Detecting Evangelists and Detractors on Twitter

Carolina Bigonha, Thiago N. C. Cardoso,
Mirella M. Moro, Virgílio A. F. Almeida, Marcos A. Gonçalves
DCC - UFMG
Belo Horizonte, MG, Brazil
{carolb,thiagon,mirella,virgilio,mgoncalv}@dcc.ufmg.br

ABSTRACT

Social networking websites provide a suitable environment for interaction and topic discussion. With the growing popularity of online communities, estimulated by the easiness with which content can be created and consumed, some of this content became strategical for companies interested in obtaining population feedback for products, personalities, etc. One of the most important of such websites is Twitter: recent statistics report 50 million of new tweets each day. However, processing this amount of data is very costly and a big part of it is simply not useful for strategic analysis. In this paper, we propose a new technique for ranking the most influential users in Twitter based on a combination of the user position in the network topology, the polarity of her opinions and the textual quality of her tweets. In addition, we develop and compare two methods for calculating the network influence. We also performed experiments with a real dataset containing one month of posts regarding soda brands. Our experimental evaluation shows that our approach can successfully identify some of the most influential users and that interactions between users are the best evidence to determine user influence.

Categories and Subject Descriptors

H.3 [**Information Storage and Retrieval**]: Systems and Software

General Terms

Influential users on Twitter

Keywords

Twitter, user influence, network topology

1. INTRODUCTION

Online Social Networks became a rich source of information, mainly due to *user-generated content*, which has been facilitated and stimulated by several supporting mechanisms in Web 2.0 applications. In these environments, users share their opinions, experiences, feelings and suggestions on various matters such as products, personalities and events.With time, the extraction and analysis of this kind of data became of great value for advertising companies, manufacturers, government, etc, in any situation in which consumer/ population feedback is desirable [6].

Among social networking websites, one of the most important is Twitter: a real-time micro-blogging network, with a huge volume of content generated daily – recent statistics report 50 million of new tweets each day [14]. Twitter presents a simplified approach for posting on the web, where each tweet has at most 140 characters, defining an easy way to produce and consume content.

Businesses want to see how well their campaigns, products and brands are being received, and understand key influencers in their areas of interest. Considering the aforementioned growth rate, it is impractical (for a social network analyst monitoring Twitter) to inspect all the data generated by every single user, even for a specific topic. The methods presented in this paper allows one to look at Twitter users and identify evangelists and detractors in a marketing campaign. Furthermore, our methods allow corporations, marketers, public relations agencies and advertisers to know who influences the target audience the most.

In this paper, we define the influence of a user based on her network position and her behavior – the interaction with other users, the polarity of her opinions and the quality of her tweets. We present a technique to automatically identify influential users on Twitter. More specifically, given a certain topic, we identify *evangelists* and *detractors*, i.e. the influential users who act in favor and against a subject, respectively. For strategic planning, the goal is to focus on a subset of Twitter content that represents the viral marketing niche for a product, personality or other subject of interest. Once a specific topic is defined, we can classify opinions as positive and negative and, consequently, divide the influential users into evangelists and detractors.

For testing our techniques, we collected information about users that posted tweets regarding a specific topic (soda brands) during one month. Each tweet and user data were manually analyzed by marketing professionals as *positive / negative / neutral* and *evangelist / detractor / irrelevant*. Our experimental results demonstrate that we can successfully identify some of the most influential users using our techniques and that interactions between users are the best evidence to determine user influence. Although the

.

experiments were performed for a specific topic, the proposed technique is applicable to any subject. Moreover, the topic-specific dataset employed has similar characteristics of some more general Twitter-based collections used in previous work, such as [7], [9], meaning that most of our results are potentially generalizable.

In summary, the main contributions of this paper are: (*i*) a technique that, given a set of keywords covering/defining a topic, provides a list of its evangelists and detractors (Section 3); (*ii*) an analysis of which metrics contribute more for the detection of these influential users (Section 4); and, (*iii*) an experimental validation and evaluation of the proposed techniques (Section 5).

2. RELATED WORK

Recently, the characterization of users in Twitter (and other similar platforms [13]) has been focus of much research. Specifically, in [8] and [9], a large effort was made for characterizing Twitter social network, its users' behavior and interactions. The concept of friend of an user as being a person that received at least two public mentions from the user was defined in [7]. The essence of such work is to show that the follower and following relation is not as meaningful as the interaction.

The study performed in [10] define "influence in Twitter" as being the potential of a user action to generate other actions (*reply*, *retweet*, *mention* or an *attribution*). The authors also define different types of users based on the follower/followee ratio and identify the relation between the posted tweets and the actions that they generated. The main objective of that work was to use the interactions in Twitter to identify influential users. Beyond the fact that the authors did not consider the tweet content, the interaction data they extract is purely quantitative. In this work, we present a graph of interactions and analyze the network topology.

