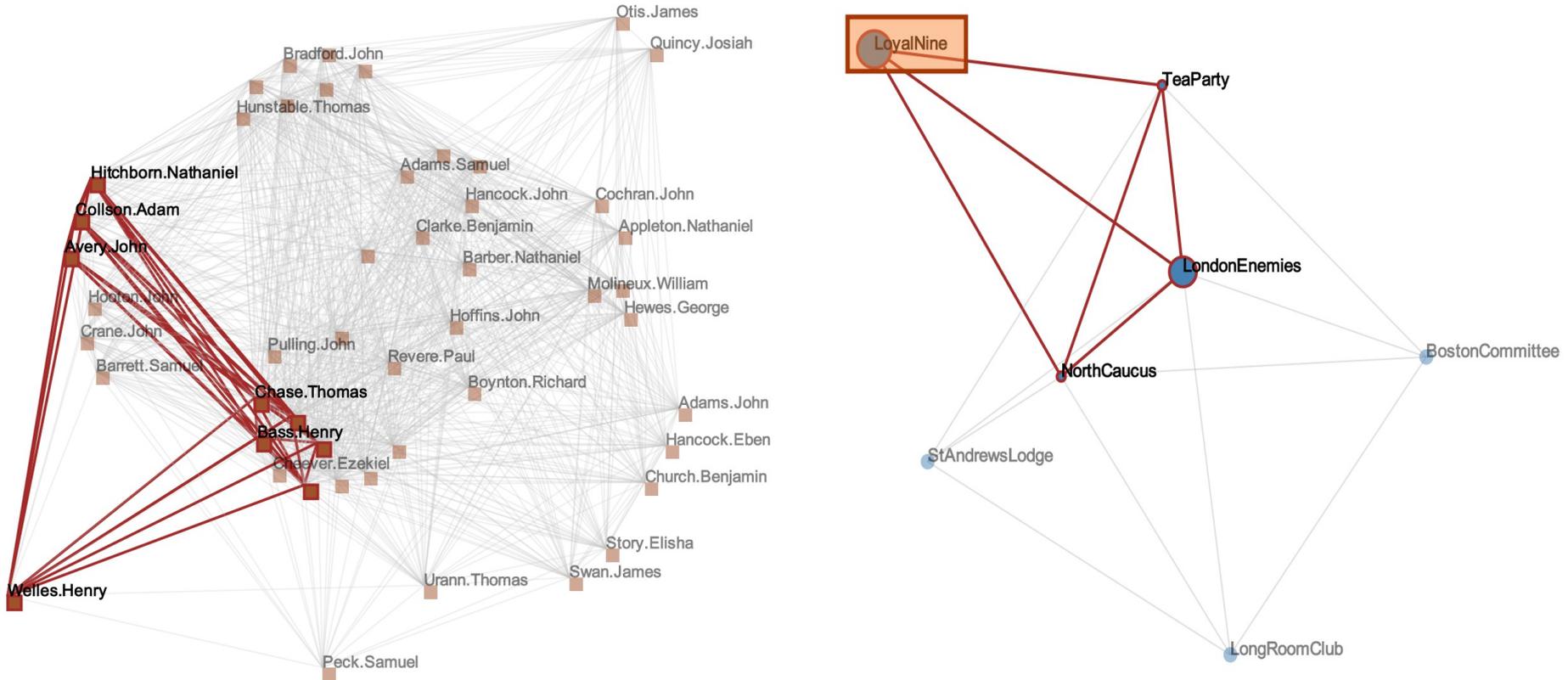


ENTANGLEMENT:

Intensity: 1
Homogeneity: 1

Using entanglement:

How loyal are the Loyal Nine ?



