

# Evaluation de l'influence sur Twitter: Application au projet "Twitter aux Elections Européennes 2014"



Lobna Azaza 

Sergey Kirgizov 

Eric Leclercq 

Marinette Savonnet 

Alexander Frame 

 Laboratoire Electronique,  
Informatique et Image

 Centre Interlangues Text, Image,  
Langage

Université de Bourgogne, Dijon

Lyon, 13 Mai

# Plan

- 1 Introduction
- 2 Influence sur les réseaux sociaux
- 3 Croyance en l'influence
- 4 Application au réseau de Twitter
- 5 Conclusion et perspectives

# Introduction

Société : Sur quoi les politiciens parlent (dans Twitter) ?  
Combien de personnes les écoutent ?  
Comment l'information se propage ?



Politiciens : Comment utiliser mieux le Twitter ?

Questions philosophiques et scientifiques :

Qu'est-ce que c'est l'influence ?  
Comment pouvons-nous la formaliser ?  
Comment vérifier notre formalisation ?

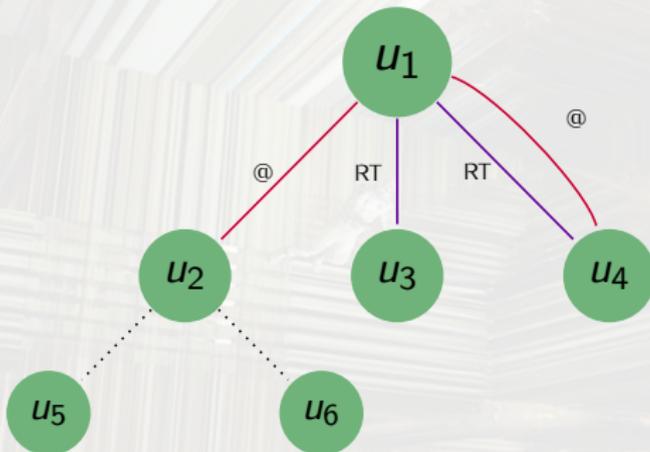
# Notre données

Projet internationale multidisciplinaire TEE 2014 :  
"Twitter aux élections européennes"

Comptes utilisateurs }  
hashtags }  $\xrightarrow[\text{pendant 50 days}]{\text{l'outil SNFreezer}}$   $\approx 40$  millions de tweets  
mots/phrases-clefs }

**Corpus français = 616 candidates et 4 millions tweets**

# Graphe multirelationnel du twitter



Nœuds : utilisateurs de twitter  
Liens : différents relations

# Influence

“The potential of a user action to initiate a further action by another user” — [Leavitt *et al.*, 2010]

Actions (relations) : tweets, retweets, follows, réponses, favoris, mentions, #hashtags, etc..

## Candidats les plus influentes selon différentes relations

Rank	<i>Retweet</i>	<i>Mention</i>	<i>Reply</i>
1	Marine Le Pen	Marine Le Pen	Christine Boutin
2	Florian Philippot	Christine Boutin	Marine Le Pen
3	Jean-Luc Mélenchon	Jean-Luc Mélenchon	Florian Philippot
4	Aymeric Chauparde	Florian Philippot	Jean-Luc Mélenchon
5	François Asselineau	Nicolas Dupont-Aignan	Louis de Gouyon Matigon
6	Corinne Morel-Darleux	José Bové	Nicolas Dupont-Aignan
7	Nicolas Dupont-Aignan	Aymeric Chauparde	Jean-Sébastien Herpin
8	Louis Aliot	Raquel Garrido	Julien Rochedy
9	Denis Payre	Jérôme Lavrilleux	Geoffroy Didier
10	Yannick Jadot	Marielle de Sarnez	Louis Aliot

# Comment mesurer l'influence ?

*"The number of retweets alone cannot be considered as a good indicator of influence"*

On the importance of considering social capitalism when measuring influence on Twitter

[Danisch, Dugué et Perez, 2014]

# Travaux existants

Auteurs	Réseau	Critères
Leavitt <i>et al.</i> , 2010	Twitter	Followers, réponses, retweets, mentions
Cha <i>et al.</i> , 2009	Flickr	Réponses + favoris
Cha <i>et al.</i> , 2010	Twitter	Followers, retweets, mentions
Wang <i>et al.</i> , 2010	Twitter	Similarité de thèmes + structure

De nombreux travaux utilisent les critères séparément, mais peu d'entre eux se sont intéressés à combiner les critères.