A method for topic-sensitive influential users detection is defined in [15]. This method considers only a pagerank-like metric in the follower-following network: it calculates the user influence based on how many people have received her tweets. As shown in previous works [7] [3], and reinforced in this paper, this metric is *not* determinant for finding influential users, since not all the users that receive a tweet are interested on its content.

In [3], influence is divided in three types: the in-degree influence (the number of followers that an user have); the re-tweet influence (the number of re-tweets containing ones name); and mention influence (the number of times an user mentioned). In our method, re-tweet and mention influences are covered as the in-degree of the interaction network. Beyond that, we apply other topology features, such as betweenness and eigenvector centrality to determine the user influence.

3. OVERVIEW OF THE TECHNIQUE

In a Twitter profile page, it is possible to obtain several pieces of personal information about the user, such as picture, real name, location, homepage and a short biography. Some quantitative data can also be obtained, including the total number of tweets posted by the user, her number of followees, followers, favorite tweets and so on. What the user is up to or with whom he interacts are examples of information that cannot be extracted based only on this raw data. According to [7], users interact with a very small number of other users compared to the number of people they follow. In this paper, the proposed technique for identifying influential users relies mainly on the behavior of users rather than on its following-follower connections, even though the presented experiments covers both approaches.

An overview of our technique is given on Figure 1. The first step (a) is the definition of the topic and time interval of interest. The goal is to monitor and identify the influential users who talk about an event, a company, a brand, a celebrity or any subject that can be synthesized into keywords.

Once these parameters were defined, tweets that fit into the specifications are collected (b) using the Twitter Streaming API [1], that allows near-realtime access to Twitter public statuses (tweets). Along with each tweet, we store profile information about its author, such as her *username*, number of *following* and number of *followers*.

Users tend to have biased opinions on certain matters. That considered, it is interesting to specify the influential users who are mainly in favor or against the selected subject. To achieve that, each item of the repository of collected tweets was analyzed by specialists (c). This analysis is performed as follows: (1) Organization of the tweets in subcategories of the topic chosen (for a presidential election, for example, the subcategories would be the candidates); (2) Classification of the polarity of each tweet according to the subcategory (positive, negative, neutral); (3) Identification of the users as evangelists, detractors or irrelevant. Any tweet that does not fit the query is eliminated from the collection. This stage is currently manually performed and future work includes the automation of this step.

As soon as the manual classification is over, two main procedures are executed: (d) fetching of the following-follower connections of the authors of each tweet; and (e) extraction of user interactions. In step (d), the goal is to construct a network where the nodes are users and the arcs represent a following-follower relation. It is important to remind that the *in* and *out-degree* of each user – the connections between the users within the collected dataset – is different from the number of following and follower users that appear on her profile. That means that each author has two values of following and follower connections: one corresponding to its connections on the whole Twitter and the other that corresponds to its relations within the network of topic-related authors. The stage represented by step (e) concerns the extraction of user interactions via tweets. It is very common for a user to interact with others in a post by using the '@' notation prefacing their username. The most common interactions are *in reply* and *re-tweet*. The former corresponds to a situation in which one user wants to answer a post from another user or simply direct the message to someone else. For example, a tweet of user A in reply to user B would be a post like '@B [content of the tweet]'. The later is used to propagate a message: A re-tweets B means that A posted a message like 'RT @B [content posted by B]'. Likewise, there are two other groups of interactions, defined by [10]: *mentions* and *attributions*. Tweets that fit into the first group are those that mention another user in the middle of the text, whereas a tweet in the second group is similar to a

[1] http://apiwiki.twitter.com/Streaming-API-Documentation

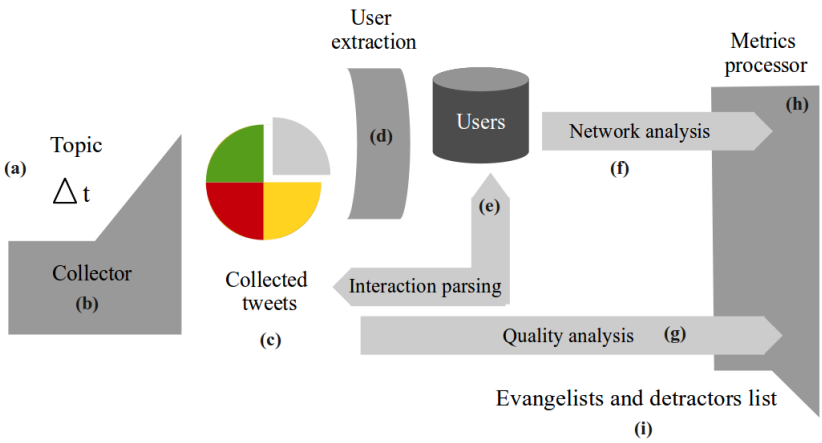

Figure 1: Overview of the technique.

re-tweet, except that an attribution cites the username following the expression 'via' instead of using the RT notation. These parsed interactions can be used as a measure of the user's influence: the more a user is mentioned, replied or re-tweeted by others, the more she causes others to react.