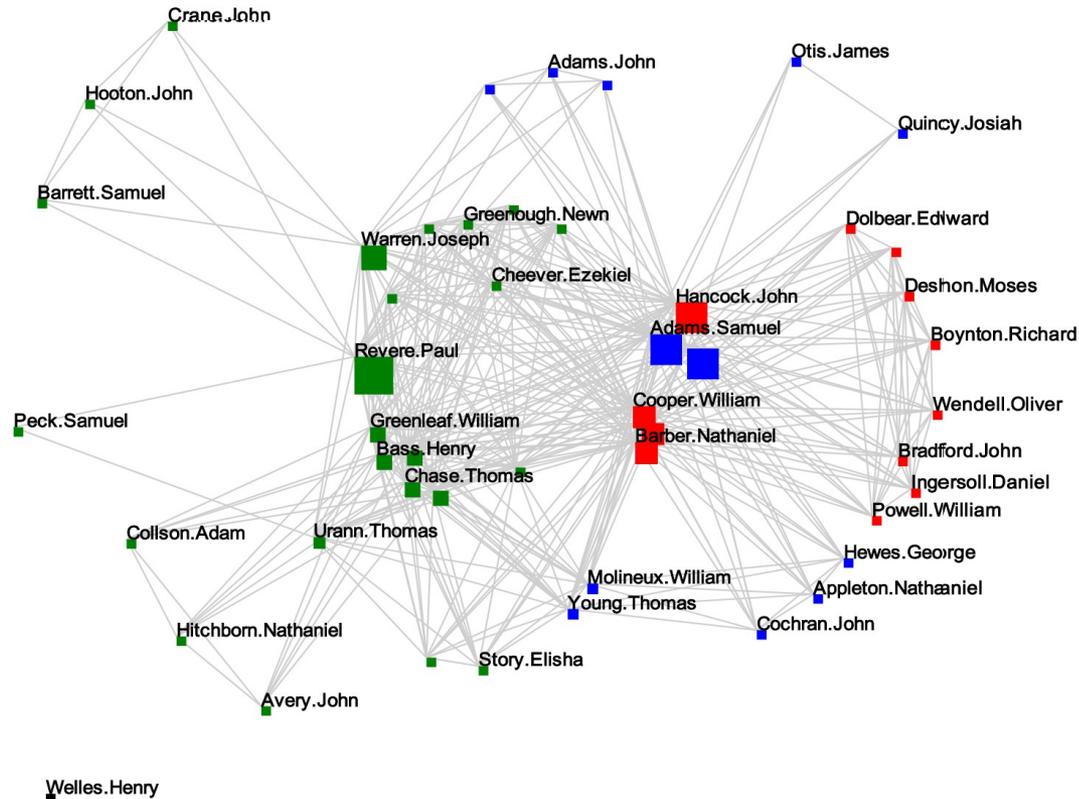
ENTANGLEMENT:

Intensity: 0.66908

Homogeneity: 0.93211

Using entanglement:

