

Social Network and Sentiment Analysis on Twitter: Towards a Combined Approach

Paolo Fornacciari, Monica Mordonini, Michele Tomaiuolo

Dipartimento di Ingegneria dell'Informazione
Università degli Studi di Parma
Parma, Italy

e-mail: paolo.fornacciari@studenti.unipr.it,
{monica.mordonini@unipr.it,michele.tomaiuolo@unipr.it}

Abstract. Twitter is a platform which may contain opinions, thoughts, facts and other information. Within it, many and various communities are originated by users with common interests, or with similar ways to feel part of the community. This paper presents a possible combined approach between Social Network Analysis and Sentiment Analysis. In particular, we have tried to associate a sentiment to the nodes of the graphs showing the social connections, and this may highlight the potential correlations. The idea behind it is that, on the one hand, the network topology can contextualize and then, in part, unmask some incorrect results of the Sentiment Analysis; on the other hand, the polarity of the feeling on the network can highlight the role of semantic connections in the hierarchy of the communities that are present in the network. In this work, we illustrate the approach to the issue, together with the system architecture and, then, we discuss our first results.

Keywords: Sentiment Analysis, Social Network, Hierarchical Classification

1 Introduction

An increasing number of people is progressively approaching to the social networking sites, which become more and more popular and complex: within their context many and various communities are originated by users with common interests or with similar ways to feel part of the community. The kinds of analysis as well as information that can be extracted from the social networking sites are varied and increasingly appealing both to the world of marketing and to the social or political one. The classical approach to Social Network Analysis allows to study the topology of a network through the connections that develop within it, giving rise to a hierarchy of communities within the main topic. Furthermore, certain types of social networks, like Twitter, allow to track relationships also in those cases in which knowledge is not mutual: simply a node is a follower of another node. The number of followers defines in part the popularity of a node within the network, but it is not able to point out if this popularity is positive or negative.

On the other hand, the explosion of data on the Web has made the research in automatic cataloging of texts increasingly interesting, as well as the extraction of

information or meta-information and the Sentiment Analysis of a review, an emotion, a tweet. Moreover, in this area the explosion of microblogging, and the use of a simple “like” or a retweet as a form of acceptance or sounding board for information as well as the dynamism and the speed with which everyone reads and writes content make the analysis of these opinions hard, if you use the methods of text mining, while they introduce, or amplify, new issues and problems for sentiment analysis (such as citations, irony, role of emoticons) that are difficult to deal with regardless the context in which they are written.

This paper presents a combined approach between Social Network and Sentiment Analysis. In particular we have tried to introduce some kind of information about sentiments on the graphs showing the results of the Social Network Analysis (SNA): in this way we hope to highlight other potential correlations among the nodes of net under examination. The idea behind it is that, on the one hand, the network topology and the selected topics of the network can contextualize and then, in part, unmask some incorrect results of the Sentiment Analysis (SA), and, the other hand, the polarity of the feeling on the network can highlight the role of semantic connections, as a possible foundation for the organization and the hierarchy of the communities highlighted by the Social Network Analysis.

In the following, after a brief description of the background, the system architecture will be showed, together with the choices which we made, then some results obtained from the initial evaluation of the system will be discussed.

2 Background

SNSs are a collection of web-based services that allow users to build a profile within the system and define a list of other users with whom they have some kind of connection [7]. The architecture of social networking platforms is very differentiated. While the most popular platforms are built as essentially centralized systems, other platforms have a distributed architecture [10][11]. The decentralized systems try to address some of the risks associated with online social networking, which are often perceived as quite serious by many users and have already led to serious incidents [6]. SNA has the objective to model social structures with different properties, starting from the mathematical theory of graphs and the use of matrix algebra, and is often augmented through computer-based simulations [11]. SA is a branch of Opinion Mining, that aims to listen and process the data that users post on social media. Generally SA classifies web comments into positive, neutral, and negative categories. To make these systems more intelligent and flexible, a deeper analysis of affective knowledge could be incorporated [9][8]. In some case an ontology driven approach is used [5][24][3].

In this research work, we built a system for social network and sentiment analysis, which can operate on Twitter data, one of the most popular social networks. The analysis of large amount of data is an exciting challenge for researchers, but it is also crucial for all those who work at different levels in the current information society: Twitter has been the subject of attention from researchers as early as 2009 [13].