Both artifacts generated in steps (d) and (e) are used in the network analysis (f). As explained later, topology features are extracted considering each user and added as input to the metrics processor (h). Simultaneously, in step (g), quality evaluators are applied to the tweets. Specifically, readability indicators [5] are computed for each tweet in order to distinguish meaningful posts from *flood* – content that does not express relevant opinion. The values are subsequently assigned to the respective users. Along with the topology features and the user information (including profile extractable and sentiment-related data), the quality factors are provided as input to the processor (h). And, finally, as a result of the metrics combination a list of evangelists and detractors is produced (i).

4. NETWORKS OF EVANGELISTS AND DETRACTORS

4.1 Metrics

This section defines the metrics used to characterize the networks we analyze to detect evangelists and detractors on Twitter. In order to understand the characteristics of the roles of users on Twitter, we adopt a complex network approach to analyze the collected data, looking at the characteristics of social network graphs that emerge from the interactions between users that interact using tweets. A social network is a set of people or groups of people with some pattern of interactions between them. Social networks are useful for analyzing interactions that involve the interactions of a large number of entities, such as users and themes. From the several networks that naturally emerge from the user interactions enabled by Twitter features, we select two of them for an in-depth analysis: Follower/Following Network and Interactions Network.

For a given subset of users involved in a specific theme, let (G_i, U) be the user directed graph, where (u_1, u_2) is a directed arc in U if user $u_1 \in G_i$ has cited (i.e., via attribution, mention, reply or re-tweet) user $u_2 \in G_i$. Similarly, let (G_r, U) be the user directed graph, where (u_1, u_2) is a directed arc in U if user $u_1 \in G_r$ follows user $u_2 \in G_r$. Figure 2 displays a visual representation of both graphs G_i and G_r, where we can observe the behavior of users involved in discussions about different brands of soda. For instance, the marked node in G_i represents a teen celebrity that commented about a specific soda brand, which generated a number of replies, as represented by the edges pointing to the node.

We use a number of graph network metrics to analyze the user interaction network in order to find influential users for a specific topic. Specifically, we look at individual node properties, such as degree, betweenness and centrality:

Betweenness: It is most often calculated as the fraction of shortest paths between node pairs that pass through the node of interest [4]. In graph G_i, users with high betweenness have important role in the information dissemination process.

Eigenvector Centrality: We use the Eigenvector Centrality (EC) algorithm [2] to assign a value to each user who was quoted or cited in tweets. The intuition is that a user has a high rank value if she received responses from many users or from users who also have a high rank value. This feature identifies the most influential users assuming that its connections are also influential. Another intuitive justification is that a user can have a high EC if there are many users that point to her, or if there are some users that point to her and have a high EC. This metric is calculated for users in both graphs, G_i and G_r.

In-degree: One key characteristic of the structure of a directed network is the in-degree. In the interaction graph, the in-degree measures the number of times a user was cited or had her tweets replied or re-tweeted. The in-degree metric is an indicator of the user relevance in the interaction graph.

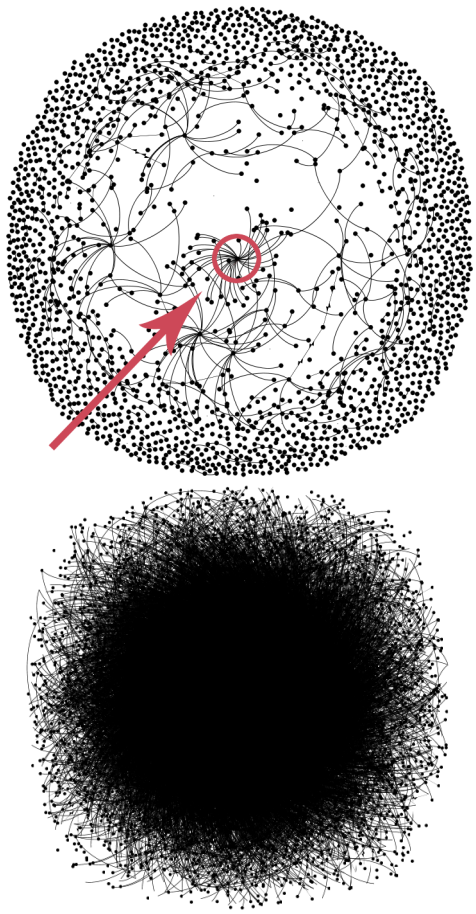

Figure 2: Graphic representation of G_i and G_r.

