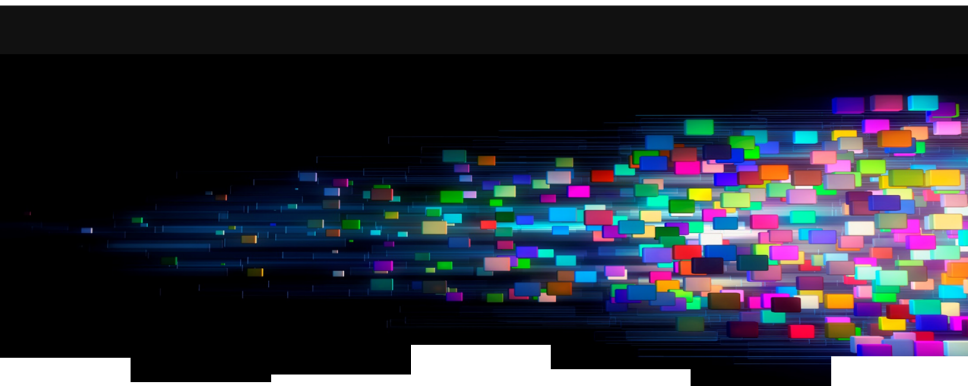


# Dynamiques dans les forums de discussions online

*Toulouse - October 2015*



Alberto Lumbreras, Marie Guegan (TCH)  
Julien Velcin (ERIC)  
Bertrand Jouve (CNRS)



1. Discussion Forums
2. Generative models
3. Some dynamics of discussion trees
4. Summary

1. Discussion Forums

2. Generative models

3. Some dynamics of discussion trees

4. Summary

# Discussion Forums

**/r/POLITIC**

commentsrelatedother discussions (1)

**62% of Republicans Feel BETRAYED by their Party** (thegatewaypundit.com)  
submitted 1 day ago by BakerANNMSU <Bakerism>  
28 comments SHARE save hide give gold REPORT other for anon comment

all 28 comments - sorted by **best**

save

reddiquette

formatting help

**[+] HugePurpleNipples** 11 points 1 day ago

They should, I used to be a Republican leaning indy, now I'm pretty staunchly liberal. It's tough to feel okay about supporting a party that does and says the awful stuff they do on a regular basis. They just come off as awful people most of the time, not that Hillary looks a lot better right now.  
permalink save report give gold reply

**[+] kevans2** 3 points 1 day ago

Sanders.  
permalink save parent report give gold reply

**[+] HugePurpleNipples** 2 points 1 day ago

Right there with you buddy.  
permalink save parent report give gold reply

**[+] HAL9000000** 1 point 1 day ago

Which parts specifically of the Republican Party do you feel betrayed by?  
permalink save parent report give gold reply

**[+] HugePurpleNipples** 1 point 1 day ago

Well, to pick a specific example, I was particularly troubled by the time a few of them said girls who were raped could decide not to get pregnant. The whole gay marriage thing is particularly troubling, I thought we were about SMALL GOVERNMENT? If you're about small government, you stay out of things that don't concern you. They're constantly on the wrong side of social issues, and they take so much money from rich people it makes me feel ill.  
  
They've totally betrayed their party values. I miss the hardcore cut the budget, shrink the government guys we used to get, we're far from that now.  
permalink save parent report give gold reply

**[+] 5bits** 0 points 1 day ago

What does it mean to lean independent, when an independent could have any political views?  
permalink save parent report give gold reply

**[+] rfinger1337** 3 points 1 day ago

It doesn't say lean independent. It said republican leaning indy. Or, "tends to vote republican, but can be convinced otherwise."  
permalink save parent report give gold reply

# Discussion Forums

Reddit, January 2015

- Power law distribution of posts/thread.