Some recent studies about American candidates are important for understanding how public sentiment is shaped and its polarization [19]. In [2] geo-spatial information related to tweets is used for estimating happiness in the Italian cities. Being Twitter a microblogging service, the techniques used generally in SA and Text Classification must be adapted to the famous 140-character tweet and this opens the way for new issues [1][17][16][26].

Another quite important problem to work on Twitter data is how to automatically collect a corpus for SA and, in general, Opinion Mining purposes: example of how to perform this task is in, for example, [21][14].

3 System Architecture

In this paragraph we describe our system for social network and sentiment analysis, which can operate on Twitter data.

Twitter is a platform which may contain opinions, thoughts, facts, references to images and other media and, recently, stream video filmed live and put online by users. So it is more than just a SNCs in which a user displays and increases their social relationships, it is a real communication channel in which a user can choose its topics and its node of reference according to his interests and culture.

A study of the network topology and the number of interconnections of a node are able to highlight the communities in the network and also in part to how the information is propagated, but they are not able to say anything about the degree of agreement and cohesion of members of a community. To solve this task you need to carry out an investigation into the semantic content of the messages.

Compared to the problems of classic data mining, sentiment analysis shows many difficulties in terms of effectiveness. This is mainly due to the subtle distinction that exists between positive and negative sentiment or between neutral and positive one. Let us suppose for example a sentence containing irony or sarcasm, where the interpretation of the meaning is strictly subjective. In this case, two human beings may be in disagreement about the real feeling that it expresses. Furthermore, not always the opinions are expressed through the use of opinion words, in many cases the special language constructs (such as the figures of speech) come into play.

Difficulties also are due to the use of non-formal expressions and slangs that do not belong to the vocabulary of a language. These terms are often used in an intensive way to express a particular opinion or a certain mood.

Additional problems are due to the domain of the subject: in particular we note that the feelings that are expressed by a word are often dependent on the topic. We look at this sentence as an example: "It's quiet!". It shall render a positive opinion if we are talking of a car engine, but it reveals a disapproval if the matter of discussion is a phone.

As a microblogging service, Twitter is used to publish short messages counting a maximum of 140 characters (tweets). This characteristic if one side it may seem easier because it forces people to take a position, on the other side the few words not allow

As a start, a tweet may contain many elements that are not significant for our classification, and can thus be dropped through a filtering process. To polish the message, we defined various filters, that can be applied in a customizable sequence. An example is shown in Fig.2.

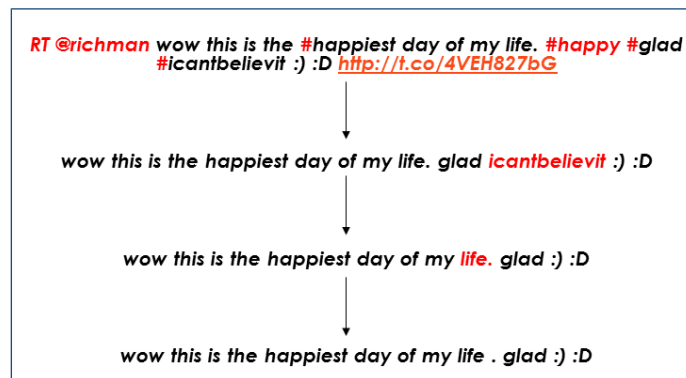


Fig. 2. The sequence of cleaning the tweet.

A first filter eliminates useless tokens such as: the “RT” sequence; the @ character and the whole following user name; the # symbol, but not the following topic name, which is kept in the message. The topic name is also removed, though, when it coincides with the name of the channel where tweets are collected from.

A second filter applies the language specific rules. It includes an orthographic correction of the message, which is used to remove unknown words (in the example: “icantbelievit”) and other filtering processes for stemming and removal of stopwords.

Finally, another filter separates all punctuation symbols from the text, and organizes them as single-character words. Even if smiles sequences, repeated question and exclamation marks are kept as aggregates because they are important patterns for the classification.

The final result of the filtering process is a word vector, which is then submitted to a set of classifiers.

