

Socio-semantic dynamics in a blog network

Jean-Philippe Cointet

CREA (CNRS/EP, France)

AND

Camille Roth

CAMS (CNRS/EHESS, France)

IEEE SOCIALCOM09, VANCOUVER, BC – AUG 29–31, 2009

A SOCIAL network

Three kinds of links for each blog...

- **citation:** post citation links
- **interaction:** comment links
- **affiliation:** blogroll links

...where contents circulate

- in terms of topics (\mathcal{W})
- in terms of cultural items (\mathcal{U})

Dataset: US blogosphere

- **scope:** 4 months of '08 campaign
- **network:** citations
- **nodes:** 1, 066 blogs (RTGI)

versac.net

citation link

» Buzzes | Accueil | Ouverture et marché de l'énergie... »

29 juin 2007

Mouhahaha

Vraiment, c'est pathétique. Après la dénonciation par Nicolas Sarkozy de la pensée unique... de l'esprit de Mai 68, voilà celle qu'on attendait tous, la **pensée unique.com** !

NB : réviser, **homme de paille**.

29 juin 2007 à 10:45 | Lien permanent

TrackBack

URL TrackBack de cette note:
http://www.lypood.com/1/trackback/365111984712

Voici les sites qui parlent de **Mouhahaha**:

Commentaires

Vous rigolerez moins, si vous aussi, ON essayait de vous bawloner pour vous empêcher d'écrire tout haut ce que d'autres gribouillent tout bas !
Mais vous préférez vous complaire dans l'échange et la pensée droite-gauchiste qui règne sur l'ensemble des blogs à LEUR solde.

Blogue par: Behet | le 29 juin 2007 à 11:09

Un travail intéressant serait de recenser le nombre de "pensées uniques" concurrentes dans l'espace public.

comment link

blogroll link

LES NOTES RÉCENTES

Supplique retribue

Hahaha

On ne s'en base pas

Aucun ne de

Changement...

Néline

Vive Rachida Dati !

Catilland a vu Louis Change

Dimanche soir : liveblogging avec l'Id

UBAMA MCOFF en ligne

LES COMMENTAIRES RÉCENTS

genies Claire sur Supplique à retribue

berf sur Supplique à retribue

Marius sur Supplique à retribue

Timothée sur Hahaha

Timothée sur Supplique à retribue

Fabriz sur Supplique à retribue

versac sur Supplique à retribue

Vivace Traquaire sur "Aucun ne de

ET SUR PUBLIUS.FR

Bavéris en Europe par la Méditerranée

Les agro-carburants, cause de famine

La BCE paye en dollars

Un modèle de société en mistral

Municipales : et l'Europe dans tout ça ?

VERSAC : À PROPOS, CONTACTS

email : versac.fat@gmail.com

skype : versacsn

:: A propos de versac

POR INTÉRIEUR

:: sans moi

ROLL

:: A l'efful de nous

:: A l'heure antichambre

:: KJambert

:: Bernard Salarié

:: Benoitphilipe

:: Big Bang Blog

A socio-SEMANTIC network

Three kinds of links for each blog...

- **citation:** post citation links
- **interaction:** comment links
- **affiliation:** blogroll links

...where contents circulate

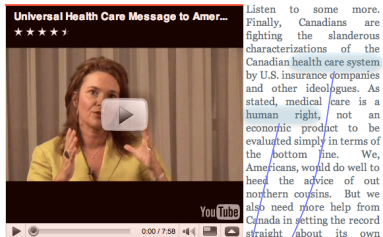
- in terms of topics (\mathcal{W})
- in terms of cultural items (\mathcal{U})

Dataset: US blogosphere

- **scope:** 4 months of '08 campaign
- **network:** citations
- **nodes:** 1,066 blogs (RTGI)

Setting the Record Straight: More from our Canadian Cousins

Written by Robert Justin Lipkin on August 28th, 2009



systematic attempt on the part of those Americans opposing health insurance reform to distort, obscure, and simply lie about the Canadian system. Click [here](#) for more.

\mathcal{U}

\mathcal{W}

semantic characterization

- “relevant” syntagms
 (“health insurance”, “climate change”, “national security”, “super Tuesday”, “human rights”...)
- urls: “www.youtube.com/x1hqweac”, etc.

A DYNAMIC socio-semantic network

Three kinds of links for each blog...

