



The Use of Social Networks for Sentiment Analysis

Trends in public opinion and changes in social movement strength are becoming easier to detect and monitor

During the last four years, social networking has been used to influence political and social change in various parts of the world. The spreading use of digital social networking went through North Africa, the Middle-East, UK, France, and even Montreal. Governments and corporations alike started to feel the need to measure factors influencing social mode that are commonly hard to capture by computer algorithms. Analysis of emotional temperature by monitoring Facebook and Twitter postings or by analyzing mobile phone traffic between particular groups requires different types of algorithms compared to typical computer science analytical approaches.

Sentiment Sensing

One classical approach is to predict the inevitable clash of people in drought situations for example. Dr. Revilla, a professor of computational social sciences at George-Mason University of Virginia, has developed a tool called Riftland for the US Navy [1,2]. Riftland focuses on the area in east Africa that includes Somalia, Ethiopia, Rwanda, Uganda, and Congo, an area that has been the scene for civil unrest for quite some time. Riftland collects data from academics, government agencies, and even charitable organizations, to identify where nomadic tribes will go and whom they may clash with at times of drought or unrest.

For instance, the Riftland model would expect tribes that own radios or mobile phones to steer away from roads during war or civil unrest, but do the opposite during drought season. With a little help from satellite data combined with a good knowledge base of tribal hostility, a quantitative system of tribal favours can be developed. Favours can be exchanged at times of austerity. One can imagine that Dr. Revilla's tool can be tweaked to

Yasser Morgan and Abdul Bais
University of Regina

model other areas of the world. However, the force binding tribal members to a common behavior is much stronger than the force binding young men on the streets of Paris for example.

Since modern uprisings have used Twitter and Facebook, it makes sense to dig through both. Germany's largest political party, the Christian Democratic Union consulted Dr. Gloor of the Massachusetts Institute of Technology to apply data mining techniques on Facebook and Twitter aiming at sentiment analysis. The bigger objective is to evaluate and predict how and when protests will evolve. Dr. Gloor developed a tool called Condor that classifies participants by their clout [3,4]. For instance, an influential twitter user is the one who has many users, but follows few people. If an algorithm similar to Google's page ranking approach is applied, we can analyze Twitter users by ranking them. An interesting finding is the language used by top-ranked users which includes bold phrases like "great", "hilarious", or "disgusting". Bold language sways followers. While positive optimistic tweets heat-up a social movement, negative tweets can be used as an indication of diminishing esteem and burn-out. If tweets like idiots, lame, and never are coming from lead users, an easy prediction of dissolving protests can be made. If similar negative tweets are exchanged within the same clout, it is a sure indication of a dead movement.

Measures used by the Condor system cannot be extended across cultures due to the different use of language and little has been done to identify

whether the positive and negative phrase indicators can be used across cultures or the entire process needs to be rethought [5]. Further, the Condor sensing of public sentiment is a slow detector. By the time influential users are using bold phrases, the demonstration is probably on its way to the street. Yet Condor remains a good tool for shorter term predictions and to monitor the evolution of existing protests.

E-MEME (Epidemiological Modeling of the Evolution of MESSAGES) is another sentiment analysis tool that expands the focus to the entire population [6]. To achieve that, data are extracted from the news to evaluate the predisposition of particular parts of the populace to certain ideas. E-MEME would have predicted the public outrage in Egypt over a border incident with Israel and could predict the effect of drought on different parts of a state.

A far more complex system developed by Lockheed-Martin is called W-ICEWS (World Integrated Crises Early Warning System) [7,8]. W-ICEWS is a social radar tool that can crunch great quantities of data from digital news media, blogs and websites, in addition to intelligence and diplomatic reports providing W-ICEWS with the ability to forecast riots, rebellions, coups, economic crises, government crackdowns and international wars months in advance [8].

An interesting feature of W-ICEWS is its ability to entertain what-if scenarios. By that, it provides the American government with a tool to influence the future, not only predict it, based on educated guesses. While, the wide multitude of influencing parameters can lead events to evolve differently than expected, W-ICEWS is our best political sentiment analysis tool developed yet.

Current systems are still vulnerable to opinion spam. User diversion from common language, the use of sarcastic language, and the spontaneous creation of new language terms have been hard to detect and analyze.

Perhaps the best news in the area is the launching of the IARPA (Intelligence Advanced Research Projects Activity) almost a year ago [9]. IARPA is an open-source application that has the potential to change our research landscape more than any other tool. The flexibility of IARPA allows programmers to customize it to analyze anything whether it is the roadside traffic cameras or the particular Wikipedia search pages that are growing in popularity, it provides social and computing researchers with newer tools in a way that was never available before.

Business Intelligence

Business corporations have been trying to extract intelligence from social networks with multiple objectives. The most direct objective is to improve business links to customers or to develop increasing loyalty as well as discovering new customers by focusing on the right demographic. But perhaps the most vital is the discovery of new business opportunities. Ambitious research is evolving even towards predicting the market value of a particular stock by analyzing the collective sentiment on that stock on the web.

In 2010 Dr. Chen from the University of Arizona analyzed the sentiment towards WalMart stock between 1999 and 2008 [10]. He presented correlation between growing disagreement and decreasing objectivity on one side and volatility on the other. Similarly low disagreements and high levels of objectivity led to higher stock volatility. In particular, sentiments of disagreement were followed by an increase in trading volume while higher levels of objectivity indicated suppressed trading volume. This analysis shows the importance of breaking down sentiment into discrete indicators. Surprisingly, efforts to influence forum participants were followed by a more conservative market. Shareholders were more likely to hold, buyers were more likely to postpone leading to a more conservative market. Generally, positive sentiments on the Web forum were correlated to lower trading volume while negative sentiment induces trading activities as shareholders defect.

