Influence Assessment in Twitter Multi-Relational Network

Lobna Azaza, Sergey Kirgizov, Marinette Savonnet, Éric Leclercq
LE2I, University of Burgundy
Dijon, France
Email: {Lobna.Azaza, Sergey.Kirgizov}@u-bourgogne.fr
Email: {Marinette.Savonnet, Eric.Leclercq}@u-bourgogne.fr

Rim Faiz
LARODEC, IHEC, University of Carthage
Tunis, Tunisia
Email: Rim.Faiz@ihec.rnu.tn

***Abstract*—Influence in *Twitter* has become recently a hot research topic since this micro-blogging service is widely used to share and disseminate information. Some users are more able than others to influence and persuade peers. Thus, studying most influential users leads to reach a large-scale information diffusion area, something very useful in marketing or political campaigns. In this paper, we propose a new approach for influence assessment on *Twitter* network, it is based on a modified version of the conjunctive combination rule in belief functions theory in order to combine different influence markers such as *retweets*, *mentions* and *replies*. We experiment the proposed method on a large amount of data gathered from *Twitter* in the context of the European Elections 2014 and deduce top influential candidates.**

***Keywords*-Twitter network, Influence, Information fusion, Belief theory**

I. Introduction

Nowadays, online social networks such as *Twitter* gather people together and empower their relationships with new forms of cooperation and communication. As a result of its massive popularity, *Twitter* is exploited as a platform for marketing or political campaigns. One of the most distinctive characteristics of *Twitter* is the information diffusion through social links. In fact, links between users determine the information flow and thus indicate the user's influence on others. Some users, called influentials, are more able than others to diffuse information to a huge number of users. Therefore, determining influential users in a network is a secret key of success for achieving a large scale information diffusion at low cost.

Influence on *Twitter* is defined as the potential of a user's action to initiate a further action by another user [1]. The term "action" means the different possible interactions between users. Hence, measuring influence on *Twitter* is not that simple as the application provides several forms of interactions. A user can *follow* another one, which allows him to see *tweets* and information about the user he follows. He is also able to *retweet* a *tweet*, this exposes the *tweet* to his followers who can also *retweet* it. A user can *mention* another one by using the "@" prefix if he wants to address him the *tweet*. And finally, a user can *reply* to another's *tweet* and thus creates a conversation with him. These different relationships are what made *Twitter* a multi-relational network [2]. Therefore, possible actions on the *Twitter* network can be *retweet*, *mention*, *reply* or *follow*. These actions are also called influence markers. While measuring influence, the choice of these actions depends on understanding the subject and domain area [3].

Influence assessment poses two main challenges. The first is the diversity of influence markers on which we can rely to compute influence. It is important to combine them in order to establish a general influence measure that takes into account the different interaction types between users. The second is related to uncertainty when the combination of influence markers is performed. In the case of multi-relational networks, due to the multiple semantics of relationships, it is difficult to assign importance weights to the different influence markers before merging their related quantitative data.

Our contributions are manifold. In order to measure influence, we combine different influence markers obtained from relationships that exist in a multi-relational network. The measure can be established between a couple of users by taking into account different influence markers between them or it may also assess user's global influence in the network considering all the influence markers that he has on his peers. We also consider uncertainty in the measurement process. We define a theoritical framework, to compute influence, based on a modified version of conjunctive combination rule for belief functions theory and Smets rule [4] to fusion and combine information about different markers. A validation through experiments is proposed. It is based on real data gathered from Twitter in the TEE 2014 project during the European Election campaign in 2014.

The rest of the paper is organized as follows. Section II presents literature review. Section III describes our proposed approach. Section IV presents the experimental results. And finally Section V concludes the paper.

II. Literature review

In this section, we review studies of influence assessment in *Twitter* and remind the basic concepts of belief functions theory on which our approach is based.

A. Influence in Twitter

While measuring users influence in *Twitter*, many criteria can be considered. Number of a user's followers is widely used to assess influence [1], [5], [6]. Authors in [1] use three features to measure influence, which are: *replies, retweets* and *mentions* in addition to number of *followers*. They only give statistics related to these measures and do not offer a global influence score based on all the proposed markers. Cha *et al.* [5] use the criteria number of *followers*, *retweets* and *mentions*. They compute the value of each influence measure for 6 million users and compare them. In order to do this, they sort users according to each influence measure, after that, they quantify how a user's rank varies across different measures. Spearman's rank correlation is used as a measure of the association strength between two rank sets. They found that *followers* number represents a user's popularity, but is not related to other important influence markers such as *retweets* and *mentions*. Their result suggests that *followers* number alone reveals very little about a user's influence. This research does not also provide a global influence measure and only measures influence according to each marker separately.

