Inferring Complex Networks of Influence

Understanding Green Investment Tipping Points

Amir Sani[*]
Centré d'Économie de la Sorbonne
Université Paris 1 Panthéon-Sorbonne
reachme@amirsani.com

Antoine Mandel
Centré d'Économie de la Sorbonne
Université Paris 1 Panthéon-Sorbonne
antoine.mandel@univ-paris1.fr

1 Extended Abstract

Green investment seeks to create suitable financial returns, while also having a (measurable) positive environmental impact. As an example of a socially responsible investment, Green investing concentrates on companies or projects that are committed to the conservation of natural resources, the production and discovery of alternative energy sources (such as biomass/biofuel, geothermal, solar, water and wind), the implementation of clean air and water projects (including renewable energy project developers), clean transportation, energy storage technologies (such as fuel cells), and/or other environmentally conscious business practices (such as energy efficiency technologies). In contrast, Brown investments produce C02 generating "dirty" energy, such as oil, gas and coal, or act in the exploration, drilling, production, processing, refining or related equipment industries.

As the job of these analysts is to influence investment, it makes sense to understand their biases for and against Green investment. Identifying key actors that promote Green and Brown investments through their analysis provides an avenue for understanding and promoting Green investment. In particular, stock analysts produce investment ratings that report on the quality of investing in specific stocks. Policy-makers can study the network of influence over these career influencers to see what policies need to be put in place to inform *key* analysts on the long term costs of Brown investment and social responsibility of Green investment. Unfortunately, the structure (topology) of the underlying network of influence is often unobservable and the success of motivating Green investment requires an effective measure or accurate simulation of influence over clean Green investments. Fortunately, identifying "influencers" in a network is central problem in Network Science (Watts, 2004; Cohen and Havlin, 2010; Estrada, 2012; Newman, 2010).

This paper infers the underlying structure of influence over stock market analysts as temporal traces, cascades. Traditionally, network-based simulation approaches from the finance and economics literature often assume the network topology apriori (Cont and Bouchaud, 2000; Panchenko et al., 2013; Toth et al., 2015) or construct it directly from the data (Schweitzer et al., 2009; Boss et al., 2004). As these edges are unobservable (Adar and Adamic, 2005), network inference methods have been introduced to reconstruct the maximum likelihood edge relationships among actors (along with their model parameters) (Saito et al., 2009; Gomez Rodriguez et al., 2010; Snowsill et al., 2011; Rodriguez et al., 2011; Du et al., 2012; Daneshmand et al., 2014). These algorithms infer the relationship between actors through a disease propagation model, where the relationship between actors is inferred through a set of infection cascades (Rodriguez et al., 2011). Cascades are assumed to represent a stochastic process generated along the edges of an underlying "unobserved" network (Valente, 1995). Further, Network inference algorithms based on disease propagation ignore reverse causality and exploit the forward time-causality of multiple cascades to infer a maximum likelihood network over actors. The edges of the network represent the actor's ability to propagate infection along the network according to a specific distribution, such as the exponential, power or Rayleigh distributions.

[*]Corresponding Author

30th Conference on Neural Information Processing Systems (NIPS 2016), Barcelona, Spain.

The state of infection in these models is limited to an “infected” or “not infected” state. Once infected, it is not possible to be uninfected, so it is not possible to represent changing opinion dynamics, as in the case of stock market analyst ratings. Wang et al. (2012) extend this simple cascade model to conditionally independent features and reinfections, but the model is still limited. Changing opinion dynamics are temporally dependent, so they can not be encoded as conditionally independent of time, and the model still assumes that the nodes can not leave the state of infection once infected. We solve these constraints by extending the NetRate algorithm introduced in Rodriguez et al. (2011) to multiple shifting states in time. Within the context of ratings cascades, these states capture the varying opinions generated by analysts over time. Some examples include "upgrade", "downgrade" and "neutral".

Ratings cascades are generated over Green and Brown stocks listed on the U.S. financial system and stock analyst ratings data downloaded from Yahoo.com. Green and Brown stock lists are drawn from traditionally classified Green and Brown sectors and do not include indirect CO2 output or downstream CO2 output. The resulting network of influence is then used to identify the top 5% GREEN-Influencers and BROWN-Influencers in the U.S. financial system through several standard metrics, including degree centrality, betweenness centrality (Freeman, 1977; Brandes, 2008), closeness centrality (Freeman, 1978), eigenvector centrality (Bonacich, 1987) and PageRank (Page et al., 1999), along with the in-degree and out-degree distributions. These metrics along with the relevant centrality ranks and graphs are reported independently for Green and Brown investments in the U.S. financial system. The resulting networks of influence over Green and Brown stocks provide an insight into the central players that drive opinions among financial actors and subsequently the financial markets.