4.1.1 User Charateristic

Another user characteristic also used in our analysis is the TFF Ratio (Twitter Follower-Foloowee Ratio): the ratio of a user's followers to followees (or people who the user follows). As discussed in [3], the number of profiles following a user indicates her popularity, but is not directly related to her influence among others. Most of the users do not interact with her connections [7]. Since it is easy and costless to follow others on Twitter, many people start following others purely as a courtesy for being followed, but they do not keep up with their tweets. In [10], the TFF ratio is presented as an adequate method for studying the collected data. We use this metric, combined with others, to identify influential users in our dataset, considering the users with higher TFF Ratio as more relevant.

4.1.2 Content Characteristics

After an initial analysis of the collected data, we noticed that some users considered as evangelists had a significant amount of *flood-like* tweets (i.e., meaningless posts only expressing their love for the brand). Thus, in order to evaluate the quality of tweets, we associated features that measure how well written and understandable is a tweet. The metric combines statistics about the number of words, syllables and sentences. In particular, we use the Flesch-Kinkaid [12] metric, which was designed to indicate comprehension difficulty when reading a passage of contemporary academic English. For each tweet, it computes the average number of syllables per word and the average sentence length – equation (1). The metric, successfully applied in the identification of high-quality Wikipedia articles [5], was adapted for Portuguese (since our dataset was collected in this language to faciliate analysis), and increased the accuracy of the influential identification. The user quality factor was determined as the average of the Kinkaid factor computed for each one of her tweets. As a measure of dispersion, the standard deviation was also calculated.

$$kinkaid = 0.39 * \frac{words}{sentences} + 11.8 * \frac{syllables}{words} - 15.59 \quad (1)$$

4.2 User Rank

We use different metrics and sentiment information to assign a single value to each user in the database, so that we can obtain a user ranking. Since each metric has its own natural range of values, we normalized each metric as a percentage of its largest value. Sentiment information was also normalized, for it is the feature that discriminates influential users as evangelists or detractors. If a user has neutral polarity, the result of the equation is zero, regardless of the other features. The user value reflects different characteristics of the user, such as how often the user emitted positive/negative opinions, the user centrality and the quality of her texts. The user rank is given by the following expression:

$$U_{rank} = \frac{(\alpha * polarity) + (\beta * network) + (\gamma * quality)}{\alpha + \beta + \gamma} \quad (2)$$

where:

Polarity: The normalized overall sentiment of the user (positive tweets - negative tweets).

Network: Network normalized component. It is the combination of different network metrics that synthesize user positioning: betweeness, eigen centrality, in-degree, out-degree (of generated networks and of the Twitter network).

Quality: Normalized average of user's text quality. In this component only the Kinkaid metric is used.

α, β, γ: Constants, greater or equal to zero, that determine the proportion between network, sentiment and quality metrics. The value of each constant was determined experimentally, as discussed in section 5.2.1.

The main idea behind the equation is that one feature by itself may not be enough to characterize a user as an evangelist or detractor. For example, someone that is well connected but does not have a biased opinion might not be a point of interest in the analysis. A user that is well connected in the graph, have biased opinion, and write often and with quality might be ranked as an influential user.

The result is a value between -1 and 1 for each user. For $U_{rank} < 0$, the given user is a detractor, for $U_{rank} > 0$ the user is an evangelist. If $U_{rank} = 0$, the user is neutral. When we classify the users in descending order we obtain the following: evangelists appear on the top of the list. Neutral users are positioned in the middle of the list with zero result.

Category	Tweets	Users
Total	14,127	12,069
Brand-specific		
Positive	3,083	2,770
Negative	824	714
Neutral	4,156	3,401
Total	8,063	6,885

Table 1: Number of tweets and users per sentiment in the data set.

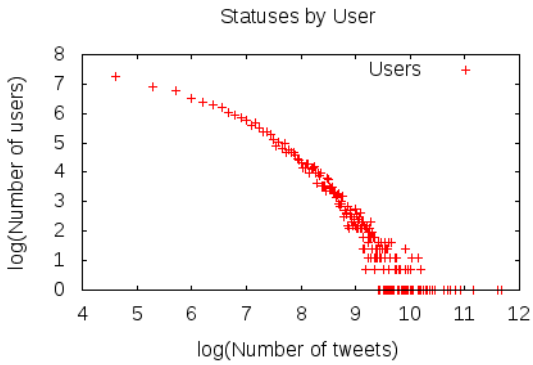

Figure 3: The number of tweets posted by each user.