In [7], authors define a measure based on topic similarity and structure in the links between users. Influence is considered as the fact of *following* other users regarding topic interests. In this context, the authors propose TwitterRank, an extension of the PageRank algorithm, in order to measure the topic-sensitive influence of the *twitterers*. Although the idea is promising, the experimental results show that there are some *twitterers follow* not because of the topic similarity between them and their friends, also the method ignored other important criteria such as *mentions* and *replies*.

In [6], the notion of social capitalism is proposed, it represents particular *Twitter* users trying to gain as many followers as possible in an artificial manner. The authors define a classifier that discriminates social capitalists from truthful users based on *Friend/Follower* ratio. The work presents some limits such as the use of number of *retweets* alone which cannot be considered as a good influence indicator. Also dataset used for experiments is quite small.

Romero *et al.* [8] propose the IP-Algorithm in order to measure influence, it is based on the HITS algorithm [9]. In this paper, influence is considered as the degree of content propagation in the network (*retweets*). In addition, authors believe that a user's influence depends not only on the size of the influenced audience, but also on their passivity. The passivity of a user is his passive information consuming without forwarding the content to the network. The algorithm showed better accuracy than other influence measures such as PageRank [10], the number of *followers* and number of *mentions*. Although passivity seems a good influence indicator, this work ignored other important influence marker such as *reply*.

In [11], authors propose a combination of two models for ranking users' influence: The PageRank algorithm [10] and HMM (Hidden Markov Model). They build a HMM to observe the influence evolution over time and use three observables: *retweet*, *mention* and *reply*. The model is evaluated using survey as ground-truth for influence ranking. The proposed model differs from the others by combining the important influence markers. However, as the purpose is to rank users' influence, a user's given influence does not reveal information about its influence degree (high or low influence), the model's output is only useful in users ranking.

Existing research proposes methods to measure influence. However, none of them presents an approach for *Twitter* influence regarding multi-criteria combination, also, uncertainty has not been considered yet in such combinations. It is important to assess influence taking into account the degrees of uncertainty about the weights assigned to different criteria according to their importance. In this purpose, we propose the use of belief functions theory.

B. Belief functions theory

Every day, a huge volume of incomplete and imperfect information is spread through the different links of social networks. Thus, reasoning with uncertainty has become a major interest in social networks [12].

The belief functions theory is considered as a general framework for reasoning with uncertainty, and has well been connected to other frameworks such as probability, possibility and imprecise probability theories [13]. The theory of belief functions, also known as evidence theory or Dempster-Shafer theory, was first introduced by A. Dempster in the context of statistical inference, and was later developed by G. Shafer as a general framework for modeling epistemic uncertainty [14].

In the following, we are going to remind the basic concepts of belief functions theory. Let Ω be a finite set, denote by 2^Ω the set of all subsets of Ω. In the context of Dempster-Shafer theory, Ω is often called a frame of discernment. A mass m is a function $m : 2^\Omega \longrightarrow [0, 1]$ such that:

$$\sum_{X \in 2^\Omega} m(X) = 1 \text{ and } m(\emptyset) = 0 \tag{1}$$

The mass $m(X)$ expresses the part of belief that supports the subset X of Ω.

Belief functions theory allows, not only the representation of the partial knowledge, but also the information fusion [15]. This is done by the conjunctive combination rule [4], it assumes that all sources are reliable and consistent. Considering two mass functions m_1 and m_2, the conjunctive combination rule is defined as:

$$(m_1 \cap\!\!\!\!\bigcirc\, m_2)(C) = \sum_{A \cap B = C} m_1(A) m_2(B), \quad A, B, C \in 2^\Omega \tag{2}$$

In order to make a decision, we try to select the most likely hypothesis which may be difficult to realize directly with the basics of the belief functions theory where mass functions are given not only to singletons but also to subsets of hypothesis. There exist several solutions to ensure decision making within belief functions theory. The most known is the pignistic probability [16]. In contrast to mass functions that are defined on 2^Ω, pignistic probability is a probability measure defined on Ω. Pignistic probability was proposed in the Transferable Belief Model (TBM) [17]. It is based on two levels: The "credal level" where beliefs are entertained and represented by belief functions and the "pignistic level" where beliefs are used to make decisions and represented as probability functions called pignistic probabilities denoted *bet*.

$$\text{bet}(x) = \sum_{x \in X \subseteq \Omega} \frac{m(X)}{|X|} \tag{3}$$

Belief functions theory has been widely used in many fields such as natural risks [18]. To the best of our knowledge, this is the first time belief functions are exploited in influence assessment.