Going further, the network of acquaintances



ENTANGLEMENT:
Intensity: 0.2719
Homogeneity: 0.72834

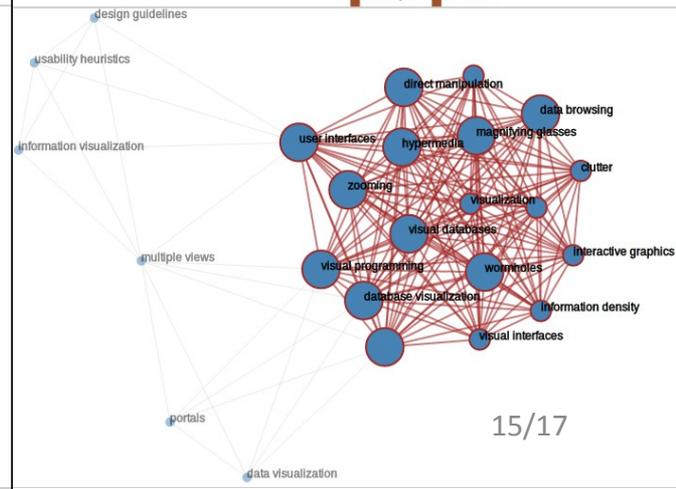
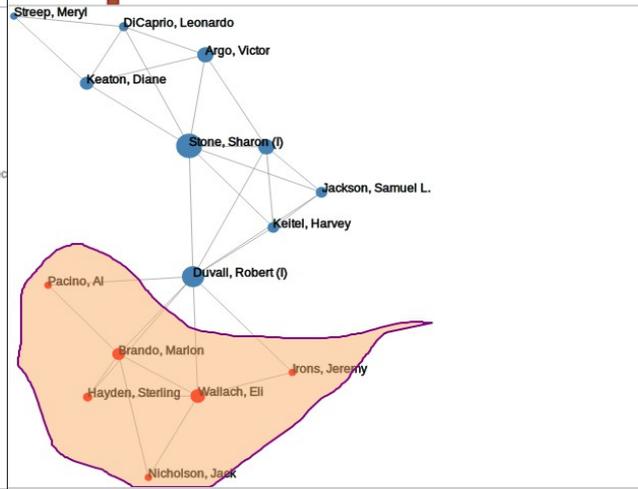
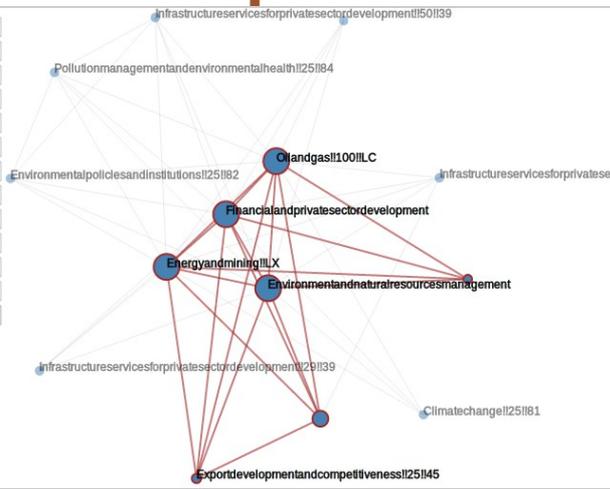
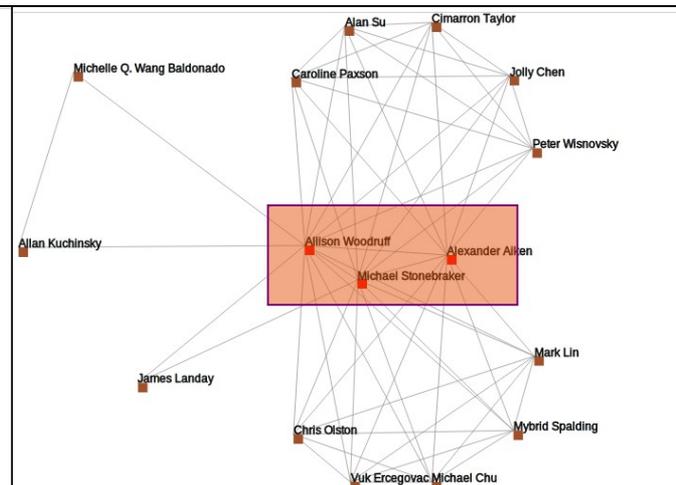
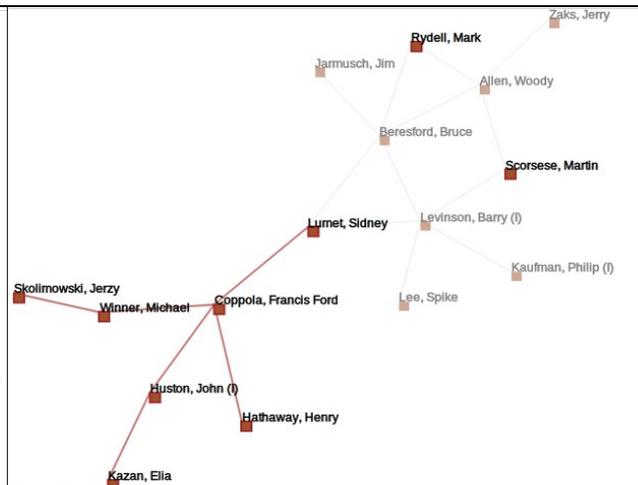
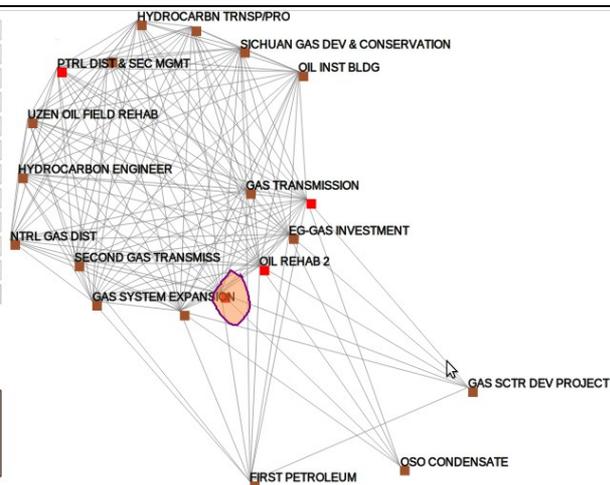
Transfer to other domains

2-mode graphs (first slides are documents/terms - OTMedia)