Another interesting application is the prediction of movie sales by mining online reviews since these reviews, collectively, reflect the wisdom-of-crowds. It is important to realize that sentiment analysis goes beyond analyzing the volume reviews or the relevant links structures. Yu et.al presented a model that used probabilistic analysis

for the reviews sentiment combined with a quality predictions model of reviews in the absence of readily available quality indicators [11]. It is possible to improve the outcome by focusing on the actionable knowledge. In all cases predicting quantifiable sales performance can be based on an artificial learning algorithm.

Conclusion

While sentiment analysis has gone beyond experimental levels into providing tangible outcomes, current research falls short of satisfying the ambitious expectations. For example, on opinion classification, there are many conceptual rules that measure opinions, and there are even more expressions (possibly unlimited) that can convey these concepts. This leads to the question of sentiment analysis accuracy of the current algorithms. To improve accuracy, systems need to include fairly large amounts of data to confirm the compliance to a certain prediction rule. But to benchmark the same rule, data needs to cover a large number of domains because a system that does well in one domain might not do well in another, as opinions in different domains can be expressed so differently.

Finally, we need to acknowledge that a completely automated system is nowhere in sight. Current systems are still vulnerable to opinion spam. User diversion from common language, the use of sarcastic language, and the spontaneous creation of new language terms have been hard to detect and analyze. Consequently, sentiment analysis systems are expected to remain captive to manual intervention, for the foreseeable future at least. Due to the real substantial need imposed by governments and industry alike, sentiment analysis is poised to remain vibrant and alive for years to come.

References:

- [1] C-R. Claudio et al., "MASON RiftLand: An Agent-Based Model for Analyzing Conflict, Disasters, and Humanitarian Crises in East Africa," in Agent-Based Models and Geographical Systems: Policy Modelling, 57th Annual Meeting of the Association of American Geographers, New York, 2012.
- [2] T. Gulden, J. F. Harrison, and A. T. Crooks, "Modeling Cities and Displacement through an Agent-based Spatial Interaction Model," in Proceedings of the Annual Conference of Computational Social Science Society of America, Santa Fe, NM, 2011.
- [3] Intelligent Collaborative Knowledge Networks (ICKN). (2012, June) Condor. [Online]. <http://www.ickn.org/condor.html>

- [4] F. Grippaa, M. Palazzolo, J. Bucuvalas, and Peter G., "Monitoring Changes in the Social Network Structure of Clinical Care Teams Resulting from Team Development Efforts," in Proceedings of the 3rd International Conference on Collaborative Innovation Networks (COINs), Basel, Switzerland, 2011.
- [5] Hui-Hsin Wu, A.C.-R. Tsai, R.T.-H. Tsai, and J.Y.-J. Hsu, "Sentiment Value Propagation for an Integral Sentiment Dictionary Based on Commonsense Knowledge," in Proceedings of the International Conference on Technologies and Applications of Artificial Intelligence, 2011, pp. 75-81.
- [6] APTIMA. (2012, June) Aptima Develops E-MEME (the Epidemiological Modeling of the Evolution of MESSAGES) to Track Infectiousness of Ideas across Groups and Geography. [Online]. <http://www.aptima.com/news/2012/aptima-develops-e-meme-track-infectiousness-ideas-across-groups-and-geography>
- [7] DARPA. (2012, June) INTEGRATED CRISIS EARLY WARNING SYSTEM (ICEWS). [Online]. [http://www.darpa.mil/Our_Work/I2O/Programs/Integrated_Crisis_Early_Warning_System_\(ICEWS\).aspx](http://www.darpa.mil/Our_Work/I2O/Programs/Integrated_Crisis_Early_Warning_System_(ICEWS).aspx)
- [8] The Economist. (2012, April) What makes heroic strife. [Online]. http://www.economist.com/node/21553006?goback=%2Egde_2550249_member_110608062
- [9] IARPA. (2012, June) Intelligence Advanced Research Projects Activity. [Online]. <http://www.iarpa.gov/organization.html>
- [10] H. Chen and D. Zimbra, "AI and Opinion Mining," IEEE Intelligent Systems, vol. 25, no. 3, pp. 74-80, May-June 2010.
- [11] X. Yu, Y. Liu, J. X- Huang, and A. An, "Mining Online Reviews for Predicting Sales Performance: A Case Study in the Movie Domain," IEEE Transactions on Knowledge and Data Engineering, vol. 24, no. 4, pp. 720-734, April 2012. ■

About the Authors

Yasser Morgan is an Associate Professor of Software Systems Engineering in the Faculty of Engineering and Applied Science at the University of Regina. He received both his Bachelor of Science and Masters of Science degree from the University of Cairo and earned his Ph. D at Carleton University in Ottawa. His current research spans various multi-disciplinary and inter-disciplinary domains in the area of mobile computing with a particular focus on vehicular mobility.

Abdul Bais is currently an Instructor in at the University of Regina in the Faculty of Electronic Systems Engineering having recently completed his post doctoral fellowship. He received his PhD from the University of Vienna and his BSc and MSc from the University of Peshawar, Pakistan.