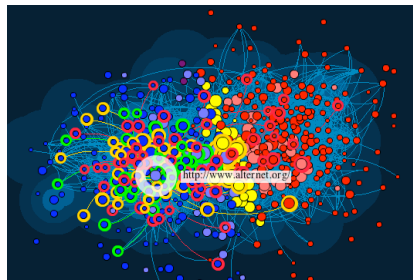
- **citation:** post citation links
- **interaction:** comment links
- **affiliation:** blogroll links

...where contents circulate

- in terms of topics (\mathcal{W})
- in terms of cultural items (\mathcal{U})

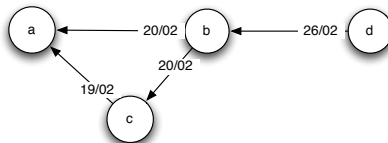
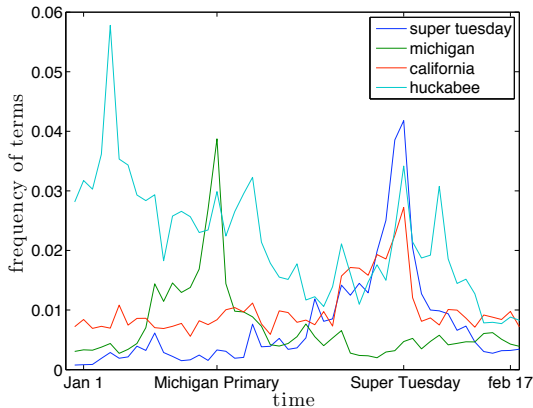
Dataset: US blogosphere

- **scope:** 4 months of '08 campaign
- **network:** citations
- **nodes:** 1,066 blogs (RTGI)



<http://presidentialwatch08.com/>

Socio-semantic configuration



Socio-semantic configuration

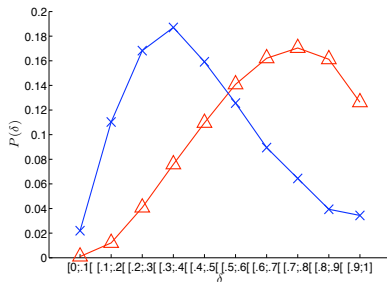
semantic profile of a blog i :

$$\hat{\mathbf{w}}_i(w) := \frac{\mathbf{w}_i(w)}{\sum_{w=1}^{|\mathcal{W}|} \mathbf{w}_i(w)} \cdot \log \frac{|\mathcal{B}|}{|\{j, \mathbf{w}_j(w) > 0\}|}$$

semantic distance

between blogs i and j :

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



Semantic distance distributions. *Triangles*: computed over the whole set of possible blog pairs. *Crosses*: distribution computed on linked blogs.

Computing link creation propensity

→ estimate the “propensity of interaction”
...that it is more or less likely for a node (or
a dyad) with property “ m ” to receive a link
...which may be simply estimated by:

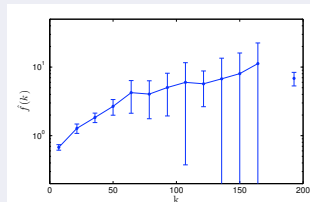
$$\hat{f}(m) = \frac{\nu(m)}{N(m)}$$

- $\nu(m)$ = number of links pointing towards an agent of type m (resp. number of new dyads of type m) during a time period,
- $N(m)$ = number of agents (resp. of dyads) of type m .

Computing link creation propensity

→ estimate the “propensity of interaction”
 ...that it is more or less likely for a node (or
 a dyad) with property “ m ” to receive a link
 ...which may be simply estimated by:

$$\hat{f}(m) = \frac{\nu(m)}{N(m)}$$



- $\nu(m)$ = number of links pointing towards an agent of type m (resp. number of new dyads of type m) during a time period,
- $N(m)$ = number of agents (resp. of dyads) of type m .