We use a set of classifier to identifying the following classes of messages: undiscriminated, objective, subjective, positive, negative. Moreover, there is a class in which the system put all the tweets that are too short to be classified. The system is organized as a simple hierarchy of agents, mimicking the hierarchy of sentiment classes. In fact, since objective messages have no polarity by definition, the classifier for positive and negative sentiments is only applied to subjective messages (see Fig. 3). One advantage of this framework for classifiers is the ease with which you can add classifiers trained to identify other emotions. In fact, hierarchical classification has been applied successfully in a number of studies, for information retrieval [23]. It has been proven effective especially in the case of classification over hierarchical taxonomies. Also in the case of sentiment analysis, a hierarchy of classes can be defined [12][5]. Accordingly, hierarchical classification has already been applied to sentiment analysis, too [23].

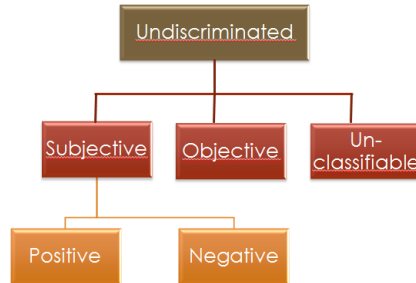


Fig. 3. Hierarchy of basic sentiment classes.

Each classifier is based on Multinomial Naive Bayes algorithm, that one of the most popular methods used in SA. We have selected it because it seems to be the most suitable to generate and process large sets of features. In fact, instead of generating a training set by hand, we aimed at realizing an automated (or at least semiautomated) process for obtaining good training sets. In our methodology, the training sets are obtained through the automatic elaboration of some particular streams of tweets and comments, obtained directly from Twitter, without any manual classification. Thus, each training set may contain an important number of wrong data. Nevertheless, we show that they can be used to obtain useful results.

About the objectivity/subjectivity classifier, we adopted a similar strategy to [20]. In fact, to obtain objective content, we gathered messages generated from popular news agencies. In our tests, we used the following list: *@ABC*, *@BBCNews*, *@BBCSport*, *@business*, *@BW*, *@cnbrk*, *@CNNMoney*, *@fox32news*, *@latimes*, *@nytimes*, *@TIME*. To obtain subjective content, instead, we gathered comments directed to the same list of users.

About the polarity classifier, we used different sources, thus generating training sets which do not overlap with those about objectivity/subjectivity. In fact, we used sources of mostly positive or negative messages, respectively. On the one hand, those sources should fit the particular setting of Twitter (short messages, idiomatic expressions, smiles, etc.). On the other hand, they should not be specific to a particular topic or context (sport, music, etc.). Thus, we dropped the idea of collecting messages about particular events, mostly generating either positive or negative sentiments. Instead, we collected messages, using generic yet polar terms as queried hashtags. In particular, we used the following channels to gather positive content: *#adorable*, *#awesome*, *#beautiful*, *#beauty*, *#cool*, *#excellent*, *#great*. We used the following channels to gather negative content: *#angry*, *#awful*, *#bad*, *#corrupt*, *#pathetic*, *#sadness*, *#shame*. Actually, such terms have been chosen quite empirically, taking into account the quality of training sets they generated. But they could be selected from WordNet-Affect [24], SentiWordNet [3], and other affective lexicons, in a more systematic way.

In this way, the training set is generated in an automated fashion, as a list of tweets. Each tweet is associated with its supposed class, in accordance to its source. In fact, the training set is not perfect, as it contains messages gathered from public channels.

However, a training set of this kind can be generated easily and in a methodical way, from real and updated Twitter messages. Moreover it is possible to extend this approach to train a classifier to recognize feelings which are written in a particular slang.

Feature	$P(F_i pos)$	$P(F_i neg)$	Feature	$P(F_i obj)$	$P(F_i sub)$
:)	0,0025	0,00055	!!!	0,000031	0,002
stupid	0,000098	0,00065	:(0,0000079	0,00030
thank you	0,0012	0,00029	%	0,0013	0,00080
!!!	0,0028	0,0019			

Fig. 4. Generated model for the classifier: example of selected feature and their probabilities in the polarity (on the left) and subjective (on the right) classifiers.

