Forecasting Social Movements By Web Buzz Analysis

Peter A. Gloor
MIT Center for Collective Intelligence, Cambridge MA 02139
pgloor@mit.edu

1. INTRODUCTION

Widespread Internet usage has led to the availability of a cacophony of fact, fiction and opinion composed by all elements of society; events such as the recent panic in Veracruz, Mexico over Twitter chatter of a violent attack on a school demonstrates the societal power of this cacophony. Governments, news organizations and most importantly individuals produce data that represents a stream of consciousness of society, albeit a stream with many competing voices, agendas, and noise. Embedded in this stream are the opinions and activities that will move society and make news.

2. COOLHUNTING WITH CONDOR

Our approach draws upon and combines tools from computer science and social sciences to accurately gather collective opinion and topic information from wide ranges of open source indicators and tracks influence through groups online to create robust early warnings. In particular, we integrate the innovative dimensions outlined below.

First, we draw upon the proven Condor Architecture [1], a flexible system for predicting societal trends based on publicly available information sources on the Internet. The flexibility of the system enables the early warning of various significant societal events. Condor is able to predict events of the type where many people express ideas and then have a greater propensity to act upon them.

Second, our Coolhunting approach [2] detects emerging group identities. Our approach is supported by key findings in social psychology, neuroscience, and the sociology of social movements. This enables us to track existing social groups, and infer the formation of new groups, in a manner that is generalizable across cultures and that does not require a priori identification of political or other affiliations.

Third, we find trends by finding the trendsetters [2]. The key differentiator of our approach is the use of social networking techniques to determine those actors that are most influential. Our approach weighs every instance of a search term with the social networking position of the actor using the search term.

Fourth, we use implicit data to interpret "honest signals"[3]. A core concept of our philosophy is that the metadata surrounding the explicit data associated with a posting (e.g. author, date, text) is as important as the data that expresses the hidden intent of the author, frequently without him/her actually being aware of it. For example, Wikipedia has a great amount of hidden, latent value in the metadata surrounding these content pages: the authors (and what else they've authored), the course of activity (edits/time), the location of links in the content, the networks of links and authorship [4].

Finally, we segment the Web data into three information spheres: Crowds, Experts, and Swarms. Condor predicts societal trends based on Web buzz time series captured through identifying and weighing the appropriate data source in the three different information spheres. For example, the crowd's output is collected by analyzing the Web through the Google/Bing/Yahoo search indices and Twitter, the experts' output is collected through Wikipedia and scientific records such as PubMed, and the swarm is found in online forums such as IMDB or Yahoo Finance

HotSocial'12, August 12, 2012, Beijing, China.

3. EXAMPLES

In a series of experiments we compared public sentiment measured through Twitter with the stock market [5]. In a recent example we were able to predict the drop in Goldman Sachs' stock price triggered by Greg Smith's March 14 article published by the New York Times, "Why I am Leaving Goldman Sachs." Twitter buzz preceded the appearance of the actual article by a few hours, the stock price dropped as soon as trading started that morning.

In another project [6] we have examined the relationship between the public mood measured through Twitter associated with the economies of Italy, Spain and Greece, and prices of Credit Default Swaps on sovereign bonds of these countries. The effect of the changes in the public mood on CDS was measured by Granger causality tests and linear regression models. Results of the Granger tests suggest that constructed mood indices convey new and meaningful information about changes in CDS prices. Moreover, the extent to which this is true varies between countries. In the analyzed timeframe, Twitter mood is a much better predictor for Spain than for Italy – there is a strong relationship between the mood associated with Spain and changes in the Italian CDS prices. This empirical evidence illustrates the spillover effect that troubles in one country might have on another country.

4. REFERENCES

[1] Gloor, P. Zhao, Y. TeCFlow - A Temporal Communication Flow Visualizer for Social Networks Analysis , Workshop on Social Networks. ACM CSCW, Chicago, Nov. 6. 2004.

[2] Gloor, P. Cooper, S. Coolhunting - Chasing Down The Next Big Thing? AMACOM, NY, 2007

[3] Gloor, P. Fischbach, K. Fuehres, H. Lassenius, C. Niinimäki, T. Olguin Olguin, D. Pentland, A. Piri, A. Putzke, J. Towards "Honest Signals" of Creativity – Identifying Personality Characteristics Through Microscopic Social Network Analysis. Procedia - Social and Behavioral Sciences, Volume 26, 2011, 166-179.

[4] Kleeb, R. Gloor, P. Nemoto, K. Henninger, M. Wikimaps: Dynamic Maps of Knowledge. International Journal of Organisational Design and Engineering, Vol. 2, No. 2, 2012.

[5] Zhang, X. Fuehres, H. Gloor, P. Predicting Asset Value Through Twitter Buzz. Altmann, J. Baumöl, U. Krämer, B. (eds) Proceedings 2nd. Symposium On Collective Intelligence Collin 2011, June 9-10, Seoul, Springer vol. 112, 2011

[6] Chekanskiy, S. Collective Intelligence in Financial Markets: Does consumer sentiment influence valuation of financial products? Sloan Master of Science in Management Thesis, 2012.