Acknowledgments

We gratefully acknowledge the support of H2020-FET proactive project DOLFINS and the support of NVIDIA Corporation with the donation of the Tesla K40 GPU used for this research.

References

Adar, E. and Adamic, L. A. (2005). Tracking information epidemics in blogspace. In *Proceedings of the 2005 IEEE/WIC/ACM international conference on web intelligence*, pages 207–214. IEEE Computer Society.

Benzi, M. and Klymko, C. (2013). Total communicability as a centrality measure. *Journal of Complex Networks*, 1(2):124–149.

Bonacich, P. (1987). Power and centrality: A family of measures. *American journal of sociology*, pages 1170–1182.

Boss, M., Elsinger, H., Summer, M., and Thurner 4, S. (2004). Network topology of the interbank market. *Quantitative Finance*, 4(6):677–684.

Brandes, U. (2008). On variants of shortest-path betweenness centrality and their generic computation. *Social Networks*, 30(2):136–145.

Cohen, R. and Havlin, S. (2010). *Complex networks: structure, robustness and function*. Cambridge University Press.

Cont, R. and Bouchaud, J.-P. (2000). Herd behavior and aggregate fluctuations in financial markets. *Macroeconomic dynamics*, 4(02):170–196.

Daneshmand, H., Gomez-Rodriguez, M., Song, L., and Schoelkopf, B. (2014). Estimating diffusion network structures: Recovery conditions, sample complexity & soft-thresholding algorithm. In *ICML*, pages 793–801.

Du, N., Song, L., Yuan, M., and Smola, A. J. (2012). Learning networks of heterogeneous influence. In *Advances in Neural Information Processing Systems*, pages 2780–2788.

Estrada, E. (2012). *The structure of complex networks: theory and applications*. Oxford University Press.

Estrada, E. and Higham, D. J. (2010). Network properties revealed through matrix functions. *SIAM review*, 52(4):696–714.

Estrada, E. and Rodriguez-Velazquez, J. A. (2005). Subgraph centrality in complex networks. *Physical Review E*, 71(5):056103.

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, pages 35–41.

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239.

Gomez Rodriguez, M., Leskovec, J., and Krause, A. (2010). Inferring networks of diffusion and influence. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1019–1028. ACM.

Newman, M. (2010). *Networks: an introduction*. Oxford university press.

Page, L., Brin, S., Motwani, R., and Winograd, T. (1999). The pagerank citation ranking: bringing order to the web.

Panchenko, V., Gerasymchuk, S., and Pavlov, O. V. (2013). Asset price dynamics with heterogeneous beliefs and local network interactions. *Journal of Economic Dynamics and Control*, 37(12):2623–2642.

Rodriguez, M. G., Balduzzi, D., and Schölkopf, B. (2011). Uncovering the temporal dynamics of diffusion networks. *arXiv preprint arXiv:1105.0697*.

Saito, K., Kimura, M., Ohara, K., and Motoda, H. (2009). Learning continuous-time information diffusion model for social behavioral data analysis. In *Asian Conference on Machine Learning*, pages 322–337. Springer.

Schweitzer, F., Fagiolo, G., Sornette, D., Vega-Redondo, F., Vespignani, A., and White, D. R. (2009). Economic networks: The new challenges. *science*, 325(5939):422–425.

Snowsill, T. M., Fyson, N., De Bie, T., and Cristianini, N. (2011). Refining causality: who copied from whom? In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 466–474. ACM.

Toth, B., Palit, I., Lillo, F., and Farmer, J. D. (2015). Why is equity order flow so persistent? *Journal of Economic Dynamics and Control*, 51:218–239.

Valente, T. W. T. W. (1995). *Network models of the diffusion of innovations*. Number 303.484 V3.

Wang, L., Ermon, S., and Hopcroft, J. E. (2012). Feature-enhanced probabilistic models for diffusion network inference. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 499–514. Springer.

Watts, D. J. (2004). The" new" science of networks. *Annual review of sociology*, pages 243–270.