A Case-Study for Sentiment Analysis on Twitter

Paolo Fornacciari, Monica Mordonini and Michele Tomaiuolo
Dipartimento di Ingegneria dell'Informazione
Università degli Studi di Parma
Parma, Italy

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

Abstract — **Microblogging platforms like Twitter can convey short messages to direct contacts, but also to other potentially interested users. They are actively exploited either by individual users or whole organizations and companies. This paper describes some results we obtained from the Social Network and Sentiment Analysis of a Twitter channel, related to a pop music event. Apart from the particular results, a methodology and some guidelines for the automatic classification of Twitter content are discussed.**

Keywords—Social Network; Sentiment Analysis; Hierarchical Classification

I. Introduction

In the common meaning of the term, an online community (or virtual community) is a group of people interested in a particular topic, or that share some ways of thinking, or that in general have some kind of link that brings them together, with the peculiarity that they interface and connect to each other through a data communication network (such as Internet). In this way, they form a social network with unique characteristics: in fact this combination is not necessarily bound to a physical place and anyone can participate wherever he is, with a simple access to networks.

The social networking sites (SNSs), as defined by Boyd and Ellison in [9], 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. According to Sunden profiles are unique pages where one can "type oneself into being" [32], as the creation of a profile is the minimum condition for joining an SNSs. What makes the SNSs unique is that their purpose is not, in most cases, to allow users to make new friends but the emphasis is on making visible their existing social networks and on the chance to describe them. On the other hand, the specific features of each social network site may depend also on the possible target (social, linguistic or geographic) to which the service is directed. The architecture of social networking platforms is very differentiated. While the most popular platforms are build as essentially centralized systems, other platforms have a distributed architecture [14][15]. The decentralized systems, in particular, often use some notion of trust and cryptography to address the risks of online social networks, which are perceived as serious by many users and have led to incidents [13][35].

Ethnicity, religion, sexual orientation, political beliefs are other factors that have led to the establishment of dedicated social network services, but probably they are also playing an active role in creating and aggregating online communities leveraging the bigger and most popular social networks. This suggests the possibility of new ways to spread information and to influence public opinion. These new scenarios can be better evaluated by a combined observation of the structure and the actual content of the network. This kind of analysis could highlight emerging social behaviors. In [6], for example, the possible differences in the sentiment polarity of female and male users, towards the discussed topic, are examined.

To investigate on the content and on the relations among the actors of a network, it could be useful to contextualize the network itself. In particular, it could be important to consider and inquiry the content of the messages that guide the relationships of the community. It is only through this kind of investigation that we can analyze the semantic meaning of a link, from which we could infer the kind of relationship. This sharpens our description of the social network in many of its facets. A useful tool for such surveys is Sentiment Analysis (SA). SA is a branch of Opinion Mining, that aims to listen and process the data that users post on social media. It is an interdisciplinary field that in recent years has had a significant growth and that makes an extensive use of machine learning techniques. A survey of the main techniques and approaches can be found in [26][7][8]. In [33], it is showed how the information about social relationships can be used to improve user-level sentiment analysis. In [25] Sentiment Analysis is mapped on social media with observations and measurable data; the results highlight the importance of SNSs (i.e. Facebook) as a platform for online marketing.

II. Background

Anthropologist John Barnes was the first to introduce the concept of social network. In 1954 in [5] he described the results of over two years of studies on the composition of classes and social groups in the town of Bremnes (today Bomlo) in Norway. James Mitchel in [22] gave a more sociological and analytical interpretation, describing a social network as "*a specific set of linkages among a defined set of actors, with the additional property that the characteristics of these linkages as a whole may be used to interpret the social behavior of the actors involved*". Mitchel is a representative of the anthropological school of Manchester, formed in the late 40s, whose founders were the first to use the concept of network in a systematic way.