In Fig. 4 there are some examples of features which are selected by the classifiers together their probabilities. It is worth noting that these are consistent with what we expected: the emoticons ‘:)’ has a high probability of being in positive phrases, while the pattern ‘!!!’ is very significant for the classifier of the subjectivity but it is a useless feature to determine the polarity of a tweet.

4 Experimental results

In this section, we will report the results of the classifiers and the analysis carried out on a couple of case studies. Using the methodology and the software which we described in Section 3, it is possible to obtain some generic training sets for the classifiers. This phase was carried out before selecting the final case studies. In our settings, they consist of:

- 86000 instances (polarity);
- 32000 instances (subjectivity).

These instances have been obtained by exploring more than 60 channels on the social network. In the generated models, the selected features are consistent with our expectations: the typical expressions of a certain feeling (such as smileys, or some words that express appreciation or disgust) show a higher probability of belonging to the class of that feeling, rather than to the class of the opposite sentiment.

The results obtained by the classifiers using cross-validation (with folds = 10) on the training sets showed an accuracy of:

- 77,45% (polarity classifier)
- 79,50% (subjectivity classifier)

These results show that the model of the classifiers contains effective features for the recognition of the sentiment of a message.

The case study which was considered in this work is the social network of the *#SamSmith* channel (the singer who won four awards at the Grammy Awards 2015). The choice of this channel is justified by the strong similarities found between the

type of the published tweets and the instances used for training the classifiers. All data were downloaded between 2015-02-02 and 2015-02-10. The awarding of the Grammy took place on 2015-02-08. The social network (shown in Fig. 5) consists of a total of 5570 nodes (users) and 6886 arcs (“follows” relationships). Nodes are deployed according to the ForceAtlas2 algorithm [15], which turns structural proximities into visual proximities, thus highlighting communities.

Looking at the figure, it is possible to notice that the network topology is consistent with the nature of the considered case. In fact, most of the channel consists of independent users (or small groups of users) that express their opinion about the artist; however, in the central part of the network there are some major communities.

As shown in Fig. 5, the prevailing sentiment detected from the classifier is the negative one. Performing an analysis on a sample of tweets in the network, we noticed that many sentences are actually quotes of songs. These messages contain melancholic and sad phrases, and are therefore classified as negative. Considering that a quote is generally an appreciation for the artist, most users classified as negative are actually positive users. This is a typical example of a classic problem of misunderstanding of the SA: the system, while classifying correctly the tweet, misses the assessment of the feeling because it can not evaluate the tweet together with its context.

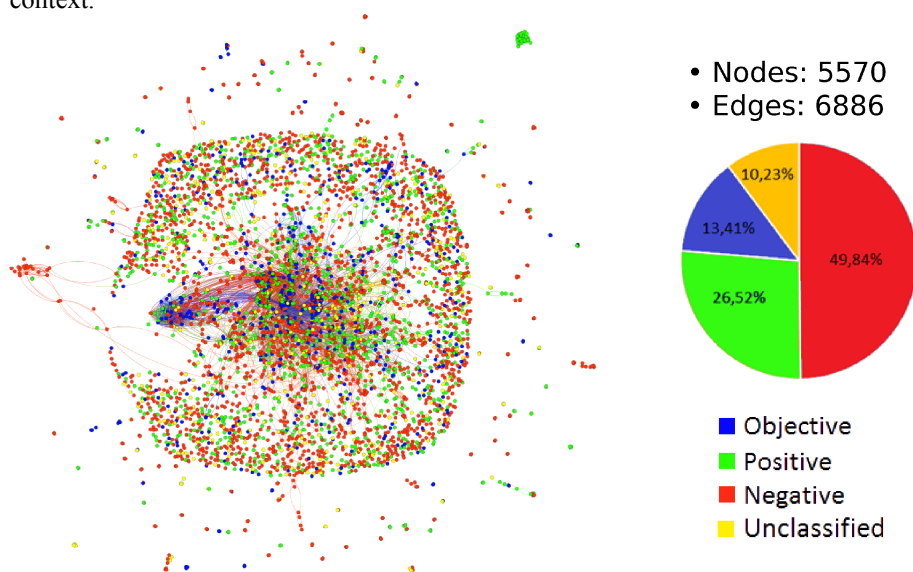


Fig. 5. Combined analysis of the *#SamSmith* channel.