III. Proposed approach

In order to assess users' influence, we propose a belief approach based on information fusion about the different possible influence relationships. Figure 1 gives an overview of the framework for the proposed approach. First, in a credal level, we associate belief masses for each relationship. Then, we combine them to obtain the influence belief mass. And finally, at the pignistic level, we compute the pignistic probability in order to make a decision about the user's influence degree. In the following section, we detail each step of the assessment process.

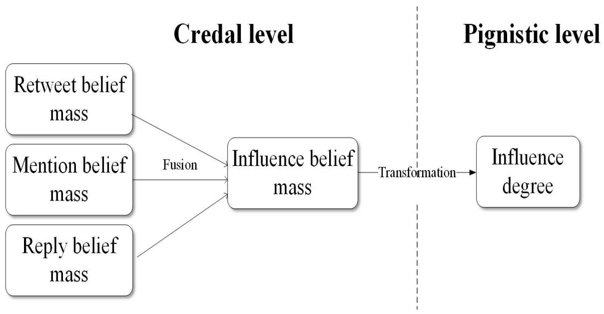

Figure 1. Framework of the proposed approach

A. Belief diffusion network

Social networks have been widely modeled as a graph [19]. A graph is usually represented as $G = (V, E)$ comprising a set V of vertices or nodes together with a set E of edges or links. To model the different relationships in a graph, the concept of multi-relational graph is used (sometimes called multi-dimensional graph or multi-layered graph). In a multi-relational graph the set of links E is divided into pairwise disjoint classes $E = \bigcup_{r \in R} E_r$, where R is the set of possible types of relationships. For example, in *Twitter* we can consider

$$R = \{\text{Retweet}, \text{Mention}, \text{Reply}\} \tag{4}$$

Recent researches have introduced uncertain graphs whose edges are labeled with a probability of existence [20], [21]. But uncertainty about the semantics of links and nodes is not introduced in this kind of graphs.

In our approach, uncertainty is injected intentionally for evaluating influence. In this context, we introduce the belief diffusion network (Figure 2) where users are represented by nodes, and links model the different relationships between them. The belief diffusion network is a labeled multi-relational graph. The links are labeled with influence degrees (e.g., Weak, Average, Strong) and belief masses that depend on the type of the relationship. Nodes are also labeled with uncertainty about their estimated influence degree resulting from the fusion of the belief masses of incident links.

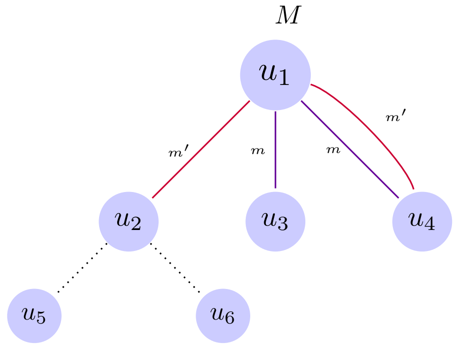

Figure 2. Belief diffusion network. Nodes represent users, and links are different relationships between them. Relationship types are represented by colors. Two different masses m and m' are associated with each relationship type. Using a combination procedure, discussed below in Subsection III-B, we obtain a global influence (denoted by M) of node u_1.

B. Masses fusion on belief diffusion network

We present here a modified version of Dempster-Shafer theory discussed above in the Section II. These modifications will allow us to fusion different mass functions defined on the multi-relational network.

Let Ω be an ordered set of possible influence degrees

$$\begin{aligned}\Omega = \{&\text{Very Weak}, \text{Weak}, \text{Average Enough},\\ &\text{Average}, \text{Strong Enough}, \text{Strong},\\ &\text{Very Strong}, \text{Extremely Strong}\}\end{aligned} \tag{5}$$

In general Dempster-Shafer theory we should use 2^Ω as a domain of mass functions. But in this paper we use only a certain subset Λ of 2^Ω, precisely:

$$\begin{aligned}\Lambda = \{&\text{Very Weak}, \text{Weak}, \text{Average Enough},\\ &\text{Average}, \text{Strong Enough}, \text{Strong},\\ &\text{Very Strong}, \text{Extremely Strong}, \Omega\}\end{aligned} \tag{6}$$

And so, mass functions are defined as follows:

$$m : \Lambda \to [0, 1] \tag{7}$$

For each relationship type, a mass function is associated. In order to estimate the influence degree of a specific node u, we take into account the local structure of the belief diffusion network around the node u and combine the belief mass functions of incident links using a modified version of conjunctive combination rule (2).

$$(m \otimes m')(z) = \sum_{y @ x = z} m(x)m'(y), \quad x, y, z \in \Lambda \tag{8}$$