More simply and more generally, in [34] Wasserman and Faust defined a social network as a finite set of actors and the relation or relations defined on them. This approach is characterized by the priority interest turned to the shape of the networks, rather than their content. According to the exponents of this line of research, the form of social relations largely determines their content. This theory (developed since the 70s at Harvard) lays the foundation for social network analysis (SNA). SNA has the objective to model social structures with different properties, starting from the mathematical theory of graphs and the use of matrix algebra [12]. All these definitions could be summarized by arguing that a social network is a group of individuals (actors) which are connected to each other through different types of social links (relationships), such as family ties, employment relationships, superficial knowledge, common interests. With the development of communication technologies and the growth of online communities, the importance of social networks has increased. The research in SNA finds application in analytical and predictive models used in sociology, anthropology, psychology, computer science and economics [29].

One of the most popular social networks is Twitter (https://twitter.com/). At the end of 2012, the company declared in a tweet: «*There are now more than 200M monthly active @twitter users. You are the pulse of the planet. We're grateful for your ongoing support*.» In this short message, the company announced what many researchers in different domains had already noticed: the information and opinions in our society go through a social network where everyone can sign up and participate. So the analysis of this 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, for example in [18]. In [24], the authors describe a recent important application for understanding how public sentiment is shaped, how it could be tracked and its polarization with respect to candidates and issues. Another kind of research in the Twitter social network is to combine data source and sentiment analysis. In [2] geo-spatial information related to tweets is used for estimating happiness in the Italian cities. Twitter is also a microblogging platform, so the techniques used generally in Sentiment Analysis and Text Classification must be adapted to the famous 140-character tweet and this opens the way for new issues. Some example of work in this sector are described in [1][21][20][36]. One of the major problems is how to automatically collect a corpus for Sentiment Analysis and Opinion Mining purposes; see, for example, [28][19].

Sentiment Analysis is traditionally focused on the classification of web comments into positive, neutral, and negative categories. But an intelligent and flexible opinion-mining system has to incorporate a deeper analysis of affective knowledge, and detecting emotions [11]. In [10] the correlation among topics and the positive or negative opinions are investigated, to automatically classify the topics themselves. An ontology driven approach is used in [4] to extract rich emotional semantics of tagged texts, by combining available computational and sentiment lexicons with an ontology of emotional categories. A similar approach can be taken into consideration for the detection of feelings in tweets: for example, a taxonomy of feelings can drive the selection of hashtags for the automatic search of tweets with a prevalent sentiment. Such tweets can be used in the training phase of an automatic classifier.

III. Sentiment Analysis on Twitter

In this research work, we built a system for social network and sentiment analysis, which can operate on Twitter data. Twitter is a popular platform for social networking and microblogging, counting hundreds of millions of active users and daily published messages. As a social networking platform, Twitter is structured as a directed graph, in which each user can choose to follow a number of other users (followees), and can be similarly followed by other users (followers). Thus, the "follow" relationship is asymmetrical, it does not require mandatory acknowledgement, and it is essentially used to receive all public messages published by any followee user. As a microblogging service, Twitter is used to publish short messages counting a maximum of 140 characters (tweets), which may contain opinions, thoughts, facts, references to images and other media. Moreover, through the @ symbol it is possible to introduce mentions, i.e. references to other users, and through the # symbol it is possible to introduce hashtags, i.e. references to discussion topics.

Consequently, in our analysis we collected three types of data. The User type represents users' profiles; from Twitter we obtain the following fields: user_id, name, location, num_followers, num_tweets. The Tweet type represents posted messages; from Twitter we obtain the following fields: tweet_id, user_id, message, date. Finally, the Friend type represents the "follow" relationships among users. Apart from data obtained directly from Twitter, we added a field to both tweets and users, to associate a sentiment with them, according to the result of our analysis.