For evaluating the performances of our system, we conducted a simple survey through a group of persons in our department. In this way, we selected and classified 100 messages that show a clear opinion on the singer. Then, we used those messages as a test. The results of the classifiers showed an accuracy of 84% for the polarity and 88% for subjectivity.

In the network periphery (at the top-right corner of Fig. 5), it is possible to notice a small group of users whose feeling is completely positive. After a careful analysis of users' tweets in this small group, it was found that these posts are mainly retweets and the original messages are only two. Of these two messages, the first is actually positive, while the other one is objective. This episode shows how some errors of assessment can have important impact on larger communities.

In addition to the *#Samsmith* channel, we considered the social network associated with the *#Ukraine* channel, trying to obtain some particularly significant results, above all from the point of view of network topology. In fact, the crisis in the region could lead to a quite sharp division on the Web. This work is still in progress, nevertheless we can show here some results which we already obtained.

At the moment we have downloaded the data, the network consisted of:

- 26131 nodes
- 1163588 edges

In Fig. 6, it is possible to see the main results of our analysis on the network.

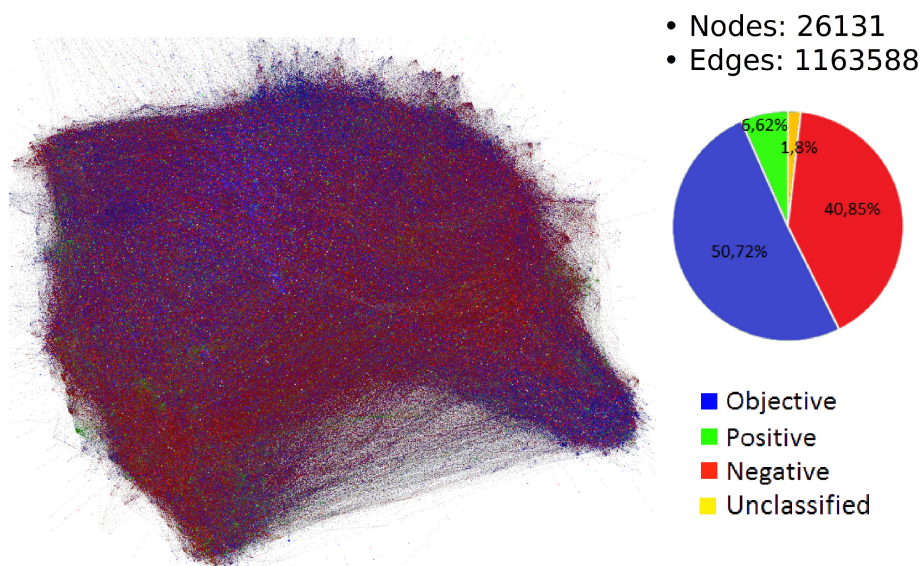


Fig. 6. Combined analysis of the *#Ukraine* channel.

The more evident thing to notice, is that the prevailing color in the network is blue (objective tweets), and the next one is red (negative tweets). Given the nature of the channel we are considering, which essentially reflects a social tragedy, the sentiment we have found through the analysis is quite plausible. However, analyzing some random messages, we have noticed a number of errors in the classification of these tweets. In particular, some objective sentences are often classified as negative ones, while some sentences expressing essentially hope (and thus positive) are classified as objective ones. In our opinion, the reason for these errors is related to the type of

features contained in the model of classifiers, which possibly are not a good fit for this particular case study.

The case of Ukraine has been discussed quite largely in traditional media, too, for the supposed role of “trolls” operating on new media to influence the public opinion [25]. In fact, this may represent, as a modern reposition, the quite classical case of opposing propaganda campaigns, this time carried on through social media. Also for this reason, we analyzed the social communities participating in the channel. We focused on the most active users, who contributed with at least 6 tweets during the whole week we considered (mid July 2014). In fact, among those it is more probable to find candidate opinion makers. The analyzed subnetwork represents around a tenth of the original network, and precisely consists of:

- 3261 nodes
- 84307 edges

We used the community detection algorithm provided with Gephi, at various resolution levels [18]. Quite interestingly, we were able to identify quite clearly two major communities. Additionally, some much smaller communities were found.

