Monday, 6 February 2017

Predictive Analytics in Social Media (Master Thesis Proposal)

Millions of people around the world perceive social media as alternative social environments, where they can find latest news, share experience and manifest their political identity. Beyond that, socially important, collective events, for example protests or social riots, become manifest in these media by huge communication activities. Figure 1 shows an exemplary Twitter communication recorded during the Euromaidan protests in Kiev in 2014. Each of the distinctive peaks represents a protest or civil unrest event.

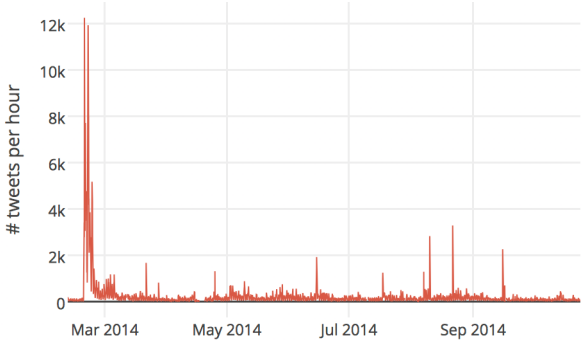

Figure 1: Twitter communication during the Euromaidan protests in Kiev.

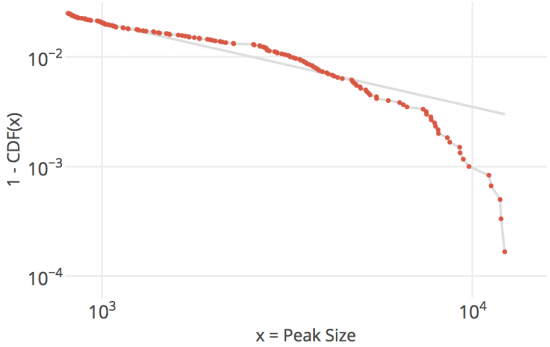

Figure 2: The complementary cumulative distribution function [1-CDF(x)] for peak sizes x during the period of the Euromaidan protests.

The proposed thesis will focus on the predictability of peak loads in social media communication. The distribution of peak sizes, shown in Figure 2, suggests that large peak events follow a different distribution than their smaller counterparts. The straight line, representing a power law or Pareto distribution in the log-log plot of Figure 2, is a good approximation for peak sizes up to about 8’000 tweets per hour. Larger peaks clearly deviate from this distribution, suggesting that large events are driven by different dynamics.

Here, we suppose that large outbreaks in the communication activity experience a transition from a power law regime towards a more coherent dynamics known as dragon-king. With respect to predictability this transition is very important, since with power law dynamics, indicating a regime of Self-Organized Criticality (SOC), extreme events are not different from their smaller siblings, making the former unpredictable. This is due to the scale invariance of power laws. In the dragon-king regime, by contrast, predictability returns.

In a preliminary study the large peaks from the time series shown in Figure 1 actually reveal Early Warning Signals (EWS) several hours before the events escalate and manifest as large social riots. The early warning signals consist of i) an increased variance, an ii) increased autocorrelation and iii) a changing skewness of the values (i.e. tweets per hour) measured within a sliding window. These EWS are widely known as indicators of an imminent, sudden transition in tipping point models, also known as catastrophic bifurcations. Most interestingly, the same system features that favour the tipping point dynamics, i) an increased coupling

strength between and ii) a reduced heterogeneity among interacting nodes, also favour the transition of a system dynamics into the dragon-king regime, mentioned above.

armasuisse Science + Technology is interested in these liaisons between dragon-king and tipping point dynamics and their implications for predictability. This should be the major topic of the thesis. We think that social networks, and in particular Twitter communication networks, offer a very nice opportunity to define and trace fundamental features such as coupling strength between and heterogeneity among interacting nodes, which are able to change the dynamics of systems profoundly.

Based on the results of the thesis we might be able to tackle a bunch of new, prevailing questions. For example, the presently hotly debated (Twitter) bots: what roles could they assume in changing the system dynamics and potentially the predictability? In the tweet collection used for the preliminary study in the context of the Euromaidan protests, it is known that up to 20% of all the tweets originate from bots.

This master thesis can be conducted at the *ETH Zürich* or in the research labs of *armasuisse Science + Technology* in Thun. In both cases, the student will get access to a large body of tweets in a computing cluster to conduct experiments with real-world events.

Contact at armasuisse S+T:

Dr. Albert Blarer
Scientific Project Manager, C4I

armasuisse Science + Technology
Feuerwerkerstrasse 39
CH-3602 Thun

Tel.: +41 (0)58 468 27 24
Mobile: +41 (0)79 568 64 31
E-Mail: albert.blarer@armasuisse.ch
Web: www.armasuisse.ch/wt