WorldBank Projects
projects/(sub)sectors
→ share same domain

IMDB
actors/directors
→ directed same actors

Coauthorship IV04 contest
authors/keywords
→ published on same topics



Future works

Weighted networks

Finding cohesive subgroups of entities

**For a live interactive demo,
please come to me after the presentation!**

Assessing group cohesion in homophily networks

Benjamin Renoust, benjamin.renoust@labri.fr

Guy Melançon, guy.melancon@labri.fr

Marie-Luce Viaud, mlviaud@ina.fr

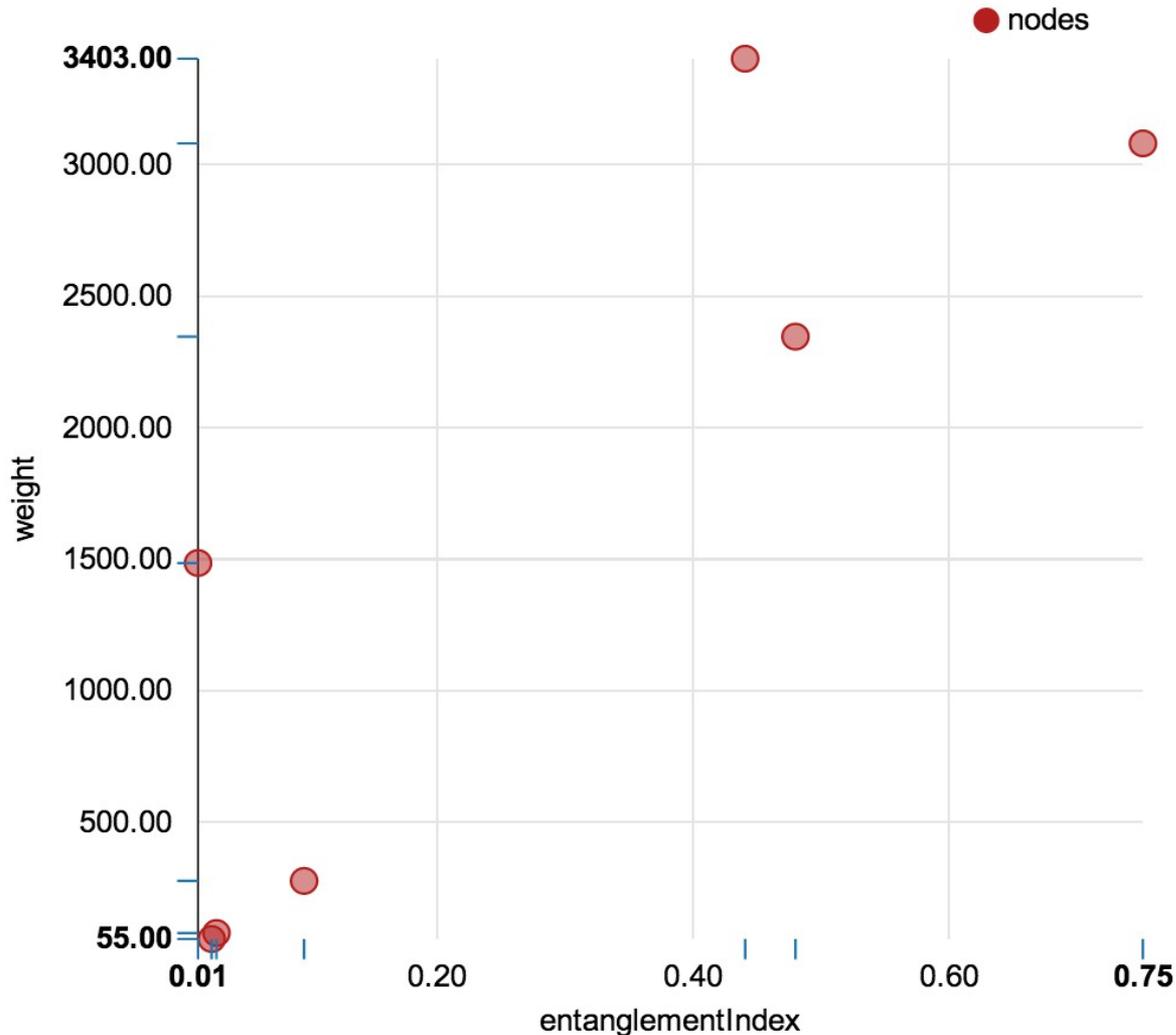
CNRS UMR 5800 LaBRI & INRIA Bordeaux Sud-Ouest

