Content-Based Visualization System For Sentiment Analysis On Social Networks

Julio Vizcarra[1], Kouji Kozaki[1,*],Miguel Torres Ruiz[2],Rolando Quintero[2]

[1] The Institute of Scientific and Industrial Research (ISIR) Osaka University Mihogaoka 8-1, Ibaraki, Osaka 567-0047. Japan.
[2] Centro de Investigación en Computación CIC , Instituto Politécnico Nacional, UPALM-Zacatenco, CIC. Building, 07738, Mexico City, Mexico.

Abstract. A content-based visualization system is presented for the sentiment analysis on social networks. The methodology implemented was focused on the semantic processing taking into account the content in the public user's opinions. In our approach the comments were handled as excerpts of knowledge. During the visualization the social graph is displayed presenting the polarity and sentiment status for each comment. Moreover a web mapping tool retrieves comments in a radius based on the location source(geographic) or concepts related to geographic entities and spatial relations in the comment(conceptual).

Keywords: sentiment analysis, knowledge engineering, conceptual similarity

1 Introduction

Nowadays the huge information transmitted on social networks has become a rich source of information for the human understanding as well as a way of expression where the users share their sentiment status and personal opinions through comments. The sentiment identification can classify comments as positive or negative(polarity) and unveil emotions such as anger, trust, sadness ,etc., on certain topics or users. Moreover the sentiments presented in the opinions can be relevant in the design of custom services, social plans for public health, marketing, e-commerce,etc. On the basis of these motivations, we developed a web system for real-time monitoring of sentiment information in social networks for specific targets(public events or users). The system is able to display the social graph structure, sentiment information related as well as retrieve comments by source's location or words geo-referenced in the text by means of a web mapping.

2 Methodology

The methodology implemented handles the comments as excerpt of the knowledge, in this gap we prioritized the semantic level, sense and meaning of the whole

* Corresponding author.
E-mail address: kozaki@ei.sanken.osaka-u.ac.jp.

comment. The proposal computed semantic similarity measures, conceptual expansion, graph theory algorithms and disambiguation using a multi domain knowledge base. The methodology is composed for the following stages: the stage of "social networks discovery" retrieves opinions from events or public profiles by reading their comments. Consequently the stage of "knowledge processing" constructs the formal representation for each comment by using an knowledge base and graph algorithms such as steiner tree [6] and shortest paths[2]. This module carries out processes of automatic knowledge graph construction enhanced by disambiguation. For processing general knowledge on specific languages we considered the English, Japanese and Multilingual Wordnets[1][4][5]. In addition the dictionaries that provided sentiment information were SentiWordnet [3] and NRC emotion lexicon [7]. On the other hand the online services for geographic information processing were WikiData and GeoNames. Finally the stage of "sentiment analysis" estimates the total polarity and main sentiment presented in the comments. For each concept the polarity is obtained from the knowledge base by average and the closets sentiment by knowledge graph expansion and shortest path. The highest polarity and main sentiment are established to a comment.

3 Visualization

This section describes the system that implemented our methodology. We present examples of the sentiment analysis and their visualization.

In order to explain the sentiment analysis an example was processed from Twitter in the CNN News account. The comment considered is: "a number of people feared dead after a dam bursts in kenya with hundreds left homeless officials say". The table 1 presents the closest sentiment and a polarity value assigned by our methodology to each concept.

Id Wordnet-Concept	Sentiment with polarity
WN:107449542-n ("flare","burst")	Sentiment:NRC_fear_NRC_anger|:Polarity:-0.25 ,
WN:107964900-n (homeless)	Sentiment:NRC_anticipation_disgust_anger|Polarity:-0.125 ,
WN:107534492-n (fear)	Sentiment:NRC_fear,sadness,anger,surprise|Polarity:-0.875 ,
WN:114509110-n (say)	NRC_surprise_anticipation|Polarity:0.5

Table 1. Sentiment-Polarity assigned to concepts

Finally the methodology estimates the total polarity and main sentiment presented in the comment. The values established were for polarity: -0.1875 and main sentiment: NRC_Anger.

Additionally some relevant results from Twitter account CNN News are presented. The table 2 defines the main sentiment and polarity value assigned by our methodology to the comments. We noticed a better and more trustworthy classification using the basic sentiments instead of polarity (average).

