Paper Review: The role of social networks in information diffusion[1]

Out of the papers that we have read for the course so far, I believe that this paper described the methodology and conclusions the best. The authors make use of the large network of Facebook users (~2.5 million user pairs) in an experimental study while also addressing some major biases along the way. I found the usage of Relative Risk (RR) to measure the association between the networks of *feed* and *no feed* was appropriate in context as a method to counteract homophily in the network. One area which confused me was when the authors talk about dealing with the bias of popular URLs by bootstrapping by clusters. I hope that this area will be addressed in the upcoming lecture because I'm interested in learning more about this.

A lot of the effects of having multiple friends sharing the same content seem to match the results that I was expecting. There is an associated decrease in RR that having more friends sharing the same URL brings about. In addition, a user is more likely to share quickly if there are friends sharing the same URL. It brings into question how the concept of News Feed, which has long since changed from 2012 to Facebook Groups and Pages, is now redundant in the scope of the study. If the study were to be conducted in the present day, how would the type of experiment change to encompass it? Would multiple networks have to be formed in order to assess the impacts of each one and their relative effects?

The authors do a good job of explaining the factors that determine the tie strength before delving into their analysis. The graphs clearly show that people with strong ties are more likely to share the content and number of weak ties determines the RR. What is interesting is how the interaction between strong and weak ties takes place. There will certainly always be far more weak ties than strong ones, thus showing that diffusion of novel information is most certainly aided by weak ties. This way, the authors come to a firm conclusion that FB sharing URLs works in a similar way to a contagion, while finding evidence to support Granovetter's weak ties theory[2].

As aforementioned, I am interested in seeing how future research addresses multiple networks within the framework of a social media platform such as Facebook. Now that there is overabundance of data (e.g. number of clicks, time spent viewing a post, etc.) the analysis will certainly be more complex. What newer techniques, compared to the traditional exposed/unexposed model using RR, would be used in such a situation?

---

[1] https://dl.acm.org/doi/10.1145/2187836.2187907
[2] https://www.jstor.org/stable/202051?seq=1#metadata_info_tab_contents