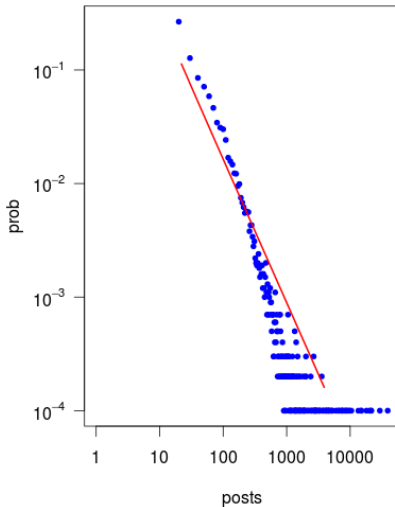
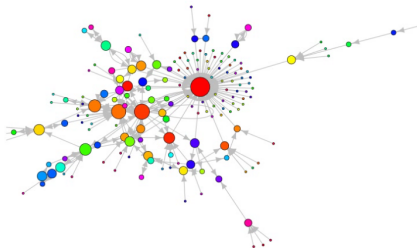


Figure 1: Posts per thread

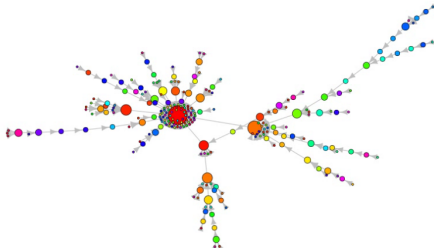
# Discussion Forums

## User graphs vs Posts graphs

Users



Posts

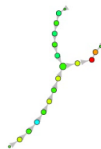


# Discussion Forums

User graphs vs Posts graphs

Users

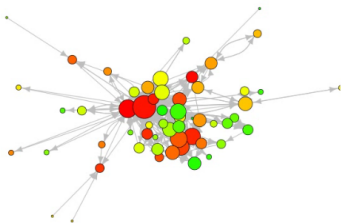
Posts



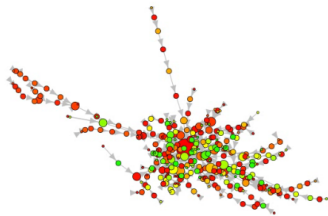
# Discussion Forums

## User graphs vs Posts graphs

Users



Posts





## Question

Can we **model** the **dynamics** of a **conversation tree**?

- To understand and **summarize** its internal mechanisms.
- To **predict** the future of a conversation.

1. Discussion Forums

2. Generative models

3. Some dynamics of discussion trees

4. Summary

# Generative models

## Modeling graphs

- Simple generative process governed by few parameters as possible.
- Artificial graphs reproduce some properties of the real graphs.
- Equiprobable edges.
- Preferential Attachment.

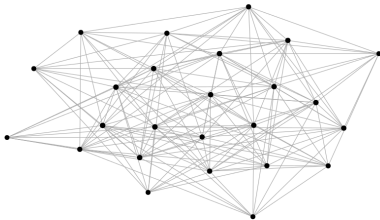


Figure 2: Erdos-Renyi  
 $n = 25$   
 $p = 0.5$

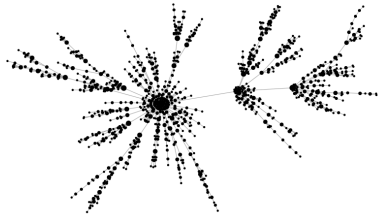


Figure 3: Barabasi-Albert  
 $n = 1000$   
 $p = 1$

# Generative models

## Modeling online conversations

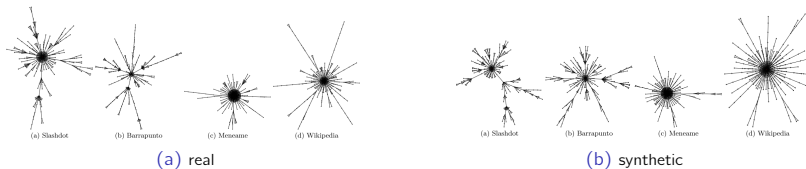


Figure 4: Gómez-Kappen-Kaltenbrunner [?]

Parameters:  $\alpha_1, \alpha_c, \beta$

- Number of responses to root. ✓
- Number of children. ✓
- Number of descendants.
- Node depth. ✓
- Size vs Proportion of responses to root. ✓

- Speeds. ✗
- Sequences (e.g.: a-b-a-b). ✗
- Text. ✗
- ...