Sentiment	Polarity	Comment
trust	0.2916667	This couple found a buried safe containing $52,000 worth of money, gold and jewelry in their backyard, but didn't keep it
trust	-0.15	In an effort to keep conversations and search results on topic, Twitter announced it will use new "behavioral signals" to push down more tweets that "distort and detract"
anger	0.04166667	A massive poaching ring in Oregon and Washington is accused of killing more than 200 animals including deer, bears, cougars, bobcats and a squirrel
anger	0.041666687	An estimated 239,000 girls under the age of five die in India each year due to neglect linked to gender discrimination, a new study finds
sadness	0.25	@CNN Her father had a heart surgery and cant walk so
sadness	-0.25	Teen develops 'wet lung' after vaping for just 3 weeks
joy	0.125	I am proud to be a woman and a feminist. The politics of Meghan Markle

Table 2. Other examples processed in twitter

In the visualization the results are displayed in the system by means social graphs and web mapping. Regarding the social graph it describes the network's structure and its sentiment information related to comments by colors. For instance the figure 1 presents the polarity and sentiment graphs for the CNN news account. Regarding the nodes the darkest blue represents the user target and light blue for users farther. Particularly in the polarity graph the nodes with gray color represents neutral comments and the scale between green and red for positive to negative polarities respectively. On the other hand in the sentiment graph each comment has a sentiment represented by a different color.

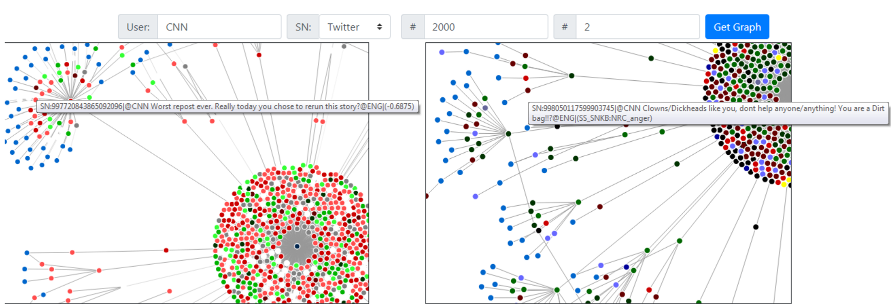

Fig. 1. polarity and sentiment graphs

In addition the web mapping tool retrieves comments by location which can be geographic(location source) or conceptual(Geo-referenced concepts and relations), The figure 2 retrieves comments by conceptual processing using the keyword “Arkansas” in a distance of 1000 km. The comment that contains in its description the concept “memphis” is retrieved.

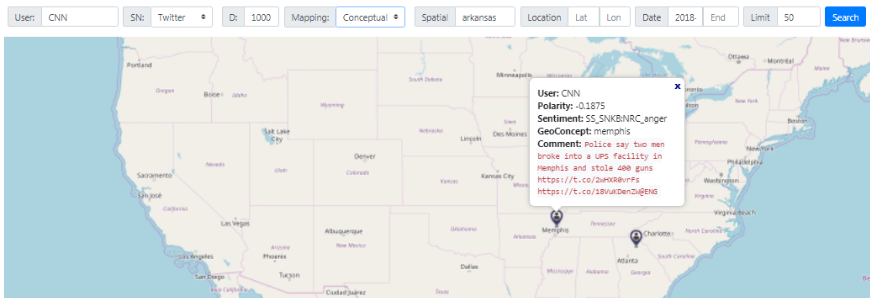

Fig. 2. web mapping

4 Conclusions

In this poster a content-based methodology and its implementation were proposed for the sentiment identification. The novelty of the presented our approach is the capability of handling the comments as excerpts of knowledge. We provided a mechanism of semantic processing using knowledge graphs, graph theory algorithms, semantic similarities and disambiguation. Our implementation can be a relevant tool for studying the impact of events and users in the society. Moreover the sentiment analysis in social networks can contribute in the public health and design of custom services.

5 Acknowledgments

This work was supported by CONACYT and JSPS KAKENHI Grant Number JP17H01789.

References

1. Princeton university "about wordnet." wordnet. princeton university (2010), <http://wordnet.princeton.edu>
2. Ahuja, R.K., Mehlhorn, K., Orlin, J., Tarjan, R.E.: Faster algorithms for the shortest path problem. Journal of the ACM (JACM) **37**(2), 213–223 (1990)
3. Baccianella, S., Esuli, A., Sebastiani, F.: Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In: LREC. vol. 10, pp. 2200–2204 (2010)
4. Bond, F., Baldwin, T., Fothergill, R., Uchimoto, K.: Japanese semcor: A sense-tagged corpus of japanese. In: Proceedings of the 6th Global WordNet Conference (GWC 2012). pp. 56–63 (2012)
5. Bond, F., Foster, R.: Linking and extending an open multilingual wordnet. In: Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). vol. 1, pp. 1352–1362 (2013)
6. Hwang, F., Richards, D.S.: Steiner tree problems. Networks **22**(1), 55–89 (1992)
7. Mohammad, S.M., Turney, P.D.: Crowdsourcing a word–emotion association lexicon. Computational Intelligence **29**(3), 436–465 (2013)