# Théorie des probabilités vs. la théorie de Dempster-Shafer

Ensemble des degrés de l'influence :  $\Omega = \{\text{Forte, Faible}\}$

**Proba**

$$p : \Omega \rightarrow [0, 1]$$

$$\sum_i p_i = 1$$

$$\text{Forte} \stackrel{p}{\mapsto} 0.3$$

$$\text{Faible} \stackrel{p}{\mapsto} 0.7$$

**Dempster-Shafer**

$$m : 2^\Omega \rightarrow [0, 1]$$

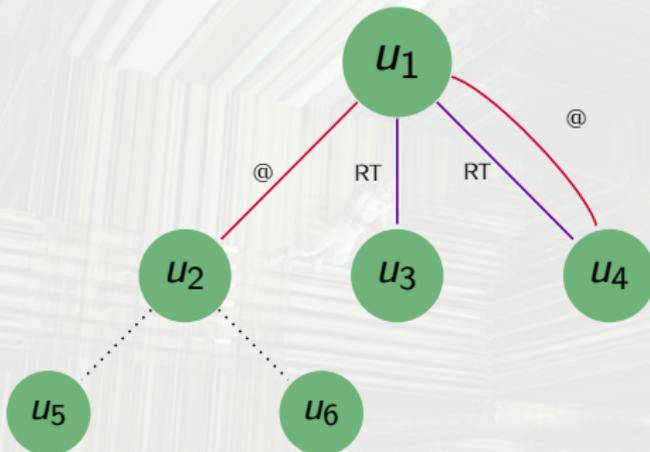
$$\sum_i m_i = 1$$

$$\text{Forte} \stackrel{m}{\mapsto} 0.3$$

$$\text{Faible} \stackrel{m}{\mapsto} 0.4$$

$$\{\text{Forte, Faible}\} \stackrel{m}{\mapsto} 0.3$$

# Graphe multirelationnel du twitter



Nœuds : utilisateurs de twitter  
Liens : différents relations

## Différents relations — différents masses de croyance

Ensemble des degrés de l'influence :

$\Omega = \{\text{Très Faible, Faible, Assez Moyenne, Moyenne, Assez Forte, Forte, Très Forte, Extrêmement Forte}\}$

$$\text{Retweet} \mapsto \begin{cases} m_{\text{retweet}}(\text{Très Faible}) = 0.6 \\ m_{\text{retweet}}(\text{Faible}) = 0.4 \end{cases}$$

$$\text{Mention} \mapsto \begin{cases} m_{\text{mention}}(\text{Très Faible}) = 0.7 \\ m_{\text{mention}}(\text{Faible}) = 0.3 \end{cases}$$

# Combinaison de masses de croyance

$$1 \text{ Retweet} \mapsto \begin{cases} m_{\text{retweet}}(\text{Très Faible}) = 0.6 \\ m_{\text{retweet}}(\text{Faible}) = 0.4 \end{cases}$$

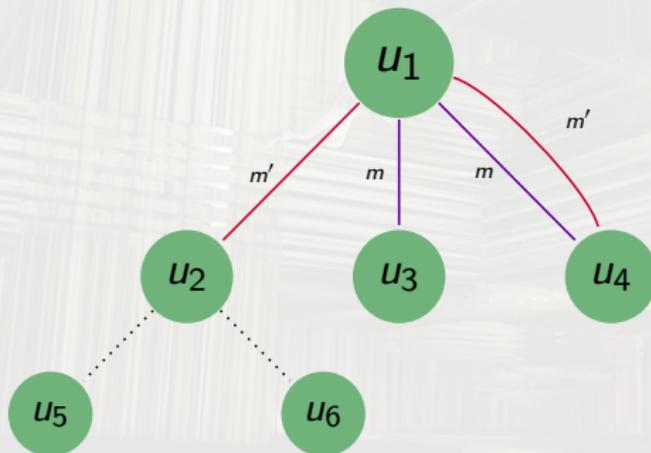
$\otimes$  : la règle de combinaison de Smets modifiée

$$2 \text{ retweets} \mapsto m_{\text{retweet}} \otimes m_{\text{retweet}}$$

$$2 \text{ Retweets} \mapsto \begin{cases} m_{\text{retweet}}(\text{Faible}) = 0.36 \\ m_{\text{retweet}}(\text{Assez Moyenne}) = 0.64 \end{cases}$$