Detractors, at the end of the list, with the smallest negative values representing the most influential detractors.

5. EXPERIMENTS AND DISCUSSION

In order to evaluate our approach, this section introduces a dataset of tweets posted by Brazilian users in section 5.1. This dataset is then evaluated and validated in section 5.2.

5.1 Dataset

In order to determine influential users, evangelists and detractors, it is first necessary to choose a subject. Focusing the influential analysis in a specific topic is a matter of *design* rather than a limitation. In this evaluation, we collected tweets regarding *soda brands*. We employed the proposed architecture to gather tweets posted by Brazilian users between August 2009 and September 2009. With such information, we are able to built a snapshot of the soda market in Brazil in such a period. The resultant dataset has 14,127 tweets from 12,069 users. In total, 13 different soda brands were tracked and the most one popular was chosen as the reference for the tweet polarity classification by the analysts. Table 1 presents the number of tweets and users for the whole collection and for the selected brand along with their respective number of positive, negative and neutral classification.

This topic-specific dataset has the same characteristics as some previously analyzed samples of the Twitter network that are not restricted to a topic [7], [9]. Figure 3 and Figure 4 show that both, statuses per user and degree distribution plots (respectively), obey a power-law behavior [11]: they appear roughly as a straight line when plotted on logarithmic scales.

Additionally, Figure 5 shows a scatter plot of the following-

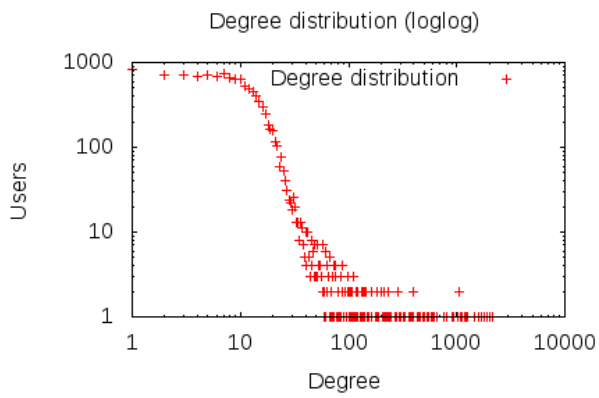

Figure 4: Degree distribution for users in the data set.

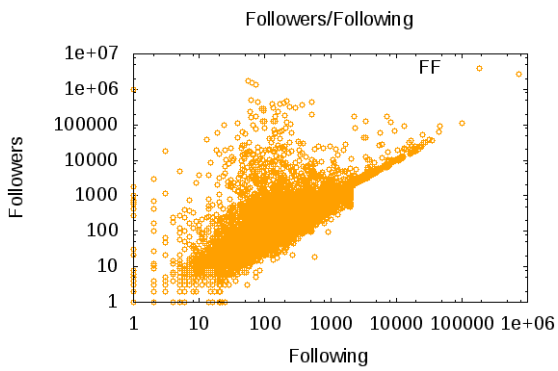

Figure 5: Followers x Following scatter plot

follower relation. As in [9], there are three identifiable groups of users: (*i*) the ones located near the $x = y$ line; (*ii*) the ones that appear in the region above the diagonal; (*iii*) and the ones located below the diagonal. The characteristic of the first group illustrates the reciprocity in the relationships: this behavior is common in regular users who follow their friends and are followed by them. The third group contains users that follow way much more people than those who follow them. This is assumed as typical behavior of spammers and other kinds of users who follow the others in order to gain attention and be followed. The only difference from our specific dataset collected following the proposed technique from an ordinary *non-topic-restricted* base is that there are fewer representatives of the third group. That happens because the set of users was built from their posted tweets. Since their main feature is that they do not tweet much [9], their representation in this dataset is smaller than usual. In order to be an influential user, the person must be an author: she must tweet.

Table 2 compares the number of vertices and arcs of the interaction and relation networks. As was shown in [7], the graph of interaction is considerably more sparse than the relation graph. Accordingly, the number of arcs in G_r is much larger than in G_i. The difference between the number of authors of the collected tweets (12,069) and the number of users in the relation network (11,641) is due to changes

	Interaction (G_i)	Relation (G_r)
Vertices	5559	11641
Arcs	3271	81911

Table 2: Graph statistics for G_i and G_r.

in the user profile. From the time the tweets were collected to the time the user information was collected, some users changed their usernames and others protected their accounts making it unavailable to collect the network.

5.2 Experiments

In order to compare different metrics for ranking evangelists and detractors, a marketing and communication specialist created a list of influential users for the studied theme. In the specialist analysis, an influential user is the one who is not only well connected but also produces content with the intention of changing people's opinions. Among the users in the dataset, the specialist identified 17 influential users.