Table I
DEFINITION OF THE OPERATION @

@	**V.Weak**	**Weak**	**Average.E**	**Average**	**Strong.E**	**Strong**	**V.Strong**	**E.Strong**	**Ω**
V.Weak	Weak	Average.E	Average	Strong.E	Strong	V.Strong	V.Strong	E.Strong	V.Weak
Weak	Average.E	Average.E	Average	Strong.E	Strong	V.Strong	V.Strong	E.Strong	Weak
Average.E	Average	Average	Strong.E	Strong	V.Strong	V.Strong	V.Strong	E.Strong	Average.E
Average	Strong.E	Strong.E	Strong	Strong	V.Strong	V.Strong	V.Strong	E.Strong	Average
Strong.E	Strong	Strong	V.Strong	V.Strong	V.Strong	V.Strong	V.Strong	E.Strong	Strong.E
Strong	V.Strong	V.Strong	V.Strong	V.Strong	V.Strong	E.Strong	E.Strong	E.Strong	Strong
V.Strong	V.Strong	V.Strong	V.Strong	V.Strong	V.Strong	E.Strong	E.Strong	E.Strong	V.Strong
E.Strong	E.Strong	E.Strong	E.Strong	E.Strong	E.Strong	E.Strong	E.Strong	E.Strong	E.Strong
Ω	V.Weak	Weak	Average.E	Average	Strong.E	Strong	V.Strong	E.Strong	Ω

where @ is a symmetric operation $@ : \Lambda \times \Lambda \to \Lambda$. Table I represents an example of an @ operation.

Proposition 1. *A combination of any two mass functions is another mass function.*

Proof. Denote $(m \otimes m')$ by m''. It is easy to see that for all x we have $m''(x) \geqslant 0$, because we compute m'' using only multiplication and addition of non-negative numbers.

Next, we show that $\sum_{z \in \Lambda} m''(z) = 1$.

Let $\Lambda_z^2 = \{(x, y) \in \Lambda : x@y = z\}$ and proceed as follows:

$$\begin{aligned}\sum_z m''(z) &= \sum_z \sum_{x@y=z} m(x)m'(y) \\ &= \sum_z \sum_{(x,y)\in\Lambda_z^2} m(x)m'(y)\,.\end{aligned}$$

Note that $\Lambda_z^2 \neq \Lambda_{z'}^2 \iff z \neq z'$, and $\bigcup_{z\in\Lambda} \Lambda_z^2 = \Lambda^2$. So, we can ommit $\sum_z$ and rewrite as follows:

$$\begin{aligned}&= \sum_{(x,y)\in\Lambda^2} m(x)m'(y) \\ &= \sum_x \sum_y m(x)m'(y) \\ &= \sum_x m(x) \sum_y m'(y)\end{aligned}$$

m and m' are mass function: $\sum_x m(x) = \sum_y m'(y) = 1$, so

$$\sum_x m(x) \sum_y m'(y) = 1$$

$\square$.

Proposition 2. *In general $\otimes$ is non-associative:*

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

Proof. Consider $\Omega = \{A, B, C\}$, and the following @:

@	***A***	***B***	***C***	Ω
A	B	B	C	A
B	B	C	C	B
C	C	C	C	C
Ω	A	B	C	Ω

$$m = m' = \begin{array}{|c|c|c|c}A & B & C & \Omega \\ \hline 1 & 0 & 0 & 0\end{array}$$

$$m'' = \begin{array}{|c|c|c|c}A & B & C & \Omega \\ \hline 0 & 1 & 0 & 0\end{array}$$

It's easy to see that:

$$(m \otimes m') \otimes m'' = \begin{array}{|c|c|c|c}A & B & C & \Omega \\ \hline 0 & 0 & 1 & 0\end{array}$$

$$m \otimes (m' \otimes m'') = \begin{array}{|c|c|c|c}A & B & C & \Omega \\ \hline 0 & 1 & 0 & 0\end{array}$$

Thus, in general:

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

$\square$.

Now, we consider multiple relationships existing between node u and its neighbors. We associate a mass function m_r to any relation $r \in R$. We denote by I_r the set of all links with relationship type r. Finally, we have the following set of mass functions $\{m_{r,i} : r \in R, i \in I_r\}$. Based on Prop. 1 we can combine these mass functions in order to obtain a global belief mass corresponding to the influence degree of node u. But the order of combinations may affect our results (Prop. 2). To be consistent in our measurements we have to choose the order. In order to simplify the expressions we will write $\bigotimes_{i\in\{1,2,3\}}$ instead of $m_1 \otimes m_2 \otimes m_3$. Thus, we consider the following order of combinations:

1) For a given relationship type r we subsequently combine the masses of the relashionships of type r in order to get r-preresult with $\hat{m}_r$ defined as follows:
$$\hat{m}_r = \bigotimes_{i\in I_r} m_{r,i}$$
2) Then we combine all r-preresults using:
$$\bigotimes_{r\in R} \hat{m}_r \tag{9}$$

Depending on the operation @ such procedure may finally converge to certain stationary mass.