# Combinaison de masses de croyance

$$(m \otimes m) \otimes (m' \otimes m')$$



## Candidats les plus influentes selon différentes relations

Rank	<i>Retweet</i>	<i>Mention</i>	<i>Reply</i>
1	Marine Le Pen	Marine Le Pen	Christine Boutin
2	Florian Philippot	Christine Boutin	Marine Le Pen
3	Jean-Luc Mélenchon	Jean-Luc Mélenchon	Florian Philippot
4	Aymeric Chauparde	Florian Philippot	Jean-Luc Mélenchon
5	François Asselineau	Nicolas Dupont-Aignan	Louis de Gouyon Matigon
6	Corinne Morel-Darleux	José Bové	Nicolas Dupont-Aignan
7	Nicolas Dupont-Aignan	Aymeric Chauparde	Jean-Sébastien Herpin
8	Louis Aliot	Raquel Garrido	Julien Rochedy
9	Denis Payre	Jérôme Lavrilleux	Geoffroy Didier
10	Yannick Jadot	Marielle de Sarnez	Louis Aliot

# Candidats les plus influentes selon notre méthode

## Résultats symboliques

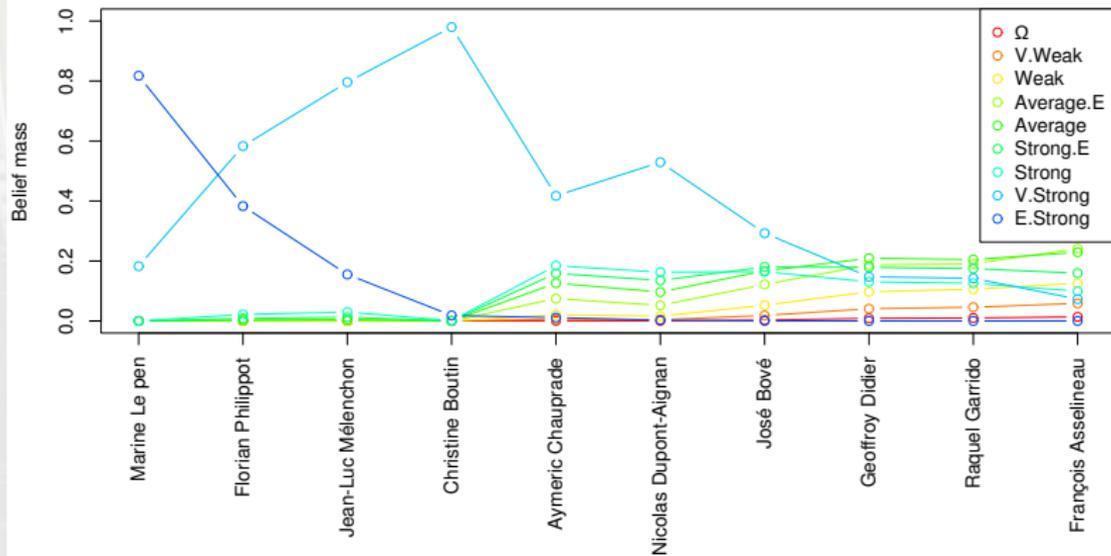
Rank	Candidats	Degré d'influence
1	Marine Le Pen	Extrêmement Forte
2	Florian Philippot	Très Forte
3	Jean-Luc Mélenchon	Très Forte
4	Christine Boutin	Très Forte
5	Aymeric Chauprade	Très Forte
6	Nicolas Dupont-Aignan	Très Forte
7	José Bové	Très Forte
8	Geoffroy Didier	Moyenne
9	Raquel Garrido	Moyenne
10	Marielle De Sarnez	Assez Moyenne

### 3 candidats les plus influentes selon notre méthode

Croyance en l'influence

	M. Le Pen	F. Philippot	J.L. Mélenchon
Très Faible	0	0,000011065	0,000030278
Faible	0	0.00007295998	0.0001832843
Assez Moyenne	0	0.0007035528	0.001403947
Moyenne	0	0.003033557	0.004954501
Assez Forte	0	0.008340205	0.01247841
Forte	0	0.02191526	0.02977818
Très Forte	0.1826552	0.5830090	0.7960571
Extrêmement Forte	0.8173448	0.3829144	0.1551143

# Candidats les plus influentes selon notre méthode



# Conclusion

- ◆ Méthode pour estimer l'influence basée sur la théorie des fonctions de croyance de Dempster-Shafer
  - △ Prise en compte des différentes relations
  - △ Prise en compte de l'incertitude
- ◆ Application au réseau Twitter

# Perspectives

- ◆ Choisir les bons paramètres initiaux (Machine learning?)
- ◆ Prise en compte de la dynamique
- ◆ Prise en compte des relations indirectes. Diffusion de l'influence

# Merci beaucoup

<https://github.com/kerzol/Influence-assessment-in-twitter>