Assuming the specialists's list as a ground truth, the proposed technique was assessed using several performance measures[1], as follows:

- *Precision*: The number of relevant documents retrieved by a search divided by the total number of documents retrieved by that search.
- *Recall*: The number of relevant documents retrieved by a search divided by the total number of existing relevant documents (those that should have been retrieved).
- *Average precision (AP)*: This metric is a form of evaluation that considers the order in which the relevant documents were ranked (Formula 3). It is the average of precisions computed at the point of each relevant document in the sequence (Formula 4).
- *Mean Average precision (MAP)*: It is defined as AP averaged over all queries.

$$P(n) = \frac{\text{number of positive instances within top n}}{n} \quad (3)$$

$$AP = \sum^{N}{}_{n=1} \frac{P(n) * position(n)}{\text{number of positive instances}} \quad (4)$$

In this paper, a document is an user and a relevant document an influential user. For MAP, the queries correspond to its evangelist and detractors result. The proposed algorithm was designed to assist the analyst on the monitoring task by providing a list of TOP evangelists and detractors. Thus, in the experiments, the goal is to maximize the *recall* in each of the generated rankings. Moreover, the formula was tested using both relation and interaction networks. Although high *precision* and AP are desired, we focus on MAP because it is important to measure the overall quality of the formula.

5.2.1 α, β and γ calculation

As stated before, a single feature may not be good enough to classify users into the different groups. In order to test this claim, different rankings were generated using only one component of the equation. Figure 6 presents recall results

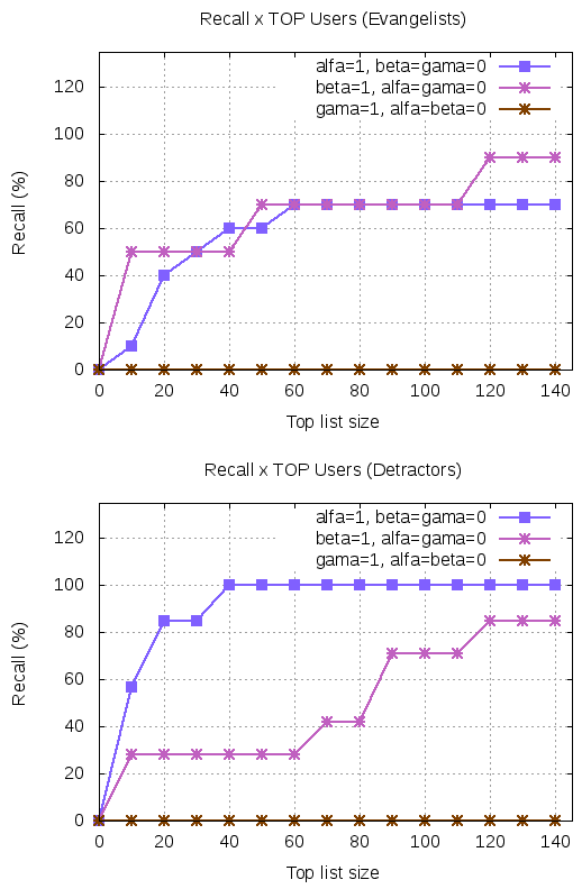

Figure 6: Comparison between equation components.

for each isolated component. As can be seen, the result for evangelists is better when using network features. This happens because of the large number of users with positive sentiment. However, this is not the best method since it was necessary to consider a 370 top users list to retrieve all evangelists. On the other hand, the best component for detractors is the sentiment since the negative polarity is clearer for the analyst to identify, which makes this information more precise.

In order to find a good set of parameters a simple method was used as follows. Every linearly independent set of α,β and γ varying from 0 to 10 was tested. For each combination, AP, MAP, precision and recall were calculated, storing the best values. The chosen parameters that were used in the following experiments were: $\alpha = 9$, $\beta = 9$ and $\gamma = 1$. Notice that here we are not worried about how to automatically identify the best parameters, which can be done with machine learning techniques. We leave that for future work.

5.2.2 AP, MAP, precision and recall

Using the calculated parameters, two ranks of users were generated: one using the interaction network as topology feature and the other using the relation network information. Table 3 shows the AP and MAP values for the generated rankings.

This generated ranking was used to create the TOP users

Metric	Interaction (G_i)	Relation (G_r)
AP-evangelists	0.173055259064	0.116782717081
AP-detractors	0.583158263305	0.577286502735
MAP	0.378106761184	0.347034609908

Table 3: AP and MAP values for generated rankings

lists. Each list contained the top ranked users for both types: evangelists and detractors. Then precision and recall were computed for each experiment. These results are shown in Table 4. Note that the Interaction Network produce better results for recall and precision.