Once we have the global belief mass on a certain node, we use a modified version of the pignistic probability defined on equation 3 in order to make the decision about the influence degree of a user. In our case the belief masses is defined on Λ = {Very Weak, Weak, Average Enough, ..., Ω}, and the pignistic probability is calculated by distributing uniformly the mass of Ω to all other elements of Λ:

$$\text{bet}(x) = m(x) + \frac{m(\Omega)}{|\Omega|}, \quad x \in \Omega \tag{10}$$

C. Illustrations

In order to illustrate our method we consider the following mass functions associated to the relationships:

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

$$\text{Mention} \mapsto \begin{cases} m_{\text{mention}}(\text{V.Weak}) = 0.3 \\ m_{\text{mention}}(\Omega) = 0.7 \end{cases}$$

The belief masses $m_{\text{retweet}}(\Omega)$ and $m_{\text{mention}}(\Omega)$ represent the partial ignorance.

Case 1: Two retweets

After initialisation of belief masses on the different relationships, we follow the proposed approach process to measure the influence resulted from combination of two *retweets* from one user to another. We first use the operation @ giving the correspondances between the influence masses, then we calculate the conjunctive combination. The combined mass function of the two *retweets* are shown in table II:

Table II
COMBINATION OF TWO *retweets*

$\otimes$	**Weak** 0.4	**Ω** 0.6
Weak 0.4	Average.E 0.16	Weak 0.24
Ω 0.6	Weak 0.24	Ω 0.36

We obtain then:
$m(\text{Weak}) = 0.24 + 0.24 = 0.48$
$m(\text{Average.E}) = 0.16$
$m(\Omega) = 0.36$

Finally, to make a decision on the influence degree, we calculate the pignistic probability using equation 10 (Table III). For example, for the degree Weak, we proceed as follows to obtain the pignistic probability:

$$\text{bet}(\text{Weak}) = m(\text{Weak}) + \frac{m(\Omega)}{|\Omega|} = 0.48 + \frac{0.36}{8} = 0.525$$

Table III
PIGNISTIC PROBABILITY FOR CASE 1

V.Weak	0.045
Weak	0.525
Average.E	0.205
Average	0.045
Strong.E	0.045
Strong	0.045
V.Strong	0.045
E.Strong	0.045

We conclude that the influence degree is Weak since it has the highest pignistic probability 0.525. This latter was 0.4 before considering the combination.

Case 2: 2 retweets + 2 mentions

In the second case, we consider two additional *mentions* existing between the same users of case 1. In order to measure influence, we use our proposed process to combine masses of the two *mentions* then we combine the obtained massess with the results of the previous case related to two *retweets* combination.

The conjunctive combination on the two *mentions* gives:

Table IV
COMBINATION OF TWO *mentions*

$\otimes$	**V.Weak** 0.3	**Ω** 0.7
V.Weak 0.3	Weak 0.09	V.Weak 0.21
Ω 0.7	V.Weak 0.21	Ω 0.49

We obtain:
$m(\text{V.Weak}) = 0.42$
$m(\text{Weak}) = 0.09$
$m(\Omega) = 0.49$

Now, we combine the obtained masses with the results of the case 1:

Table V
CASE 2: 2 *retweets* + 2 *mentions*

$\otimes$	**Weak** 0.48	**Average.E** 0.16	**Ω** 0.36
V.Weak 0.42	Average.E 0.2016	Average 0.0672	V.Weak 0.1512
Weak 0.09	Average.E 0.0432	Average 0.0144	Weak 0.0324
Ω 0.49	Weak 0.2352	Average.E 0.0784	Ω_{Inf} 0.1764

We obtain:
$m(\text{V.Weak}) = 0.1512$
$m(\text{Weak}) = 0.2676$
$m(\text{Average.E}) = 0.3232$
$m(\text{Average}) = 0.0816$
$m(\Omega) = 0.1764$

We note that, by combining the four relationships, the belief mass on the degree Weak has decreased compared to the first case, this is due to the fact that the mass of the Average.E degree has increased and became equal to 0.3232. We also notice that the degree Average appeared with a mass equal to 0.0816. We can conclude that the more we have relationships and the more we combine them, the highest influence we get.

