

Modelling information diffusion in faculty hiring networks

Dimitrios Economou and Allison Morgan
CSCI 5352 - Project Proposal

October 4, 2016

1 Proposal

Recent research suggests that within the faculty hiring network there is a core-periphery structure, with prestigious universities occupying central (or core) positions and less prestigious universities living in the periphery [2]. This strong assortative mixing implies that information originating at the core could more easily spread than that originating from the periphery [2]. Similarly, from an epidemiological perspective, the location of an infected individual within a core-periphery structure can determine the magnitude of an outbreak [6]. As such, we propose to run simulations of information flow in the faculty hiring network using models from sociology and epidemiology.

The models we primarily intend to explore are the standard epidemiological models SI, SIR, SIS, and SIRS [6]. These models represent the transmission of infection - from those susceptible to (S), infected by (I), and recovered from being infected (R). For example, within the faculty hiring network, universities can be infected by the research interests of their faculty members. Universities may recover from this infection if they decide to no longer participate in this research (perhaps faculty doing work in this field retire or leave). We would also like to explore a few sociological models of idea diffusion like SEIZ [1], independent cascade (IC), and linear threshold (LT) [3, 4]. A general survey of epidemic models can be found here [7].

In our simulation, we will assign each vertex in the faculty hiring network a state: susceptible, infected, recovered, etc. We initially choose a vertex to infect and let all the other vertices be susceptible. Let S , I , and R be the number of susceptible, infected, and recovered vertices, respectively. The *classic epidemic model* of SIR says that the dynamics of these quantities is given by:

$$\begin{aligned}dS/dt &= -\beta IS/N, \\dI/dt &= \beta IS/N - \gamma I, \\dR/dt &= \gamma I, \\N &= S + I + R,\end{aligned}$$

where β is the contact rate, γ is the recovery rate, and N is the (constant) size of the population [5, 6]. For example, we may consider β to be some function of the number of infected neighbors a vertex has and γ to be some function inversely proportional to the prestige of the initially infected university. Using these equations, we'll be able to observe how the initial position and prestige of the vertex determines how infectious (e.g., the number of infected individuals or time it takes to infect the whole network) an idea can be.

2 Data Description

The faculty hiring network is represented by a collection of records, each of which corresponds to a unique faculty member. The records have the following fields: faculty member name, sex, department, place (university), and current position (such as assistant professor), as well as their education and employment history (degree, place, field, and years active).

References

- [1] L. M. A. Bettencourt, A. Cintrón-Arias, D. I. Kaiser, and C. Castillo-Chávez. The power of a good idea: quantitative modeling of the spread of ideas from epidemiological models. 2005.
- [2] A. Clauset, S. Arbesman, and D. B. Larremore. Systematic inequality and hierarchy in faculty hiring networks. *Science Advances*, 1(1):e1400005, Feb. 2015.
- [3] A. Dotey, H. Rom, and C. Vaca. Information diffusion in social media, 2011.
- [4] E. Gaussier. Models of information diffusion in social networks. In *Proceedings of the Second Symposium on Information and Communication Technology*, SoICT '11, pages 2–2, New York, NY, USA, 2011. ACM.
- [5] H. W. Hethcote. The mathematics of infectious diseases. *SIAM Rev.*, 42(4):599–653, Dec. 2000.
- [6] M. Newman. *Networks: An Introduction*. Oxford University Press, Inc., New York, NY, USA, 2010.
- [7] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani. Epidemic processes in complex networks. *Rev. Mod. Phys.*, 87:925–979, Aug 2015.