Figure 7 shows a comparison between the two approaches (relation and interaction networks). The y axis represents the Recall for the Interaction Network and the Relation Network in each experiment. The interaction network always leads to better results for both types of users, which is confirmed by the confidence interval of the recall difference: (6.65, 14.77) for detractors and (14.13, 27.3) for evangelists with 95% of confidence.

The difference in accuracy for both network approaches is once again showed by the number of users needed to retrieve all evangelists and detractors. For the Relation Network equation, 570 users were needed in comparison with 140 needed by the Interaction Network. The AP values for evangelists reflects the fact that is harder to classify this kind of user. This characteristic is explained further.

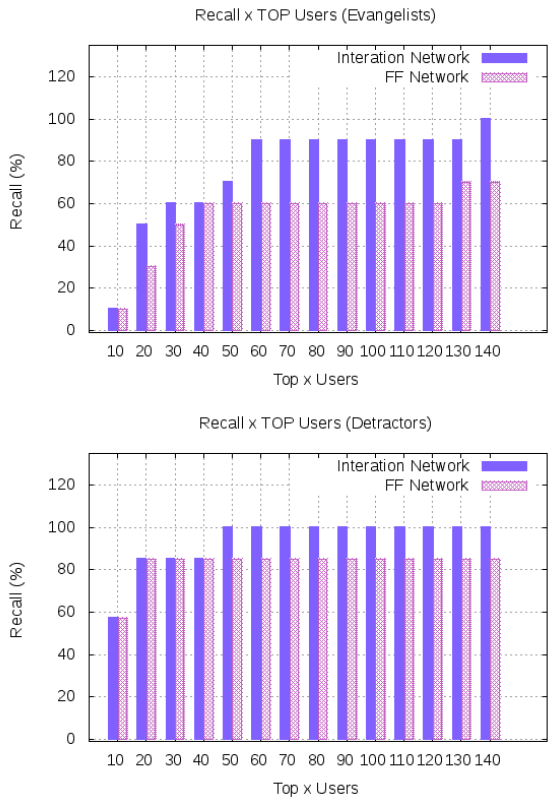

Figure 7: Comparison between Interaction network and Relation network (Recall G_i - Recall G_r).

Figure 8 shows the recall for each list of top users using

Figure 8: Recall for x retrieved users

the Interaction Network Features. Considered the 20 users retrieved, the recall for detractors is 85% and for evangelists it is 50%. This difference is mainly because it is easier for an analyst to classify a detractor: tweets with a content that is not really clear if it has positive or neutral polarity are very common. The last retrieved evangelist, in both methods, is one example of this problem. The specialist himself noted that this user is between evangelist and neutral.

Finally, the computational complexity of the extraction of betweeness and eigenvector centrality (EC) for both G_i and G_r was analyzed. Each metric was extracted 10 times for each network and Table 5 exhibits the average mean cost (μ) and the standard deviation (s) obtained. As expected, given the number of vertices and arcs shown in Table 2, the cost to compute features in the interaction network is low. G_i expresses only the real content-based connections between users reducing the problem complexity.

	Betweeness		EC	
	μ (sec)	s (sec)	μ (sec)	s (sec)
Interaction (G_i)	0	0	1.96	0.21
Relation (G_r)	123.17	5.34	8.84	0.4

Table 5: Computing time comparison, in seconds, of betweeness and eigenvector centrality in G_i and G_r.

6. CONCLUSION

This paper analyzes user behavior and connections in order to determine their influence in the Twitter network. Specifically, for each user, her tweets' readability features and polarity are extracted, and her position in two diferent networks (interaction and relation) of people that talk about the same topic is analyzed. Once a specific subject is defined, the evangelists and detractors can be determined. The identification of these users is crucial for social media analyzers that behold on Twitter network a potential viral marketing environment and want to idenfity the ones who are the influentials concerning one subject.

The obtained results are a proof of concept that our proposed techniques are able to rank users, by an influcence factor, as evangelists or detractors. The results for finding detractors are visibly more accurate than the evangelist's.