Now to make the decision on the influence degree, we compute the pignistic probability:

Table VI
PIGNISTIC PROBABILITY FOR CASE 2

V.Weak	0.17325
Weak	0.0.28965
Average.E	0.34525
Average	0.10365
Strong.E	0.02205
Strong	0.02205
V.Strong	0.02205
E.Strong	0.02205

We conclude that the influence degree is Average.E with a pignistic probability of 0.34525. This latter was 0.205 before considering the two *mentions*.

IV. EXPERIMENTS AND RESULTS

A. Data description

The research work takes place in the project TEE 2014 whose exact title is: "*Twitter* in the European Elections: An international contrastive study of *Tweets* use by candidates in elections to the European Parliament in May 2014". This international project led by the House of Human Sciences (MSH) in Dijon, brings together nearly 45 researchers (political scientists, sociologists, communication researchers and computer scientists), 10 research laboratories spread across 6 European countries (France, Germany, Belgium, Italy, Spain and the UK). The overall objective of this project is to observe and analyze the *Tweets* communication policies during the election period in May 2014 in various countries of Western Europe.

The *tweets* collection during the election period has build a corpus which is then analyzed. To collect information from *Twitter*, we used our developed tool *SNFreezer*[1]. Three types of information (generalized under the term "source") are taken as a parameter in this collection: user accounts; *hashtags* and words or phrases. The purpose of gathering is to retrieve *tweets mentioning* the designated users, those containing a *hashtag*, a word or phrase, or *tweets* sent by candidates. In addition, we collect information on these *tweets* such as *retweeted tweets*, users *mentioned* in *tweets* and *replies* to *tweets*. These sources were chosen by political scientists, and among them we find the names of the leading candidates, their *Twitter* accounts and their parties. The collection allowed us to have a large number of tweets (37 million) retrived for 50 consecutive days, and to massively process these data.

[1]https://github.com/SNFreezer

B. Experiments and results

Our experimental goal is to measure candidates influence on the network. Unlike illustrations given in the previous section, we do not consider the case of measuring influence between two users but rather global candidates' influence in the network. In order to study their influence, we affect masses to the considered relationships, then we take each candidate's number of *retweets*, *mentions* and *replies* and combine their masses.

The choice and the affectation of the masses in the initialisation step is an important issue while dealing with real data. In some domains such as politics, users have very high number of relationships. When we used the masses initialized in the illustration's section on our data, influence converges to the highest influence E.Strong after a short number of iterations ($\simeq$ 45 iterations). In this way, we can not compare candidates' influence as we have the same influence with similar masses for most of them. To deal with this, we perform a rescaling and use the following masses afectation:

$$\textit{Retweet} \mapsto \begin{cases} m_{\text{retweet}}(\text{V.Weak}) = 0.55 \cdot 10^{-3} \\ m_{\text{retweet}}(\Omega) = 1 - 0.55 \cdot 10^{-3} \end{cases}$$

$$\textit{Mention} \mapsto \begin{cases} m_{\text{mention}}(\text{V.Weak}) = 0.45 \cdot 10^{-3} \\ m_{\text{mention}}(\Omega) = 1 - 0.45 \cdot 10^{-3} \end{cases}$$

$$\textit{Reply} \mapsto \begin{cases} m_{\text{reply}}(\text{V.Weak}) = 0.45 \cdot 10^{-3} \\ m_{\text{reply}}(\Omega) = 1 - 0.45 \cdot 10^{-3} \end{cases}$$

Tables VII and VIII show the combination results for the French candidates "Marine Le Pen", "Florian Philippot" and "Jean-Luc Mélenchon" and the English candidates "Katie Hopkins", "Nigel Farage" and "Patrick O'Flynn". For example, we conclude that the influence degree on the network for the candidate "Marine Le Pen" who has: 14678 *retweets*, 66798 *mentions* and 4003 *replies*, is E.Strong with the belief mass 0.8173448. Results given do not only provide the influence degree but also give indication of our belief in the given results which is performed by the belief masses on the different degrees. The source code is available on github[2].

C. Towards ranking users influence

The proposed approach of influence assessment can also be exploited to rank users' influence. Table IX presents the top influential candidates according to the relationships taken individually. They are ranked by their numbers of *retweets*, *mentions* and *replies*. The presented results do not provide the global candidates influence in the network since different rankings for each relationship are used. Users are also ranked according to their centrality degree. It is computed using the number of the candidates' neighbors in the multi-relational network. This enables to have a global ranking for the candidates but do not offer any indication on the influence degree of each candidate.