# Generative models

## Modeling online conversations

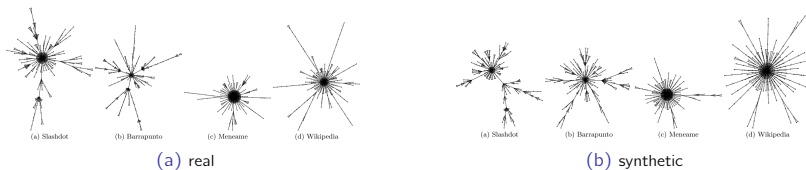


Figure 5: Gómez-Kappen-Kaltenbrunner [?]

Parameters:  $\alpha_1, \alpha_c, \beta$

$$p(\pi_t = k | \pi_{1:t-1}) \propto (\beta_k d_{k,t})^{\alpha_k}$$

$\beta_k$ : root bias

$\alpha_k$ : Preferential Attachment exponent.

- Estimate model parameters after conversations are finished (once and forever!)

## Generative models

### Question

Can we improve this model to simulate **real time**? (speed changes, bursts, slow down...)

Can we create a non-parametric growth model that continually re-evaluates its **predictions** on where the next post will be?

(that will allow for predictions on  $G_{t+1}$  given  $G_{1:t}$ )

("link prediction" for conversations)

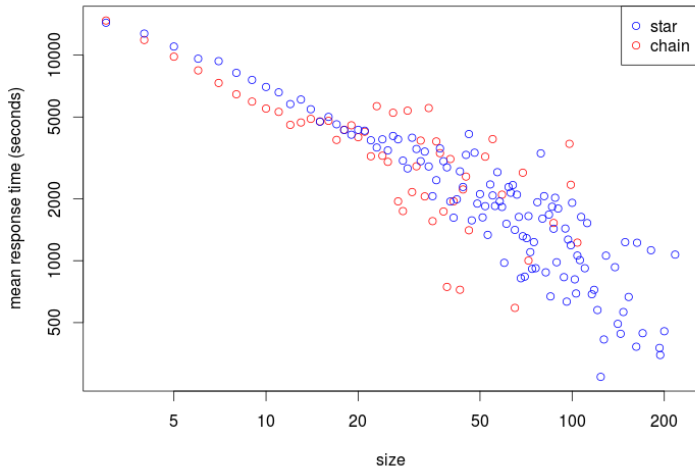
1. Discussion Forums
2. Generative models
3. Some dynamics of discussion trees
4. Summary

# Discussion trees

Different structures, different speeds?

- Time between posts is shorter in chains.

speed of of starts and chains

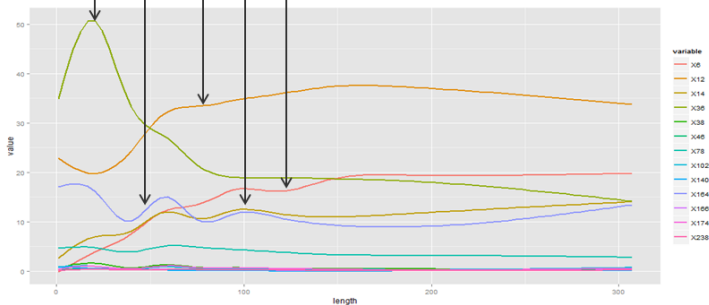




# Discussion trees

Different thread lengths, different triadic structures

Motif															
Motif ID			36	164	12	14	6	78	38	174	166	46	238	102	140



?

1. Discussion Forums
2. Generative models
3. Some dynamics of discussion trees
4. Summary

# Conclusions

Current:

- Good graph models that explain degree distributions.
- Good thread models.
- Parametric models.

But:

- Changes of rhythm, burst, etc not modeled (time is homogeneous).
- $a-b-a-b-a-b = a-b-c-d-e-f$ .
- Can not predict (therefore recommend posts to users).

Therefore we would like:

- Growth model that considers relation between structure and speed.
- Growth model that considers sequences (authorship).
- Link prediction for conversation trees.