Université de Bordeaux I, France

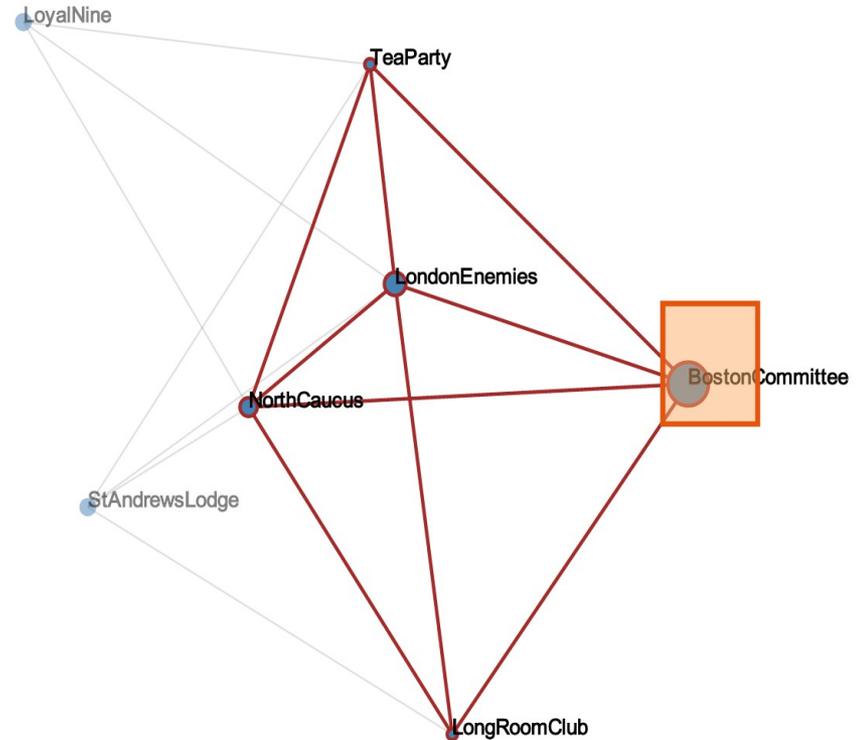
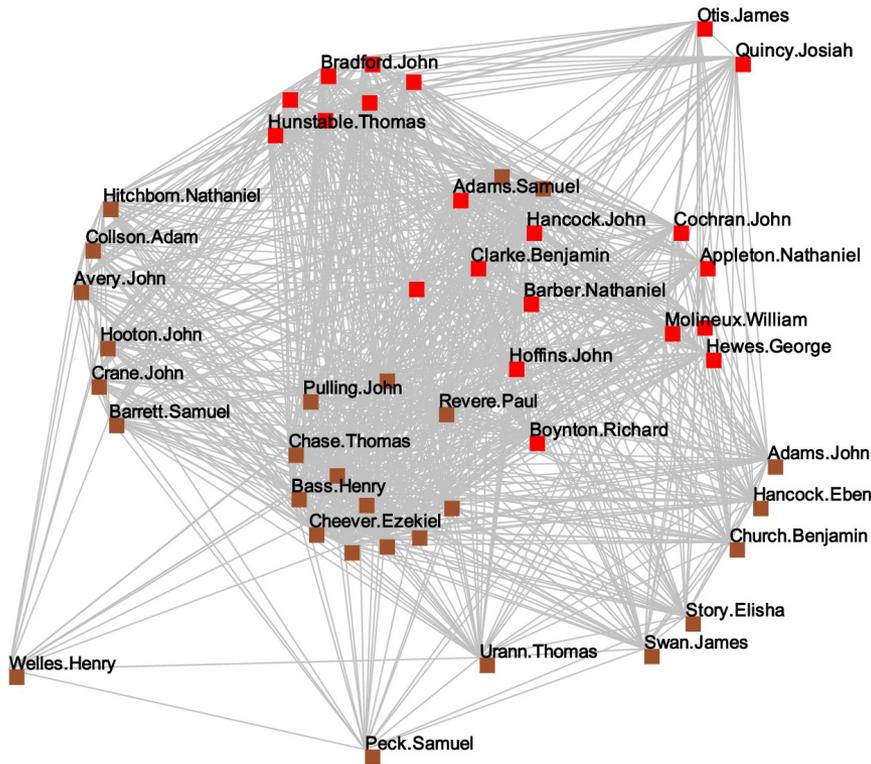
Institut National de l'Audiovisuel, France

Using entanglement:

Entanglement index differs from edge occurrence



Using entanglement: Query from a society



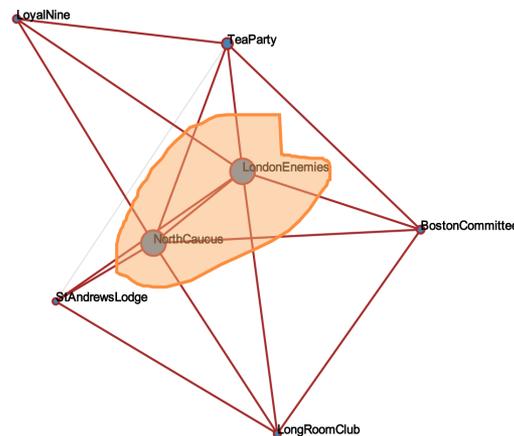
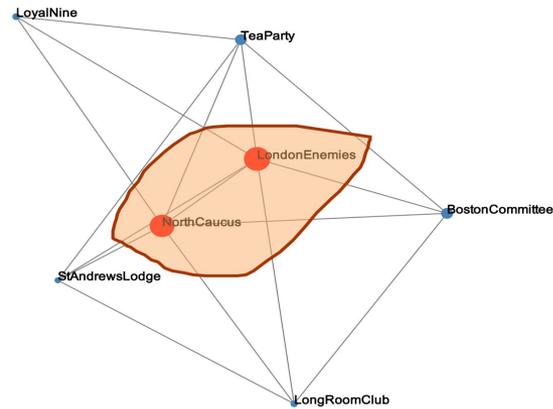
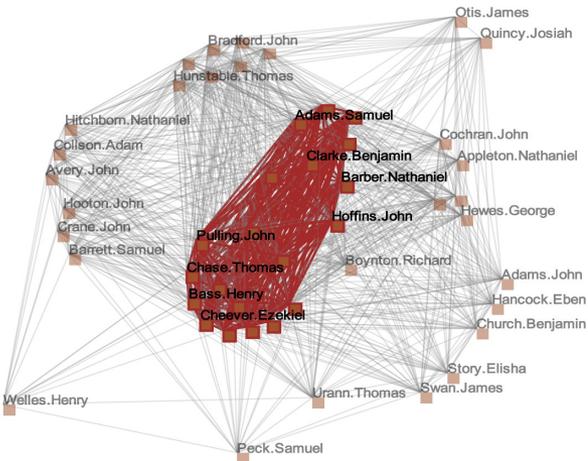
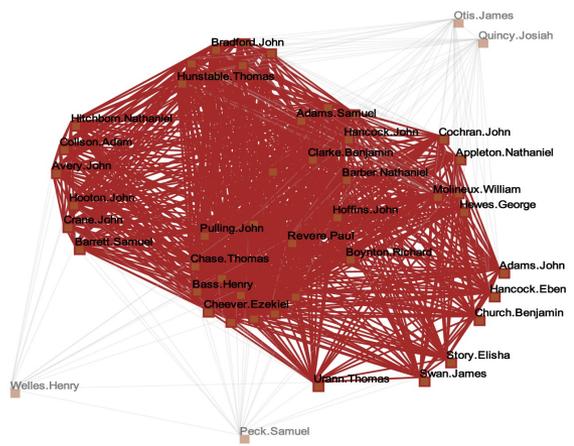
ENTANGLEMENT:

Intensity: 0.35507

Homogeneity: 0.71351

Using entanglement:

Two main societies, North Caucus and London Enemies



ENTANGLEMENT:

Intensity: 0.49502

Homogeneity: 0.9925

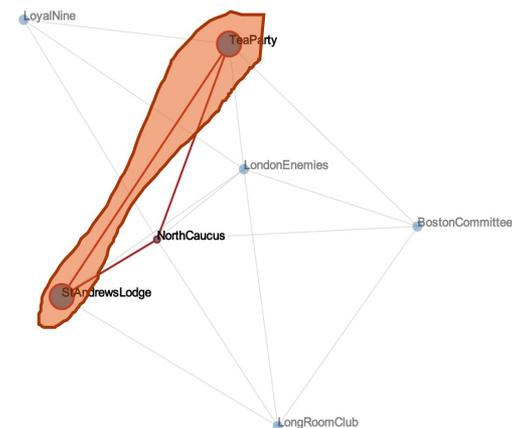
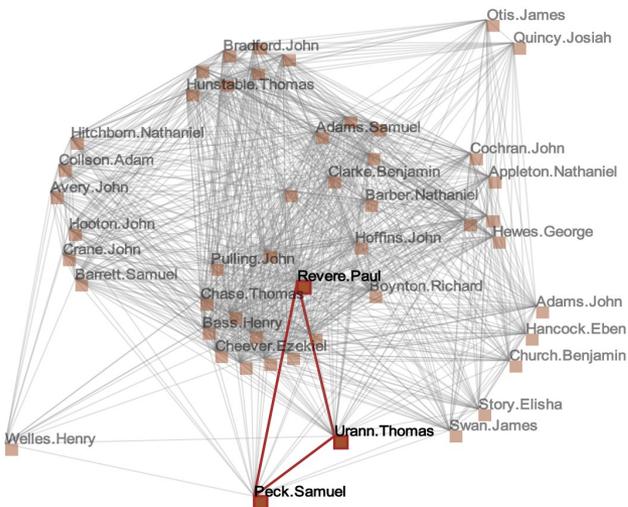
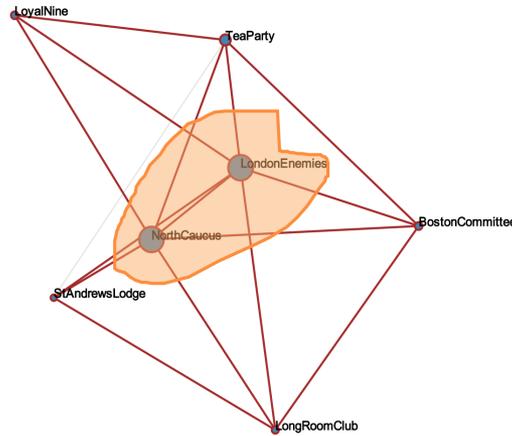
ENTANGLEMENT:

Intensity: 0.35105

Homogeneity: 0.64452

Using entanglement:

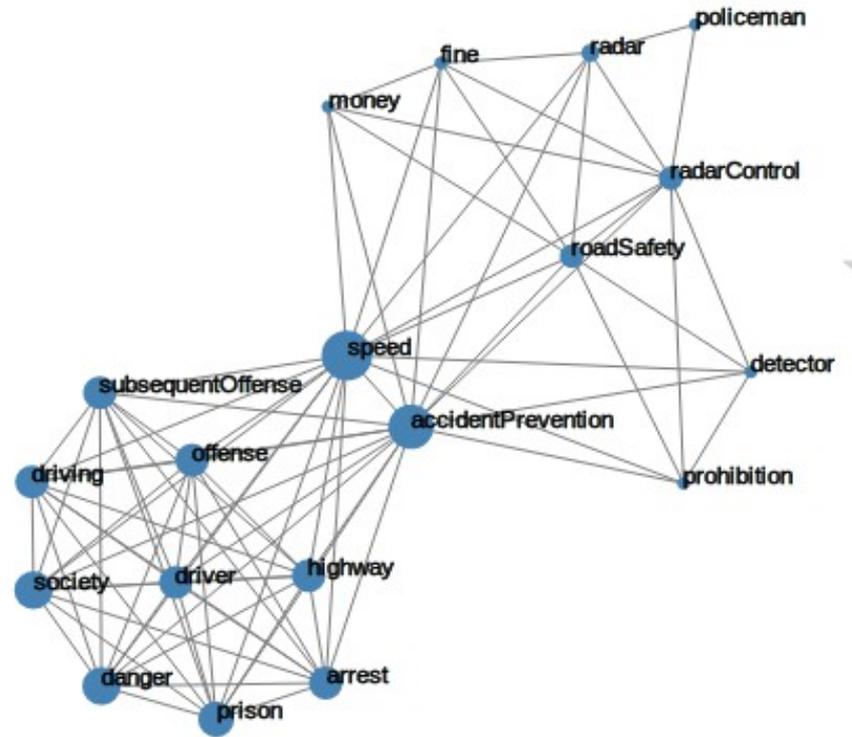
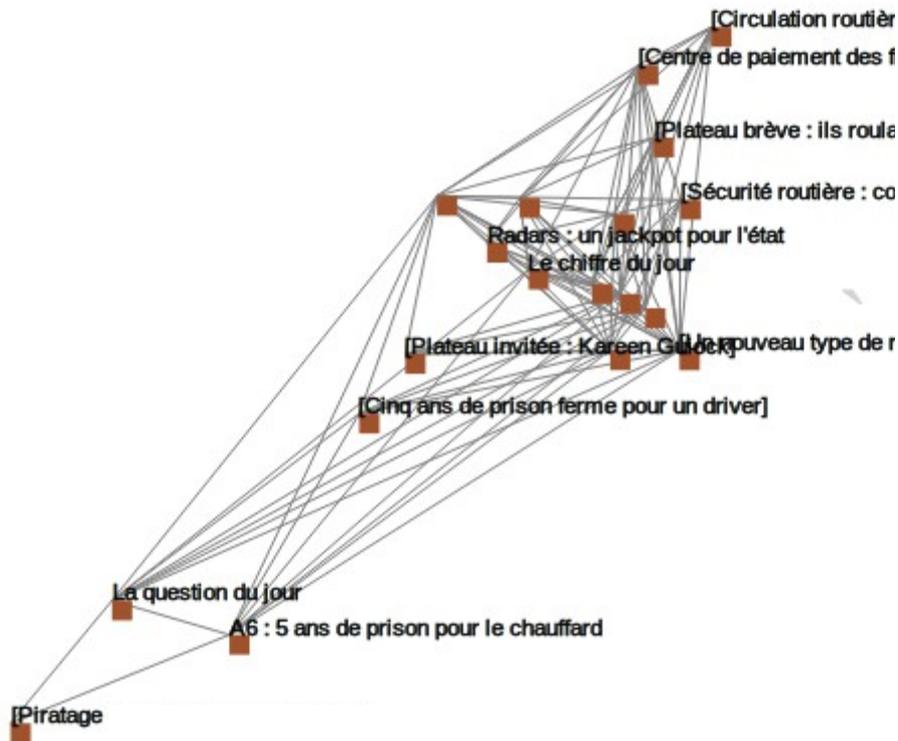
The missing link and Paul Revere