	Full Network	Community 1	Community 2
<i>Average degree</i>	51.706	53.021	42.649
<i>Diameter</i>	7	7	6
<i>Radius</i>	1	4	4
<i>Avg path length</i>	2.511	2.248	2.334
<i>Shortest paths</i>	10591776	3152400	2014980
<i>Graph density</i>	0.016	0.030	0.030
<i>Clustering coeff.</i>	0.420	0.480	0.414
<i>Total triangles</i>	873460	540526	281524

Table 1. Features of the main communities detected on the #Ukraine channel.

Looking at data reported in Table 1, it is easy to notice that the two communities, corresponding to opposing factions in the crisis, have a quite similar size. Moreover, also their main features are quite similar. This seems to indicate that the two camps have a quite similar internal social organization, at least at the macroscopic level. Nevertheless, both the communities have high density, almost doubling the value of the whole network. This means that, in fact, there is a quite clear separation between those two communities, which have relatively few shared connections.

Our sentiment analysis has not highlighted significant differences in the emerging opinions in the two communities. In fact, they largely share the same negative outlook of the whole network. This is an issue that we plan to analyse in deeper detail in future. The emerging sentiment in each camp may also vary during time, and in particular in correspondance with major events and turnpoints in the crisis.

5 Conclusions

This study reports the initial results we obtained from the synthesis of Social Network Analysis and Sentiment Analysis. We experimented our approach on a couple of Twitter channels, as case studies. In particular, we considered the *#SamSmith* channel during the Grammy Awards in 2015, and the *#Ukraine* channel during the crisis of 2014. Apart from the particular results, a methodology and some guidelines for the automatic classification of Twitter content have been discussed.

The implemented software allows: (i) to get a training set for the classifiers that deal with Sentiment Analysis, and (ii) to make a thorough study of the network topology. The study of the global sentiment within the network has highlighted the typical problems of Sentiment Analysis (irony, sarcasm, lack of information, etc.). Additionally, some peculiar problems of the considered channel were also detected (such as the quotes of songs). Also, the analysis of biased channels, may pose additional difficulties.

The performances obtained by the classifiers during tests conducted on the training set and the analysis of the case studies have shown good and promising results.

References

1. A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R. Passonneau. "Sentiment analysis of Twitter data", Procs of the Workshop on Languages in Social Media (LSM '11), Association for Computational Linguistics, USA, pp. 30-38, 2011.
2. L. Allisio, V. Mussa, C. Bosco, V. Patti, and G. Ruffo, "Felicittà: Visualizing and Estimating Happiness in Italian Cities from Geotagged Tweets," Proc. of the 1st Int. Workshop on Emotion and Sentiment in Social and Expressive Media: Approaches and Perspectives from AI (ESSEM 2013), Turin, Italy, 2013.
3. A. E. S. Baccianella and F. Sebastiani, "Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining," Proc. of the 7th Conf. on Inter. Language Resources and Evaluation (LREC'10), ELRA, 2010.
4. F. Bergenti, A. Poggi, and M. Tomaiuolo, "An Actor Based Software Framework for Scalable Applications," Lecture Notes in Computer Science, 8729, pp. 26-35. 7th International Conference on Internet and Distributed Computing Systems (IDCS), 2014.
5. M. Baldoni, C. Baroglio, V. Patti, and P. Rena, "From tags to emotions: Ontology-driven sentiment analysis in the social semantic web," *Intelligenza Artificiale*, vol. 6(1), pp. 41-54, 2012.
6. E. Franchi, A. Poggi, and M. Tomaiuolo, "Information and Password Attacks on Social Networks: An Argument for Cryptography," *Journal of Information Technology Research (JITR)*, 8(1), 25-42, 2015. doi:10.4018/JITR.2015010103
7. D. Boyd, N. Ellison, "Social Network Sites: Definition, History and Scholarship," *Journal of Computed-Mediated Communication*, vol. 13 (1), pp. 210-230, 2008.
8. K. Ca, S. Spangler. Y. Chen; L. Zhang, "Leveraging Sentiment Analysis for Topic Detection," *Web Intelligence and Intelligent Agent Technology (WI-IAT '08)*, IEEE/WIC/ACM Int. Conf. on, vol.1, pp. 265-271, 2008.
9. E. Cambria, B. Schuller, Y. Xia, C. Havasi , "New Avenues in Opinion Mining and Sentiment Analysis", *IEEE Intelligent Systems*, vol.28, no. 2, 2013.

