NETS 112 F2016 HW1 Solutions

1. [Diameter, "Economic Altruism Model"] - graded by Brad

a. Diameter = average shortest path distance in the network - 25/15 = 5/3 ~ 1.67
(Note: longest distance path on network = 3 —> this is not the definition of diameter we want)

b. Eq. values using w(i) = W*(d(i) / D) where W = total wealth = 6 and D = total degree = 14
w(A) = 3/7
w(B) = 9/7
w(C) = 6/7
w(D) = 9/7
w(E) = 9/7
w(F) = 6/7

c. Edges A —> C and A —> F

d. Key = network with 4 different degree values

2. [Contagion & Immunization] - graded by Adel

a. EV = 7(7/25) + 3(3/25) + 7(7/25) + 2(2/25) + 6(6/25) = **5.88**

b. Immunize hub of the "star"
New EV = 7(7/25) + 3(3/25) + 0(1/25) + 6[1(1/25)] + 2(2/25) + 6(6/25) = **4.12**
OR
New EV = 7(7/24) + 3(3/24) + 0(1/24) + 6[1(1/24)] + 2(2/24) + 6(6/24) = **4.33**

[*Note: you needed to explain why you used 24 vs 25 in your calculations in order receive credit for using the second equation]

c. Connect the groups of 2 and 3 vertices
New EV = 7(7/25) + 5(5/25) + 7(7/25) + 6(6/25) = **6.36**

3. [WikiLeaks Game] - graded by Brad

The most common strategy was to zoom out, move in the general space, then zoom back in. This is similar to the travel analogy that Dr. Kearns discussed in class: you typically travel between international airports to cross large distances and taxis to cover the smaller distances at either end of the trip. The other common idea was to stay in topics that you know about, which makes sense because for such topics, you can predict what links they will have. This is like already knowing who your friends friends are when navigating through a Facebook network.

4. [Forest Fire Demo] - graded by Adel

For green level wind speed: tipping point around 60% forestation

For yellow level wind speed: tipping point around 50% forestation

For red level wind speed: 1/2 burned around 70% forestation (arguably less tipping point behavior, more of a spread)

Possible answers for how to use network structure to model wind speed: orientation of dense (or less dense depending on desired “wind speed) connectivity in NE direction; more or less clusters with connectivity to nodes located along this axis, etc. We accepted any well formed explanation as long as it included information on wind direction and/or directed networks, connected components, degrees, etc.

5. [Synthesis of Contagion Articles] example answer (longer than ideal) - graded by MP

While the three of the reports differ in their methodologies and motivations, each one investigates some aspect of structural virality in social contagion.

The first paper, “Can Cascades be Predicted?” provides the result of tracking a number of cascades over time in order to determine whether or not their ultimate sizes can be predicted and if they can be, which cascade features are most predictive. They found that cascade size can indeed be predicted and that structural and temporal features have the highest predictive value. However, they also found that the salience of structural features declines over time. They concluded that events with large initial breadth will produce bigger cascades than events with large initial depth. This result relates to the third essay’s results on “broadcast” contagions: the largest broadcast drives an event’s popularity (proxy for reshares in this context). One limitation of this study is mentioned at the end of the paper. The authors remind us that as cascades get larger and larger, Facebook’s algorithms actually affect the cascade in a way that could artificially augment it. In other words, there is a confounding variable affecting the ultimate size of the cascade.

The second paper, “Structural Diversity in Social Contagion,” looks into the claim that the probability of contagion is determined by the structural virality of an individual’s “contact neighborhood,” rather than by the size of the neighborhood. The authors investigated this claim by analyzing the conversion rate of Facebook invitations sent to nonmembers with a list of members selected from their email contact list. They found that the invitations with the largest acceptance rates were those that listed members with networks of low structural virality. In other words, emails that included member networks in which the individuals were from different social contexts had the highest likelihood of acceptance. One possible limitation on this study is the nature of the inviters and invitees. The inviters must have used a specific contact importer tool and selected at most twenty individuals on their contact lists to invite. This means that the selection process was probably quite personal and socially oriented. This feature of the study might detract from the usefulness of the results by only offering information on a very niche group of users. However, the authors state that the study accounts fro potential confounding variables such as the one just mentioned.

The third paper, “The Structural Virality of Online Diffusion,” seeks to determine whether or not the biological contagion model is an appropriate model for all social networks. The authors examined the retweet patterns of the most popular Twitter events (100+ retweets) over a twelve month period. The authors found that structural virality is low in the most popular events observed, meaning that broadcast events have the largest popularity. They also found only a weak relationship between size and virality which makes the problem of determining how content spread much harder and proves the necessity of a way to classify cascades based on their structural virality. At the end of the paper, the authors discuss a number of alternative definitions of structural virality (i.e. relative size of the largest broadcast, the average depth of nodes in the tree, the probability that two randomly selected nodes will have distinct parent nodes). They also concede that the definition they use may not be the best one for the purposes of their study. However, they did find that the way they defined it is correlated with the other measures they cite and that the results of their study are robust with respect to the way they chose to define structural virality.

The source of information for the first and second papers is the Facebook data base. The third relies on Twitter's retweet date of popular events for its analysis. The first and third papers are similar in that they both examine the patterns inherent in resharing content on their respective platforms. Both of these papers also determine interesting results for the structural virality of broadcast social contagions, external applications for their results, and contributions to the progress of computational social science. The second paper reveals how structural virality affects people's behavior in social networks and is also very informative for both computer science and social science fields. The authors of all three papers prove something about contagion in social networks by examining historical data to draw their conclusions. These papers also show us how studies of relatively niche characteristics of social networks on specific platforms can tell us a lot about the larger dynamics at play within many different diffusion settings.

Acceptable answers to this question took many forms. However, in order to earn full credit, it was necessary for students to go beyond just summarizing the articles. Students needed to include a synthesis/comparison of the articles and their findings, as well as discuss the weakness and limitations of each.