How Does Metoo Movement Spread in Twitter?
——investigating the complex contagion model in social networks

Literature Review

Jielu He

Social contagion is a mechanism that explain the spread of ideas and behaviors in social networks. In particular, complex contagions are thought to dominate the spread of social movements and other important social phenomena. When adopting an idea or behavior that may influence the whole social group, the possibility of one person's adoption always depends on the number of different neighbors adopted. Metoo Movement, for example, opening up social interactions on a large scale and providing an opportunity for us to observe the spread of social communication "in the wild".

Much of the work in searching for complex phenomena in real world contagions focuses on measuring user adoption thresholds. In this work, I want to show an alternative method for fitting complex contagion models to empirical twitter data that avoids measuring thresholds directly. My research focuses on what we can observe - the empirical adoption curve for Metoo hashtag - and determines which contagion model best fits the twitter data about Metoo movement.

I. Collective behaviors and user adoption thresholds

Granovetter (Granovetter 1978) describes a class of group behaviors. Only after finding a certain number of others choose to participate in the group behavior, then an individual may engage in as well. A slight change in the threshold distribution across the entire population may lead to completely different results. Part of this heterogeneity of the threshold depends on local differences in network structure and temporal and spatial factors. Riots, mentions the spread of rumors and innovation all serve as examples of collective behavior that relies on adoption thresholds. Centola and Macy (Centola and Macy 2007) described Granovetter's collective behavior threshold model as a complex contagion and studied how they spread across the network. In response to Granovetter's (Granovetter 1973) work on the importance of long-distance connections in networks for spreading thoughts and behaviors, they used simulation to prove that although complex contagions can saturate dense areas of networks with many redundant connections, unless there is enough number of redundant long-distance connections, otherwise it will not be able to spread to uninfected areas of the network. Centola (Centola, 2010) designed an experiment using a constructed real-world online network and tested the threshold behavior on a highly clustered lattice network and a randomly rerouted lattice network. All nodes have the same degree, and the experiment is repeated for different node degrees and network scale. He found that contagions spread faster and farther in clusters than in random networks, which led his research results to prove the importance of redundant relationships for the transmission of complex contagions. In related work, Barash et al. (Barash, Cameron, and Macy 2012) Simulated complex contagions on meshless and scale-free networks. It was found that complex contagions require a certain number of adopters before sufficient redundant long relationships can propagate it to the network. Theoretical studies on complex contagions indicate that, in addition to densely connected dense networks, the obstacles to the success of complex contagions are also obvious, which illustrates Granovetter's original description of these processes.

II. Empirical evidence and data analysis

In view of the growth of user-generated content on the Web in recent years, researchers have turned to social media data to seek empirical support for complex contagions theory. Barash (Barash, 2011) viewed photo tags on the photo sharing service Flickr, which allows users to comment and tag photos. Some tags related to offline behavior will bring certain risks (for example, the "jailbroken" iPhone is an example), showing some signs of complex contagions. Romero et al. (Romero, Meeder, and Kleinberg, 2011) used 8 million tweets from 2009 to 2010, selected the top 500 hashtags based on user mentions, and sorted by topic category (politics, celebrity, sports, music, And Twitter idioms). They built a network based on tweets from other users. For the case where users initially used hashtags, they calculated p (k): the label adopted by users with k or more exposures contain exactly the percentage of k exposures. They found that the p (k) curve of the political label (effectively the probability range of the threshold) is always different because p (k) is higher and remains higher as k increases. They also found that those who initially adopted the political label formed a denser network, which they believed was consistent with complex contagions and the need for social strengthening that adopted controversial ideas or behaviors. Using the p (k) value calculated for the user, they simulated the cascade on the constructed network using the user's original seed set and random seed. They found that the experience cascade of political and idiom tags grew faster than the cascade based on random seeds.

Goel et al. (Goel et al., 2015) conducted a larger study of Twitter, using 1.2 billion tweets from 2011 to 2012, and each tweet contained a URL pointing to one of a specific set of websites. Their definition of adoption is that users write a tweet that contains one of the 622 million unique URLs they identified. They defined a measure of "structural virality" for the cascade used to balance the width of the cascade (due to the broadcast from a single node) and the depth of the cascade. Most cascades are small, so they focus on 0.025% of a cascade of 100 or more nodes. They discovered various cascading structures, including large broadcasts, deep tree structures, and many hybrids of the two. These URLs are classified as links to videos, news or online petitions. They found that 25% of the petitions have a high degree of structural virality, while other types of petitions have a high degree of virality, and petitions of different levels have a high structural virality. This result is interesting and consistent with the findings of (Romero, Meeder, and Kleinberg 2011) that the initial density of adopters of political labels is higher. One explanation is that, similar to the use of political tags, links to online petitions require social reinforcement, while a denser local network structure facilitates this reinforcement through redundant bonds. Finally, they used a susceptible infection (SI) infection model and a constant infection rate to simulate on the simulated priority attachment map. The range of cascade showed many distributional aspects of the empirical cascade. Their simulation results show that a simple contagion model can explain the diversity of the real-world cascades found in its Twitter dataset.

State and Adamic (State and Adamic 2015) investigated the evidence for complex contagion on Facebook in the context of people registering support for same-sex marriage. They used anonymous data from 3 million Facebook profiles to work, and they used voluntary demographic information related to the profile to investigate the link between these factors and the number of friends who had to change photos before users decided to change their photos. They found that, for example, there are more users who indicate same-sex attractiveness in their profile than users who have few or no friends indicating same-sex attractiveness (for example, changing their profile picture). Adoption increases with the number of adoptions by friends, which is especially evident when users register to be interested in their friend 's adoption by clicking a posted link or “like” posts related to same-sex marriage. They looked at the adoption thresholds of other content and found "cut and paste memes". If users might expose themselves to doubt or ridicule, they need a higher adoption threshold. Finally, a simulation was conducted on the experience network of 800,000 users. They model infections based on user sensitivity (randomly drawn from an exponential distribution) and the probability of infection as the number of friends used decreases logarithmically. For higher infection rates, as shown by empirical data, the adoption threshold increases with the number of adopted friends.

References

Barabasi, A.-L. 2005. The origin of bursts and heavy tails in human dynamics. *Nature* 435(7039):207–211.

Barash, V.; Cameron, C.; and Macy, M. 2012. Critical phenomena in complex contagions. *Social Networks* 34(4):451 – 461.

Centola, D., and Macy, M. 2007. Complex contagions and the weakness of long ties. *American Journal of Sociology* 113(3):702– 734.

Centola, D. 2010. The spread of behavior in an online social network experiment. *Science* 329(5996):1194–1197.

Coleman, J. S.; Katz, E.; and Menzel, H. 1966. *Medical innovation: A diffusion study*. Indianapolis: The Bobbs-Merrill Company.

Goel, S.; Anderson, A.; Hofman, J.; and Watts, D. J. 2015. The structural virality of online diffusion. *Management Science* 62(1):180–196.

Romero, D. M.; Meeder, B.; and Kleinberg, J. 2011. Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter. In *Proc. of the 20th Internat. Conf. on World Wide Web*, 695–704. ACM.

State, B., and Adamic, L. 2015. The diffusion of support in an online social movement: Evidence from the adoption of equal-sign profile pictures. In *Proc. of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 1741– 1750. ACM.

Valente, T. W. 1996. Social network thresholds in the diffusion of innovations. *Social Networks* 18:69–89.