10. E. Franchi, A. Poggi, and M. Tomaiuolo, "Open Social Networking for Online Collaboration," *Int. J. e-Collab*, IGI Global Publisher, vol. 9(3), pp- 50-68, 2013.
11. E. Franchi, and M. Tomaiuolo, "Distributed social platforms for confidentiality and resilience," *Social Network Engineering for Secure Web Data and Services*, IGI Global Publisher, pp 114-136, 2013.
12. V. Francisco, P. Gervas, and F. Peinado., "Ontological reasoning to configure emotional voice synthesis," *Procs of Web Reasoning and Rule Systems*, vol. 4524 of LNCS, pp. 88–102. Springer, 2007.
13. A. Go, L. Huang, and R. Bhayani, "Twitter sentiment analysis," Final Projects from CS224N for Spring 2008/2009 at The Stanford Natural Language Processing Group, 2009.
14. S. Hassan, F. Miriam, H. Yulan, and A. Harith, "Evaluation datasets for Twitter sentiment analysis: a survey and a new dataset, the STS-Gold," *Proc. of 1st Int. Workshop on Emotion and Sentiment in Social and Expressive Media: Approaches and Perspectives from AI (ESSEM 2013)*, Turin, Italy, 2013.
15. M. Jacomy, T. Venturini, S. Heymann, M. Bastian, "ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software," *PLoS ONE* 9(6), 2014. doi:10.1371/journal.pone.0098679
16. S. Kiritchenko, X. Zhu, and S. Mohammad, "Sentiment analysis of short informal texts," *J. Artif. Int. Res.*, vol. 50(1), 2014.
17. A Kowcika, A. Gupta, K. Sondhi, N. Shivhre, R. Kumar, "Sentiment Analysis for Social Media", *Int. J. of Advanced Research in Computer Science and Software Engineering*, vol. 3(7), 2013.
18. R. Lambiotte, J.C. Delvenne, M. Barahona, "Laplacian dynamics and multiscale modular structure in networks". arXiv preprint arXiv:0812.1770, 2008.
19. S. Mohammad, X. Zhu, S. Kiritchenko, J. Martin, "Sentiment, emotion, purpose, and style in electoral tweets", *Information Processing & Management*, Elsevier, vol 50 (1), 2014.
20. B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up?: sentiment classification using machine learning techniques". *Procs of the ACL-02 conf. on Empirical methods in natural language processing (EMNLP '02)*, Association for Computational Linguistics, vol. 10, USA, pp. 79-86, 2002.
21. A. Pak and P. Paroubek, "Twitter as a corpus for sentiment analysis and opinion mining," in *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)*. Valletta, Malta: European Language Resources Association (ELRA), May 2010.
22. M. Shams, M. Saffar, A. Shakery, and Faili., "Applying sentiment and social network analysis in user modeling" *Procs of the 13th int. conf. on Computational Linguistics and Intelligent Text Processing – Vol. Part I (CICLing'12)*, Springer-Verlag, Berlin, 2013.
23. Jr. Silla, and A. Freitas, "A survey of hierarchical classification across different application domains," *Data Mining and Knowledge Discovery*, vol. 22(1-2), pp. 31-72, 2011.
24. C. Strapparava and A. Valitutti. "WordNet-Affect: an affective extension of WordNet," *Procs of 4th Int. Conf. on Language Resources and Evaluation (LREC'04)*, vol. 4, pp 1083–1086, 2004.
25. S. Walker, "Salutin' Putin: inside a Russian troll house", *The Guardian*, 2015-04-02.
26. X. Zhu, S. Kiritchenko, and S. Mohammad, "NRC-Canada-2014: Recent improvements in sentiment analysis of tweets," *Procs of the Int. Workshop on Semantic Evaluation*, Dublin, Ireland, 2014.