ENTANGLEMENT:
Intensity: 0.77778
Homogeneity: 0.92717

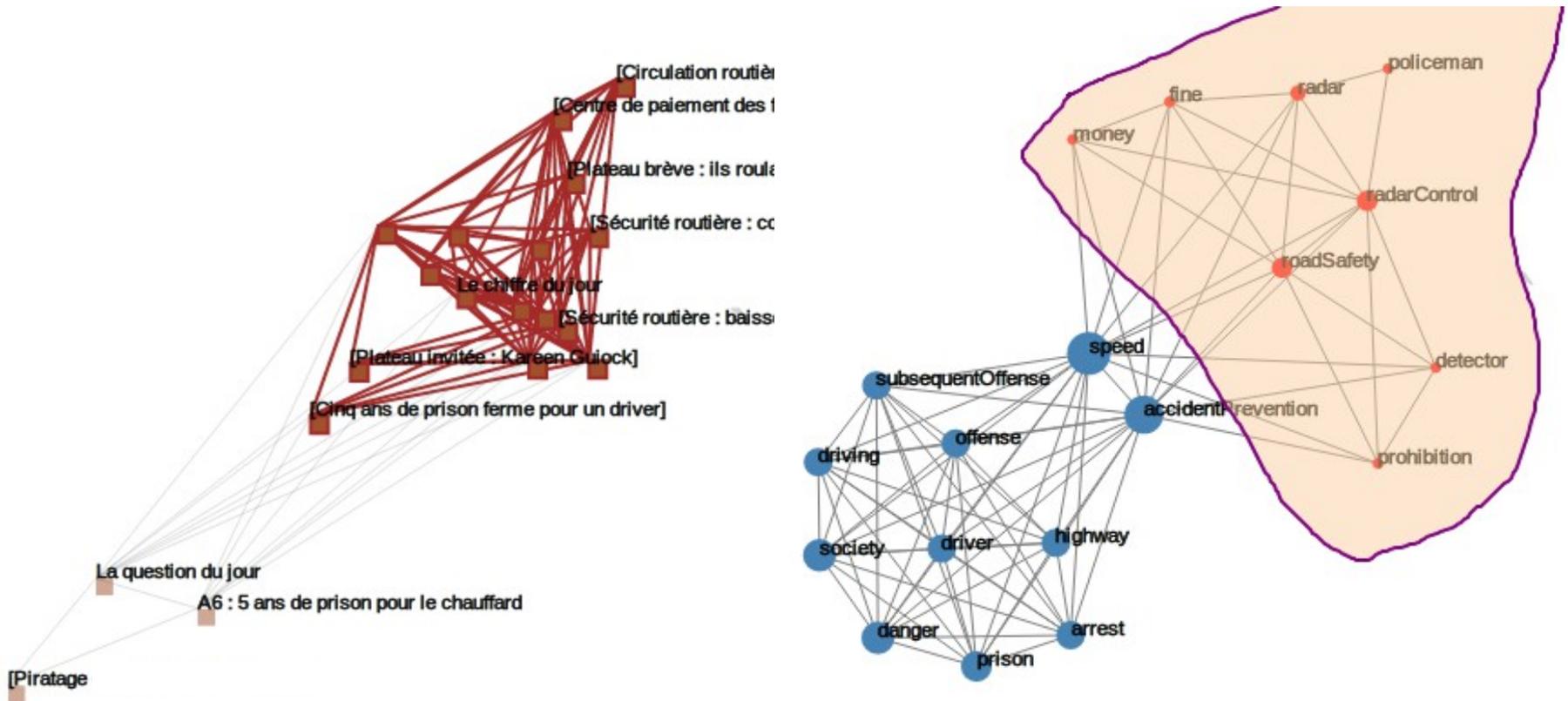
Using entanglement: Example

Intensity = 0.33 Homogeneity = 0.81



Using entanglement: Example

Intensity = 0.29 Homogeneity = 0.63



Using entanglement: Example

Intensity = 0.72 Homogeneity = 0.99