As a communication medium, tweets have a quite peculiar nature. Some distinguishing features of communication on Twitter are related to technical aspects; those include length of text, tags, urls, etc.. Other features may be classified as idiomatic use of the medium, and create a sort of Twitter culture; those features include typical content and most discussed topics, idiomatic expressions, abbreviated forms, etc. For example, a tweet may have the following form:

«*RT @richman wow this is the #happiest day of my life. #happy #glad #icantbelievit :) :D http://t.co/4VEH827bG7*»

The peculiar nature of tweets requires specialized analysis techniques. As a start, a tweet may contain many elements which are not significant for our classification, and can thus be dropped though a filtering process. To polish the message, we defined various filters, which we have applied in a customizable sequence.

A first filter eliminates useless tokens. Removed tokens include: the starting "*RT*" sequence, which indicates a republished messages from a different user (i.e. a retweet); 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, which may not appear in any other tweet (in the example: "*icantbelievit*"). Ideally, the filter at this level should also support stemming and removal of stopwords. However, those operations can be easily performed by Weka, which we used for analysis.

Finally, another filter separates all punctuation symbols from the text, and organizes them as single-character words. However, some typical patterns are kept as aggregates, including smiles sequences, repeated question and exclamation marks.

The final result of the filtering process is a word vector, which is then submitted to the classifier agents. As we have mentioned, our analysis aims at identifying the following classes of messages: undiscriminated, objective, subjective, positive, negative.

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. If a message fails to be classified at the first stage, then it simply remains undiscriminated. If it fails to be classified at the second stage, then it is marked as generically subjective.

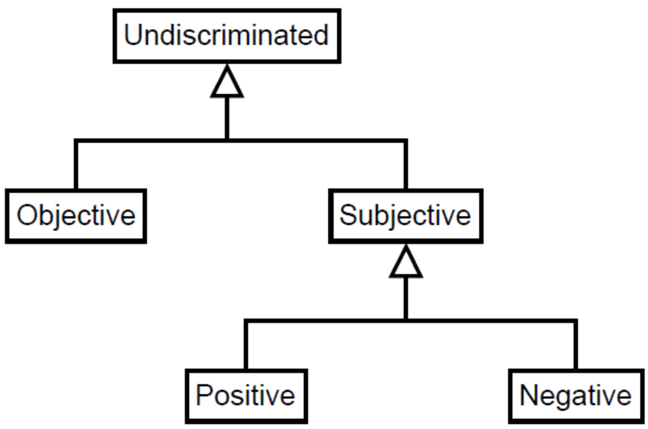

Fig. 1. Hierarchy of basic sentiment classes.

Currently, the classifier agents apply the Multinomial Naive Bayes algorithm, but other methods can be used and different agents can be plugged in the system. However, instead of generating a training set by hand, we aimed at realizing an automated (or at least semiautomated) process for obtaining good training sets.

About the objectivity/subjectivity classifier, we adopted a similar strategy to [27]. 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, @cnnbrk, @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 decided to search for 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 [31], SentiWordNet [3], and other affective lexicons, in a more systematic way.

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. In the next section, we will also discuss the quality of results that can be obtained, using it as a basis for sentiment analysis.

The training set can be provided directly to the classifier agents. In the present form, the system is based on Weka, and can thus be configured for performing additional preprocessing steps on the messages, including common TF-IDF transformations, stemming, elimination of stopwords, exclusion of infrequent words, etc.

Currently, we analyze tweets for discriminating the basic classes of objectivity and polarity, at two levels. However, we designed the system for more complex hierarchical classification, with the application of various types of classifiers, as an alternative to current Naive Bayes.

In fact, hierarchical classification has been applied successfully in a number of studies, for information retrieval [30]. It has been proven effective especially in the case of classification over hierarchical taxonomies. Moreover, it has the advantage of being modular and customizable, with respect to the classifiers used at different levels. Using the same probabilistic classifier and a maximum likelihood estimator, instead, does not provide advantages for the hierarchical approach over the flat approach. Mitchell [23] has proved that the same feature sets represent documents in both approaches. Consequently, the whole hierarchical classifier system is equivalent to the corresponding flat system.