[2]https://github.com/kerzol/Influence-assessment-in-twitter

Table IX
TOP INFLUENTIAL FRENCH CANDIDATES ACCORDING TO DIFFERENT RELATIONSHIPS AND CENTRALITY DEGREE

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

Table VII
RESULTS FOR THE TOP 3 INFLUENTIAL FRENCH CANDIDATES

	M. Le Pen	**F. Philippot**	**J.L. Mélenchon**
E.Weak	0	0.000011065	0.000030278
Weak	0	0.00007295998	0.0001832843
Average E	0	0.0007035528	0.001403947
Average	0	0.003033557	0.004954501
Strong.E	0	0.008340205	0.01247841
Strong	0	0.02191526	0.02977818
V.Strong	0.1826552	0.5830090	0.7960571
E.Strong	0.8173448	0.3829144	0.1551143
Ω	0	0	0

Table VIII
RESULTS FOR THE TOP 3 INFLUENTIAL ENGLISH CANDIDATES

	Katie Hopkins	**Nigel Farage**	**Patrick O'Flynn**
E.Weak	0	0	0.0022350807
Weak	0	0	0.0067817003
Average E	0	0	0.0292114368
Average	0	0	0.0707907792
Strong.E	0	0	0.1140117014
Strong	0	0	0.1635190125
V.Strong	0.0260529	0.2663876	0.5804666241
E.Strong	0.9739471	0.7336124	0.0325333152
Ω	0	0	0.0004503499

In this section, our experimental goal is to detect most influential candidates on the network based on our proposed approach. We focus on the French candidates in the elections, we have 616 candidates with 4 million *tweets*.

In order to rank users we proceed as follows:

1) For each candidate we take the influence with maximal belief mass (for example, Marine Le Pen $\mapsto$ E.Strong).
2) We rank candidates by their "maximal influence degree".
3) When two candidates have the same "maximal influence degree":

$$\text{Florian Philippot} \mapsto \text{V.Strong}$$

$$\text{Jean-Luc Mélenchon} \mapsto \text{V.Strong}$$

we compare belief masses of next-greater influence degree[3]:

$$m_{\text{Philippot}}(\text{E.Strong}) > m_{\text{Mélenchon}}(\text{E.Strong})$$

We proceed this way since it is unfair to rank candidates by maximal belief masses they have on the degrees. This is because we may have a user more influential than another although he has a weaker belief mass than him on the same degree. This is due to the fact that, the belief mass on the next-greater degree has increased and became quite important. For example, the candidate Florian Philippot has a belief mass on the degree V.Strong weaker than the belief mass of Jean-Luc Mélenchon on the same degree as we can see in table VII. Inspite of this, he is ranked before Jean-Luc Mélenchon (table X) as he has a greater belief mass on the degree E.Strong.

We do the combination procedure for all candidates and deduce their ranking by influence degree. Findings are shown in table X. The results are general, taking into account the possible influence markers in one same measure unlike results shown in table IX.

Table X
TOP INFLUENTIAL CANDIDATES ACCORDING TO BELIEF FUSION

Rank	**Candidates**	**Influence degree**	**Belief mass**
1	Marine Le Pen	E.Strong	0.8173448
2	Florian Philippot	V.Strong	0.5830090
3	Jean-Luc Mélenchon	V.Strong	0.7960571
4	Christine Boutin	V.Strong	0.9796956
5	Aymeric Chauprade	V.Strong	0.4171324655
6	Nicolas Dupont-Aignan	V.Strong	0.5293170700
7	José Bové	V.Strong	0.2925722297
8	Geoffroy Didier	Average	0.2092645352
9	Raquel Garrido	Average	0.2048485
10	Marielle De Sarnez	Average.E	0.2074260

V. CONCLUSION

In this paper, we proposed an influence assessment approach for the *Twitter* social network. This approach addresses limitations of existing systems such as lack of

[3]The next-greater influence degree for V.Strong is E.Strong.

markers combination and uncertainty ignorance on the given measures. In our work, we proposed a diffusion belief network allowing us to observe different relationships in the network, we considered three influence markers: *retweet*, *mention* and *reply*. Based on belief functions theory, we established a general influence measure for a given user by information fusion of the different markers. We experimented our approach on real data gathered from *Twitter* in the context of the project TEE'14. The experiments show that markers combination under uncertainty leads to a quite interesting results.

Interesting perspectives emerge to further strengthen the proposed approach. Indirected influence in our belief network will be taken into consideration in future works, the influence measure should consider influence exercised on indirected nodes (e.g. a user may *retweet* another's *tweet* indirectly through an intermediate user). This will span the influence on multiple levels based on a multi-level diffusion network. Besides, the method for users ranking will be improved and we will compare the results with those obtained with well known algorithms in the literature such as TwitterRank and HITS algorithms. And finally, we hope to consider more interesting relationships in the method such as *hashtags* or *favorites*.