Dynamics of the social network

in-degree effects

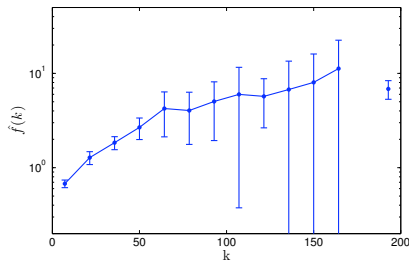
→ increasing, plateauing

topological distance effects

→ strong trend to repetition and local interaction

semantic distance

→ strong trend to homophily



Dynamics of the social network

in-degree effects

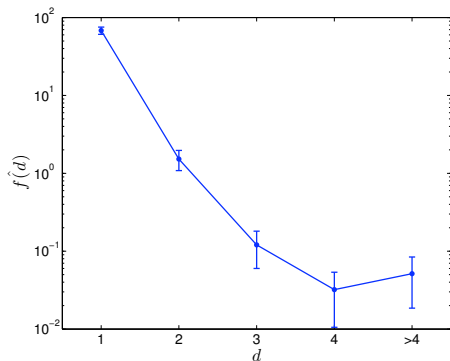
→ increasing, plateauing

topological distance effects

→ strong trend to repetition and local interaction

semantic distance

→ strong trend to homophily



Dynamics of the social network

in-degree effects

→ increasing, plateauing

topological distance effects

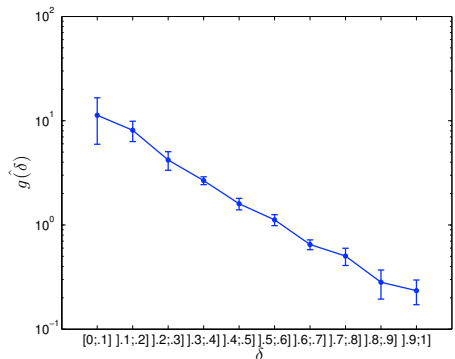
→ strong trend to repetition and local interaction

semantic distance

→ strong trend to homophily

primarily “social”?

→ increased distance and degree



Dynamics of the social network

in-degree effects

→ increasing, plateauing

topological distance effects

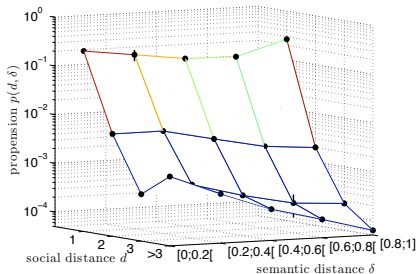
→ strong trend to repetition and local interaction

semantic distance

→ strong trend to homophily

primarily “social”?

● social distance and degree



Dynamics of the social network

in-degree effects

→ increasing, plateauing

topological distance effects

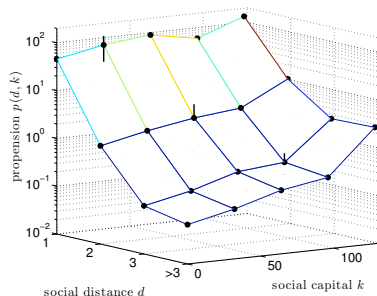
→ strong trend to repetition and local interaction

semantic distance

→ strong trend to homophily

primarily “social”?

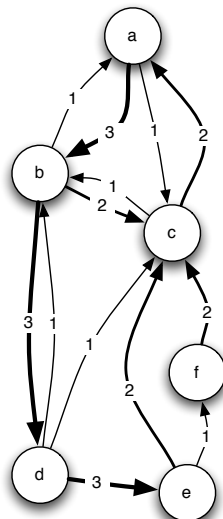
- social distance and degree



Information flows: measures on the post network

Dyadic measures:

- raw, weighted network, aggregated on 4 months
- attentional matrix a_{ij} → and total attention
 $\alpha_a = 5/6$
- detachment matrix



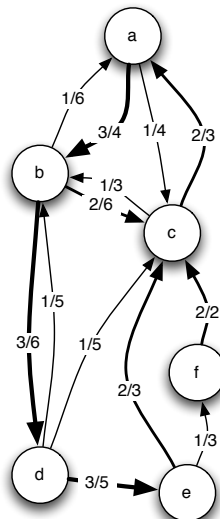
Information flows: measures on the post network

Dyadic measures:

- raw, weighted network, aggregated on 4 months
- attentional matrix **a**...
→ and total attention
 $\alpha_a = 5/6$
- detachment matrix

“edge range”:

- quantifying shortcuts



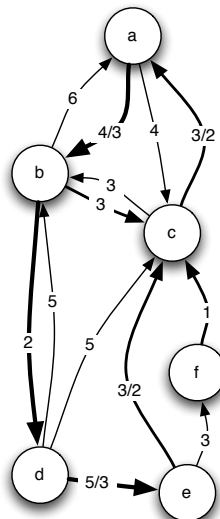
Information flows: measures on the post network

Dyadic measures:

- raw, weighted network, aggregated on 4 months
- attentional matrix **a**...
→ and total attention
 $\alpha_a = 5/6$
- detachment matrix

“edge range”:

- quantifying shortcuts



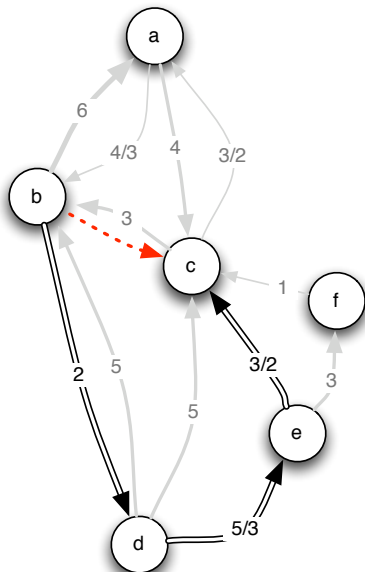
Information flows: measures on the post network