Also in the case of sentiment analysis, a hierarchy of classes can be defined [16][4]. Accordingly, hierarchical classification has already been applied to sentiment analysis, too [17].

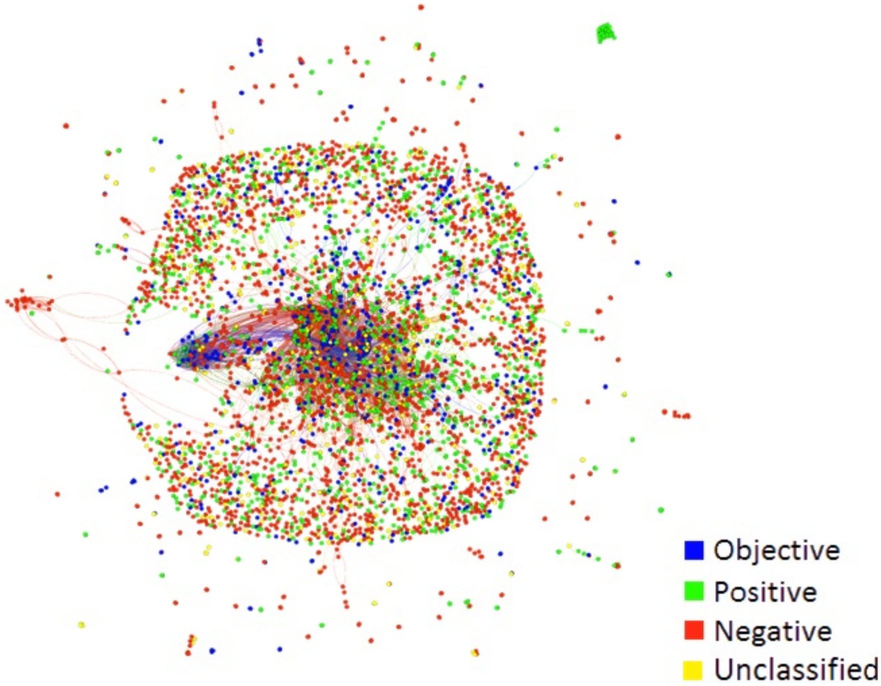

Fig. 2. Communities participating in the *#SamSmith* channel.

IV. A Case-Study: the *#SamSmith* Channel

This section will show the results of the classifiers and the analysis carried out on a case study.

With the above described software, it is possible to obtain some training sets for the classifiers. In our case study, 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 obtained results 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 network (shown in Fig. 2) consists of a total of 5570 nodes and 6886 arcs.

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. 3, the prevailing sentiment detected from the classifier is the negative one. Performing an analysis on a

NEG	49,84%
POS	26,52%
OBJ	13,41%
NC	10,23%

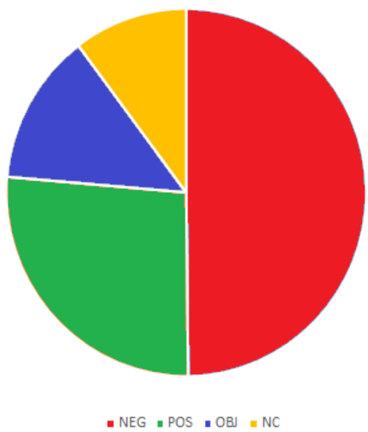

Fig. 3. Sentiment analysis on the *#SamSmith* channel.

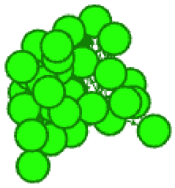

Fig. 4. A small community, showing positive sentiment.

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.

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, it is possible to notice a small group of users whose feeling is completely positive (Fig. 4). 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.

Another kind of analysis we made concerns with the grade of the users. Fig. 5 shows that two nodes have a key role within the social network:

- *@samsmithworld*
- *@TheGRAMMYs*

These users are the main sources of news about the singer Sam Smith and the event Grammy Awards 2015. This explains their importance within the social network which we considered.