	Evangelists				Detractors			
	Precision		Recall		Precision		Recall	
Rank size	IT	FF	IT	FF	IT	FF	IT	FF
TOP 10	10.0%	10.0%	10.0%	10.0%	40.0%	40.0%	57.0%	57.0%
TOP 20	25.0%	15.0%	50.0%	30.0%	30.0%	30.0%	85.0%	85.0%
TOP 30	20.0%	16.7%	60.0%	50.0%	20.0%	20.0%	85.0%	85.0%
TOP 40	15.0%	15.0%	60.0%	60.0%	15.0%	15.0%	85.0%	85.0%
TOP 50	14.0%	12.0%	70.0%	60.0%	14.0%	12.0%	100.0%	85.0%
TOP 60	15.0%	10.0%	90.0%	60.0%	11.6%	10.0%	100.0%	85.0%
TOP 70	12.9%	8.6%	90.0%	60.0%	10.0%	8.6%	100.0%	85.0%
TOP 80	11.3%	7.5%	90.0%	60.0%	8.8%	7.5%	100.0%	85.0%
TOP 90	10.0%	6.7%	90.0%	60.0%	7.8%	6.7%	100.0%	85.0%
TOP 100	9.0%	6.0%	90.0%	60.0%	7.0%	6.0%	100.0%	85.0%
TOP 110	8.2%	5.5%	90.0%	60.0%	6.4%	5.5%	100.0%	85.0%
TOP 120	7.5%	5.8%	90.0%	70.0%	5.8%	5.0%	100.0%	85.0%
TOP 130	6.9%	5.4%	90.0%	70.0%	5.4%	4.6%	100.0%	85.0%
TOP 140	7.1%	5.0%	100.0%	70.0%	5.0%	4.3%	100.0%	85.0%

Table 4: Precision and Recall (Evangelists)

This happens due to the occasional difficulty for distinguishing between a neutral and a positive-biased tweet during the manual classification. For the negative tweets, this boundary is usually clearer.

Since there is no benchmark for influential users detection (a default dataset with tweets and users previously classified), one significant effort of this work was to built such a test collection. This is not a trivial task due to the difficulty to classify posts as positive or neutral (this is a subjective problem by nature).

The experiments results also demonstrate that the interactions (mentions, replies, re-tweets, attributions) of an user with others is a better representation of her influence than her connections (follower, following). The precision and recall values for the generated ranks, using the interactions, were always better. Another substantial remark is that the interaction network is more sparse than the relations network. This turns the computational cost must cheaper and with more accurate results.

As future work, we can address the problem of finding α, β and γ parameters. The developed technique also needs further testing in real environments (with evangelists and detractors identified) and on data bases with different themes.

Acknowledgments. The authors would like to thank Rodrigo Pazzini for the analysis of user influence and insightful comments. This work is partially supported by the projects INCT-Web (MCT/CNPq grant 57.3871/2008-6) and by the authors' individual grants and scholarships from CNPq, CAPES and FAPEMIG.

7. REFERENCES

[1] R. A. Baeza-Yates and B. Ribeiro-Neto. *Modern Information Retrieval*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1999.

[2] P. Bonacich. Some unique properties of eigenvector centrality. *Social Networks*, 29(4):555 – 564, 2007.

[3] M. Cha et. al. Measuring User Influence in Twitter: The Million Follower Fallacy. In *Procs. Intl. AAAI Conf. on Weblogs and Social Media (ICWSM)*.

[4] L. F. Costa et. al. Characterization of complex networks: A survey of measurements. *Advances In Physics*, 56:167, 2007.

[5] D. H. Dalip et.al. Automatic quality assessment of content created collaboratively by web communities: a case study of wikipedia. In *Procs. Joint Conf. on Digital libraries (JCDL)*, pages 295–304, 2009.

[6] N. A. Diakopoulos and D. A. Shamma. Characterizing debate performance via aggregated twitter sentiment. In *Procs. Intl. Conf. on Human Factors in Computing Systems (CHI)*, 2010.

[7] B. A. Huberman, D. M. Romero, and F. Wu. Social networks that matter: Twitter under the microscope. *ArXiv e-prints*, December 2008.

[8] A. Java et. al. Why we twitter: understanding microblogging usage and communities. In *Procs. of WebKDD and SNA-KDD Workshop on Web Mining and Social Network Analysis*, pages 56–65, 2007.

[9] B. Krishnamurthy, P. Gill, and M. Arlitt. A few chirps about twitter. In *Procs. Workshop on Online Social Networks (WOSP)*, pages 19–24, 2008.

[10] A. Leavitt et. al. New approaches for analyzing influence on twitter. Technical report.

[11] M. E. J. Newman. Power laws, pareto distributions and zipf's law. *Contemporary Physics*, 46(5), 2005.

[12] S. Ressler. *Perspectives on electronic publishing: standards, solutions, and more*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1993.

[13] T. Rodrigues, F. Benevenuto, V. Almeida, J. Almeida, and M. Gonçalves. Uma análise contextual de conteúdo duplicado no youtube. In *WebMedia 2009*, Fortaleza, CE, Brasil, 2009.

[14] Twitter Blog. Measuring Tweets. `http://blog.twitter.com/2010/02/measuring-tweets.html`, 2010.

[15] J. Weng et. al. Twitterrank: finding topic-sensitive influential twitterers. In *Procs. ACM Intl. Conf. on Web Search and Data Mining (WSDM)*, pages 261–270, 2010.