Information Flow in Networks — Trendsetters, Bellwethers and Shepherd Dogs

Yaniv Altshuler[1] , Alex (Sandy) Pentland[1]

Human Dynamics Group
MIT Media Lab
`yanival@media.mit.edu` , `sandy@media.mit.edu`

1 Abstract

The rapid expansion of "social network research" is an exciting and unique phenomenon taking place in the last decade. Networks, from their essential definition, serve as an infrastructure for communication — either explicit (e.g. phone calls or emails), implicit (e.g. social signals), or a combination thereof (e.g. online social service such as *Facebook* or *Flicker*). As such, a fundamental aspect that must be addressed in order to truly comprehend networks, is the particular way they interact with and influence information that flows through them. Specifically, when addressing this issue, our goals should be threefold.

First, we must aspire for the *understanding* the underlying mechanisms that control information flow in networks. This should be done by developing analytic models that describe the properties (both local and global) of this process.

Second, using these models (or other methods that will be developed based on them) to generate *predictions* concerning the evolution of information propagation processes. This may include for example the ability to predict the probability that some trend or idea will epidemically spread throughout a given network, . Another example would be the assessment of the *openness* of networks to the proliferation and assimilation of new information or knowledge that are introduced to them by some of their members.

Third, we are interested in developing techniques for efficiently *intervening* in this process. For example, such intervention may take the form of external stimuli that may be injected to the system in order to generate (or strengthen) certain social links, and subsequently facilitate the assimilation of future information that would be introduced to the network (i.e. the creation of "Shepherd dogs", that increase the cohesion of the network). Alternatively, by efficiently recognizing the implicit "roles" of members of the network, certain members can be engaged (e.g. by informative or monetary means) in a way that influence the global behavior of the entire network (i.e. "Bellwethers" and "Trendsetters").

In order to formally discuss the ability of a new idea or a new piece of information to gain popularity over a network, we formally define two terms : "*appeal factor*" and "*persistence factor*", denoting an idea's *local* spreading and deletion expected probabilities, respectively. Using these terms we show that the ability of an idea to epidemically spread throughout a network can be analytically

predicting, resulting in a predictor that tightly thresholds between decaying and prospering spread processes. We then demonstrate this method using several datasets containing information for real world online social networks.

We then discuss the challenges that are involved with the implementation of this technique, and suggest ways of overcoming them. Specifically, we propose a way to generate fast-to-calculate local-information-based predictors that may be able to approximate the results that can be achieved by a theoretical optimal algorithm. We conclude by presenting preliminary results concerning the implementation of such methods using real world online social networks.