V. Conclusion

In this article, we describe some results obtained from the synthesis of Social Network Analysis and Sentiment Analysis applied to the channel *#SamSmith* during the Grammy Awards in 2015. 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 topology of the networks.

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).

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.

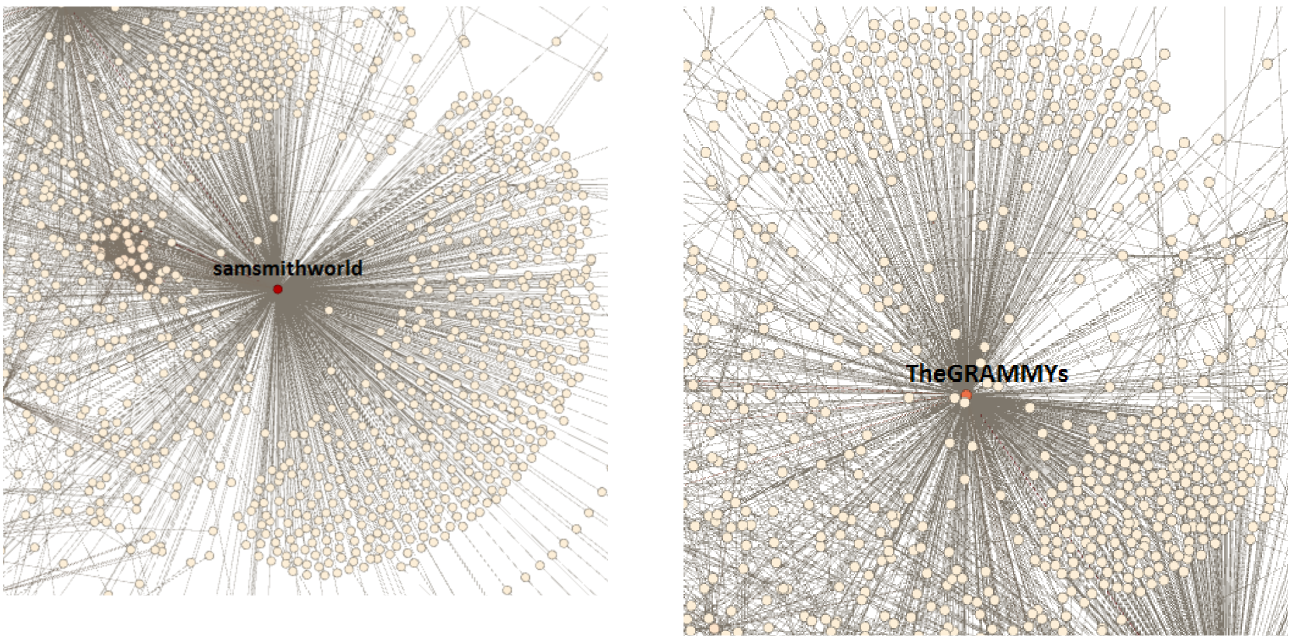

Fig. 5. The most followed nodes in the *#SamSmith* channel.

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), 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 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] M. Baldoni, C. Baroglio, V. Patti, and P. Rena, P. "From tags to emotions: Ontology-driven sentiment analysis in the social semantic web," Intelligenza Artificiale, vol. 6(1), pp. 41-54, 2012.

[5] J. A. Barnes, "Class and Committees in a Norwegian Island Parish, Human Relations", Human Relations, vol 7(1), pp. 39-58, 1954.

[6] A. Bermingham, M. Conway, L. McInerney., N. O'Hare, and F. Smeaton, "Combining Social Network Analysis and Sentiment Analysis to Explore the Potential for Online Radicalisation", in Procs of the 2009 Int. Conf. on Advances in Social Network Analysis and Mining (ASONAM '09). IEEE Computer Society, pp 231-236, 2009.

[7] L. Bing, "Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data," Data-Centric Systems and Applications, 2nd ed., Springer, pp. 1-603, 2011.

