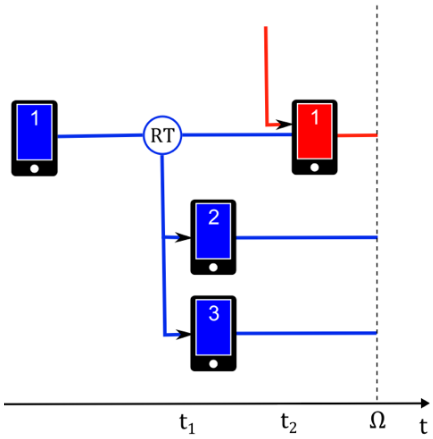
Kevin Ryczko
100386879
February 12, 2014

Topics in Contemporary Physics - Writing Project 2

Abstract

J. Ward *et al.* have derived simple mechanistic models to describe the popularity of memes (tweets) in an online social network environment. Heavy tailed (power law) popularity distributions emerge from these models, indicating that some memes become extraordinarily popular, while the majority of the memes do not. This phenomena can be found in real life online social networks and examples include apps on Facebook, or hashtags on Twitter. The topologies of the popularity models agree well with the topology found in experimental data obtained from Twitter, indicating that the popularity of hashtags or retweeted URLs may follow the simple model predicted by J. Ward *et al.*

Modelling tweet popularity through a simple mechanistic model



While working at the University of Limerick, J. Ward *et al.* derived simple mechanistic models to describe the popularity of memes (tweets) in an online social network like Twitter. The first model included only 2 memes and no innovation (i.e. the users in the network can not produce their own meme), and the second model better resembled a realistic online social network by allowing innovation and multiple initial memes. The popularity distributions produced by these models are power law distributions, indicating that only a very few memes become very popular while the majority of the memes do not gain any popularity. "It is very interesting to see that such simple models produce such accurate descriptions of real life popularity phenomena", says Ward in an exclusive interview.

The inflation of massive online social networks has sparked a recent interest in the physics community to understand the topology and dynamics of such networks; popularity distributions obtained by J. Ward *et al.* give a keen insight into how items in these complex networks become so popular so quickly. "When people select from multiple items of roughly equal value, some items quickly become extremely popular, while other items are chosen by relatively few people", said J. Ward *et al.* This same phenomena can be found in online social networks through apps on Facebook, or retweeted URLs and hashtags on Twitter. The popularity distribution of the world wide web (or degree distribution) defined by the degrees of webpages (the number of connections a certain web page has to other web pages) has been shown to be a power law distribution with an exponent of 2. The topology of the preferential attachment model, which correctly resembles the topology of the world wide web, is similar to the topology given by the models used by J. Ward *et al.*, but the popularity distributions found give a power law exponent less than 2. The popularity distributions in the examples mentioned earlier regarding Facebook and Twitter give a power law exponent of approximately 1.5, and interestingly this is the same power law exponent for previously studied self organized criticality models (an example of a self organized criticality model is the evolution and extinction of competing species); this suggests that the heavy tailed distributions of popularity (distributions that are not exponentially bound, like a power law distribution) may be due to the criticality of the system.

The models used in the simulations are built upon the notion of 'screen space', where every user in the model has their own 'screen'. At every time step, when a user is chosen to 'act', they have the ability to innovate (by creating their own meme), or retweet a meme that is currently on their screen. The user will then send (either the retweet or their new tweet) to all of their followers; if a user is retweeted their popularity is incremented by 1. For the 2 meme model with no meme innovation, an analytical solution was obtained by solving a pair of differential equations. The dynamics behind the fraction of non-empty screens was found to obey the logistic differential equation which is the mean-field approximation for the infected population fraction in a susceptible-infected epidemic model. The mean-field approximation does not consider finite N effects (where N is the total number of users), and therefore the number of screens displaying each meme oscillate about an average value due to stochastic fluctuations. These fluctuations cause one meme to become extremely popular, while the other meme becomes extinct. A further study of these fluctuations was used in the multiple meme model, which also allowed for the innovation of new memes (with a certain probability). The limited resource of user attention (or screen space) lead naturally to create the power law popularity distribution; some memes became extremely popular, while the other memes became moderately popular or ignored. This extremely competitive process causes each meme to follow a critical branching process. On average the number of innovations produced by users will be equal to the number or overwrites by other users; this balancing between the 'birth' and 'death' of memes gives a critical branching process. The differential equation derived governing the popularity distribution for this model was easily solved using standard numerical techniques, and the simulated popularity distribution agreed well with actual Twitter data from 15 million Spanish users.

Future work for J. Ward *et al.* includes generalizing the current model by increasing the capacity of the screen space (allowing more than 1 meme per screen), allowing followers to reject a meme, and permitting users to retweet a meme at most once. Dr. Ward hopes to work closer with researchers in the social sciences to further understand human behaviour in large complex social networks.