Dyadic measures:

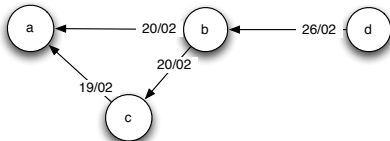
- raw, weighted network, aggregated on 4 months
- attentional matrix **a**...
→ and total attention
 $\alpha_a = 5/6$
- detachment matrix

“edge range”:

- quantifying **shortcuts**

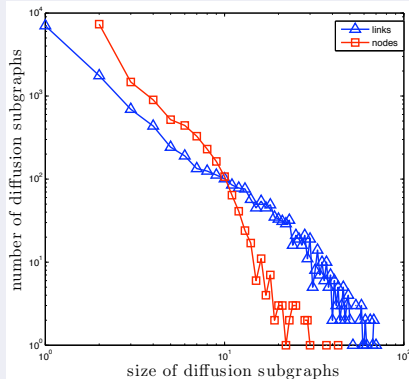


Information cascade



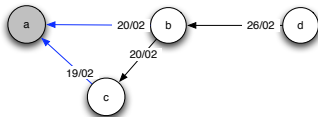
An example of diffusion subgraph, a common “resource” and a set of citation links between blogs

Diffusion subgraphs



⇒ heterogeneous cascade sizes

An ego-centered perspective

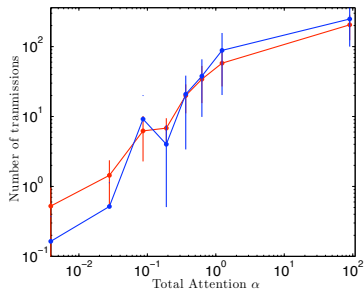


We focus on the total number of “transmissions” generated by blogs with a given total attention α

a bit more “global”...

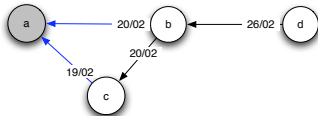
second transmissions: we focus on “later transmissions”, i.e. after a first transmission event

role of the *total attention* on the number of *diffusion links*



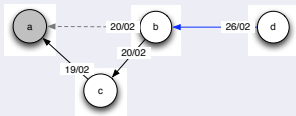
Larger *active readership* \Rightarrow larger number of diffusion links, **yet not linearly**

An ego-centered perspective



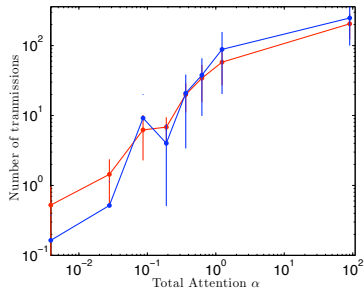
We focus on the total number of “transmissions” generated by blogs with a given total attention α

a bit more “global”...



second transmissions: we focus on “later transmissions”, i.e. after a first transmission event

role of the *total attention* on the number of *diffusion links*

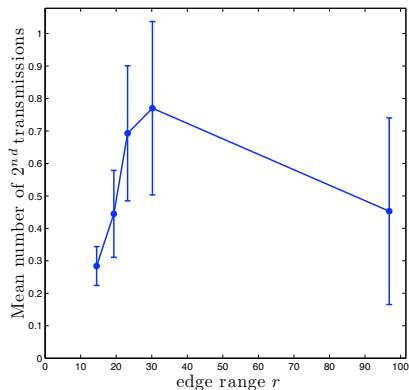
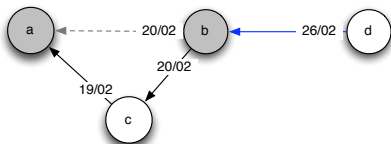


Larger *active readership* \Rightarrow larger number of diffusion links, **yet not linearly**

A more global perspective

→ role of *edge-range* on the number of grand-children

We focus again on transmissions occurring after a first transmission event



An information which has been transmitted through a “*median*” link generates a larger number of grandchildren

Concluding remarks

Co-evolution of content and relationships

- Patterns *not necessarily linked to authority only*
- Patterns *not necessarily ego-centered only*
 - divergent from the “neighbor-based-influence” perspective

Concluding remarks

Co-evolution of content and relationships

- Patterns *not necessarily linked to authority only*
- Patterns *not necessarily ego-centered only*
 - divergent from the “neighbor-based-influence” perspective

Thank you!

cointet@poly.polytechnique.fr & roth@ehess.fr