[8] L. Bing. Sentiment Analysis and Opinion Mining, Morgan & Claypool Publishers, 2012.

[9] D. Boyd, N. Ellison, "Social Network Sites: Definition, History and Scholarship," Journal of Computed-Mediated Communication, vol. 13 (1), pp. 210-230, 2008.

[10] 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.

[11] E. Cambria, B. Schuller, Y. Xia, C. Havasi , "New Avenues in Opinion Mining and Sentiment Analysis," IEEE Intelligent Systems, vol.28, no. 2, 2013.

[12] P. Carrington, J. Scott, S. Wasserman, Models and Methods in Social Network Analysis, Cambridge University Press, 2005.

[13] 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.2014070104.

[14] E. Franchi, A. Poggi, and M. Tomaiuolo, "Open Social Networking for Online Collaboration," International Journal of e-Collaboration (IJeC), 9(3), 50-68, 2013. doi:10.4018/JITR.2015010103.

[15] 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. doi:10.4018/978-1-4666-3926-3.ch006.

[16] 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.

[17] D. Ghazi, D. Inkpen, and S. Szpakowicz, "Hierarchical approach to emotion recognition and classification in texts," Advances in Artificial Intelligence, vol. 6085 of LNCS, Springer Berlin, pp 40-50 2010.

[18] 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. [Online].

[19] 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 ESSEM 2013, Turin, Italy, 2013.

[20] S. Kiritchenko, X. Zhu, and S. Mohammad, "Sentiment analysis of short informal texts," J. Artif. Int. Res., vol. 50(1), 2014.

[21] 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.

[22] J.C. Mitchell, Social Networks in Urban Situations: Analysis of Personal Relationships in Central African Towns, Manchester Univ. Press, 1969.

[23] T. Mitchell, "Conditions for the Equivalence of Hierarchical and Flat Bayesian Classifiers". Technical report, Center for Automated Learning and Discovery, Carnegie- Mellon University, 1998.

[24] S. Mohammad, X. Zhu, S. Kiritchenko, J. Martin, "Sentiment, emotion, purpose, and style in electoral tweets", Information Processing & Management, Elsevier, vol 50 (1), 2014.

[25] F. Neri, C. Aliprandi, F. Capeci, M. Cuadros., and T. By, "Sentiment Analysis on Social Media", Proces of the 2012 Int. Conf. on Advances in Social Networks Analysis and Mining (ASONAM '12), IEEE Computer Society, pp: 919-926, 2012.

[26] B. Pang, and L. Lee, "Opinion Mining and Sentiment Analysis," Found Trends Inf. Retr., vol. 2 (1-2), pp. 1-135, 2008.

[27] 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), vol. 10, USA, pp. 79-86, 2002.

[28] A. Pak and P. Paroubek, "Twitter as a corpus for sentiment analysis and opinion mining," in Proc. of the 7th conf. on International Language Resources and Evaluation (LREC'10). Valletta, Malta, May 2010.

[29] J. Scott, Social Network Analysis, 3rd ed., London and Beverley Hills Sage Publications, 2012.

[30] 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.

[31] 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.

[32] J. Sunden, Material Virtualities. Approaching Online Textual Embodiement, New York: Peter Lang., 2003.

[33] C. Tan, L. Lee, J. Tang, L. Jiang, M. Zhou, and P. Li, "User-level sentiment analysis incorporating social networks", Procs of the 17th ACM SIGKDD Int.. Conf. on Knowledge Discovery and Data Mining (KDD '11), ACM, pp 1397-1405. 2011.

[34] S. Wasserman, K. Faust, Social Network Analysis: Methods and Applications, Cambridge University Press, 1994.

[35] M. Tomaiuolo, "dDelega: Trust Management for Web Services," International Journal of Information Security and Privacy (IJISP), 7(3), 53-67, 2013. doi:10.4018/jisp.2013070104

[36] 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.