ACKNOWLEDGEMENTS

Authors are grateful to Nicolas Gastineau and Armen Petrossian for useful discussions about possible formalisations of influence-belief functions and to Elisabeth Gavignet for helpful comments and proof-readings.

REFERENCES

[1] A. Leavitt, E. Burchard, D. Fisher, and S. Gilbert, "The Influentials: New Approaches for Analyzing Influence on Twitter," *Webecology Project*, September 2009.

[2] Z. Wu, W. Yin, J. Cao, G. Xu, and A. Cuzzocrea, "Community detection in multi-relational social networks," in *Web Information Systems Engineering WISE 2013*, 2013, pp. 43–56.

[3] J. Sun and J. Tang, "A Survey of Models and Algorithms for Social Influence Analysis," in *Social Network Data Analytics*, C. C. Aggarwal, Ed. Boston, MA, USA: Springer US, 2011, ch. 7, pp. 177–214.

[4] P. Smets, "Imperfect Information: Imprecision and Uncertainty," in *Uncertainty Management in Information Systems*, A. Motro and P. Smets, Eds., 1997, pp. 225–254.

[5] M. Cha, H. Haddadi, F. Benevenuto, and K. Gummadi, "Measuring user influence in twitter: The million follower fallacy," in *4th International AAAI Conference on Weblogs and Social Media (ICWSM)*, 2010.

[6] M. Danisch, N. Dugu, and A. Perez, "On the importance of considering social capitalism when measuring influence on Twitter," in *Behavioral, Economic, and Socio-Cultural Computing*, Shanghai, China, Oct. 2014.

[7] J. Weng, E.-P. Lim, J. Jiang, and Q. He, "TwitterRank: Finding Topic-sensitive Influential Twitterers," in *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, ser. WSDM '10. New York, NY, USA: ACM, 2010, pp. 261–270.

[8] D. M. Romero, W. Galuba, S. Asur, and B. A. Huberman, "Influence and passivity in social media," in *Proceedings of the 20th International Conference Companion on World Wide Web*, 2011, pp. 113–114.

[9] J. M. Kleinberg, "Authoritative sources in a hyperlinked environment," *J. ACM*, pp. 604–632, Sep. 1999.

[10] L. Page, S. Brin, R. Motwani, and T. Winograd, "The pagerank citation ranking: Bringing order to the web," in *Proceedings of the 7th International World Wide Web Conference*, 1999, pp. 161–172.

[11] D. Simmie, M. Vigliotti, and C. Hankin, "Ranking twitter influence by combining network centrality and influence observables in an evolutionary model," in *Signal-Image Technology Internet-Based Systems (SITIS), 2013 International Conference on*, 2013, pp. 491–498.

[12] A. P. Dempster, R. Yager, and L. Liu, *Classic Works on the Dempster-Shafer Theory of Belief Functions (Studies in Fuzziness and Soft Computing)*, 2007.

[13] T. Denoeux and M.-H. E. Masson, "Belief Functions: Theory and Applications," in *Proceedings of the 2nd International Conference on Belief Functions, 9-11 May 2012*, ser. Advances in Intelligent and Soft Computing, 2012, p. 444.

[14] S. Kotz and W. N. L. Johnson eds., "Belief functions," *Encyclopedia of Statistical Sciences 1 209*, 1982.

[15] V. Nimier and A. Appriou, "Utilisation de la théorie de dempster-shafer pour la fusion d'informations," *GRETSI, Groupe dEtudes du Traitement du Signal et des Images*, pp. 137–140, 1995.

[16] P. Smets, "Constructing the pignistic probability function in a context of uncertainty." in *UAI*, vol. 89, 1989, pp. 29–40.

[17] P. Smets and R. Kennes, "The transferable belief model," in *Classic Works of the Dempster-Shafer Theory of Belief Functions*, ser. Studies in Fuzziness and Soft Computing, 2008, pp. 693–736.

[18] J. Tacnet, M. Batton-Hubert, and J. Dezert, "Analyse multicritères et fusion d'information pour l'expertise et la gestion intégrée des risques naturels en montagne," in *Colloque LambdaMu 17*, 2010, p. 10 p.

[19] J. A. Barnes, "Graph Theory and Social Networks: A Technical Comment on Connectedness and Connectivity," *Sociology*, pp. 215–232, 1969.

[20] A. Khan, F. Bonchi, A. Gionis, and F. Gullo, "Fast Reliability Search in Uncertain Graphs," in *Proceedings of Proceedings of the International Conference on Extending Database Technology (EDBT 14)*, 2014.

[21] P. Parchas, F. Gullo, D. Papadias, and F. Bonchi, "The Pursuit of a Good Possible World: Extracting Representative Instances of Uncertain Graphs," in *Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data*, 2014, pp. 967–978.