

The Impact of Influence on the Dynamics of Complex Information Cascades

ADRIANA DIAKITE, EMILY GLASSBERG, EYUEL TESSEMA

Stanford University

diakite@stanford.edu, ecgberg@stanford.edu, etessema@stanford.edu

I. INTRODUCTION

I. Motivation

How do behaviors, contagions, products, and ideas spread through networks? These questions can all be addressed using the conceptual framework of information cascades. The impact of network characteristics on information cascades is of great interest to sociologists, politicians, public health professionals, marketing executives, and many others. Propagation through networks can be simple, meaning that contact with one "infected" node is enough for transmission. Many diseases spread in this fashion. However, propagation can also be complex, meaning that a certain proportion of a node's neighbors must have adopted a behavior or idea before the node will decide to adopt it as well. Current literature suggests that cascades involving social movements, attitudes, and ideas are likely to be transmitted in this way (Granovetter 1978). As a result, the dynamics of complex information cascades differ from simple cascades. Furthermore, the dynamics of social networks are influenced by the specific relationships between individuals within that community. For example, it has been shown that people's behavior in social networks is influenced by an individual's status (Leskovec et al, 2010). Existing models of complex transmission are limited in that they treat the relationships between all pairs of nodes as equal, not accounting for potential effects of homophily, interpersonal relationships, and status on an individual's decision to adopt a behavior or idea (Watts 2001, Centola et. al 2007, Centola 2010). This project aims to extend

current research on the dynamics of complex information transmission by adding variation in influence across a navigable, small-world network model. We then explore the impact of variation in influence on real world information cascades by comparing the results of our simulated cascades to a cascade of retweets on the Twitter network.

II. Prior Work

Current research on complex propagation has focused on random networks, regular lattices, and small-world networks. A 2007 paper by Centola et. al. investigated the impact of network topology on the "critical threshold," the threshold above which no cascades occur. The authors found that, in random networks, increasing the degree of each node decreases the critical threshold (promotes propagation). However, in regular lattices, increasing the degree increases the critical threshold (inhibits propagation). Most interestingly, increasing the randomness of regular lattices by rewiring edges also increased the critical threshold. Previous literature has shown that increasing randomness in the form of long-range links allows for increased transmission between neighborhoods (Granovetter, 1973), however, the above results suggest that these conclusions do not hold in the case of complex transmission. Additionally, a 2010 paper by Centola found that complex transmission occurred more rapidly and with greater success (as determined by the proportion of nodes that adopt the cascading trait) in clustered lattices than in random networks. The author suggests that this difference

results from increased social reinforcement due to the high degree of clustering. These findings support the claim that the dynamics of complex propagation are distinct from those of simple transmission. Specifically, these results demonstrate that complex and simple information cascades are impacted differently by network topology. However, the network topologies currently studied in the context of complex information cascades may be insufficient to capture the dynamics of cascades in real-world networks. For example, random graphs do not reproduce the clustering of real-world networks, regular lattices do not capture the small diameter seen in real-world situations, and small-world networks are not navigable in the way that real-world networks are. Furthermore, these analyses assume that each node in the network exerts equal influence on its neighbors. A significant body of research on the impact of properties of social groups such as status and homophily suggests that this assumption is highly unrealistic in real-world settings. In this analysis, we expand existing research on complex information cascades in two ways. First, we investigate the behavior of complex information cascades on navigable, small-world networks. Second, we explore two models of variable influence and their impact on the rate and success of complex information cascades.

II. METHODS

I. Description of Models

a) Basic Model - The basic model we will use is an extension of the small-world network model (Watts and Strogatz, 1998). The network will be represented as a balanced binary tree of height H . The leaves of the tree are the nodes of the network. Nodes will be formed into completely connected "neighborhoods" by placing an edge between every pair of nodes in subtrees of height h . Edges will be added between pairs of nodes in different neighborhoods with probabilities that decrease as a function of the dis-

tance between the nodes. The distance between a pair of nodes will be the number of steps to the most recent common ancestor of that pair on the tree. This process will be repeated until r edges have been added from each node to nodes in different neighborhoods (Model I). By varying the size of the neighborhoods ($2h$), and the inter-neighborhood connectivity (r), this general model will allow for the investigation of variable influence on a variety of network topologies.

b) Models of Variable Influence - Common sense suggests that, in real world networks, distance between nodes is likely to affect the influence that one node exerts on another. For example, you are more likely to adopt the behavior of a family member or close friend with whom you are often in contact than the behavior of a distant acquaintance that you rarely see. Therefore, we model influence as a decreasing function of the tree distance between a target node and its neighbor (Model II). Specifically, the influence of a node v on a node w is calculated as $[\frac{1}{d(v,w)}]$, where $d(v,w)$ is the number of steps to the most recent common ancestor of v and w on the tree. It is worth noting that Model I also indirectly incorporates a distance-based effect in terms of the ability of one node to influence another (nodes that are further away from a target node are less likely to have an edge to the target and so are unlikely to influence it). For purposes of comparison, we implement a third model wherein influence is modeled as an increasing function of the tree distance between a target node and its neighbor (Model III). In this case, the influence of a node v on a node w is calculated as $d(v,w)$. This model could be interpreted as a scenario in which, as you become familiar with the people around you, you become immune to their views and no longer adopt behaviors from them. By contrast, when you meet someone new, their behaviors or attitudes are surprising or interesting enough that

you are more likely to adopt them.

II. Description of Simulations

We generate a series of graphs each of $H = 10$, varying h from 0 – 5 and r from 1 – 10. We then calculate the critical threshold (T_c) for each network topology (as specified by h and r) by simulating complex cascades on each of these graphs.

We simulate an information cascade in a network by seeding a random neighborhood with a new state. At each of one thousand time steps, each node will change state if a certain proportion of its neighbors had previously adopted the new state. This proportion is called the adoption threshold (discussed in greater detail below). Each graph contains 2^{10} (1024) nodes and, at every iteration, a cascade will either infect at least one additional node or the cascade will cease permanently. 1000 nodes is $\approx 98\%$ of the 1024 nodes in the original graph, so, if a cascade has not succeeded by 1000 iterations, we know that the cascade must have ceased prior to the final iteration. Therefore, performing more than 1000 iterations will have no impact on the final cascade size.

On a single network topology, we simulate information cascades as described above at varying adoption thresholds. Because our graphs are symmetric, every node in a network will have roughly the same degree. The degree of each node (k) is approximately the number of neighbors within the node's neighborhood plus the number of neighbors in different neighborhoods $((2h - 1) + r)$. Given that the maximum h of our graphs is five, the maximum r is ten, the maximum degree of nodes in our network is 41, and the maximum distance between two nodes is ten. Therefore, the minimum contribution of any one node towards the adoption threshold is less than $(\frac{1}{41}) * (\frac{1}{10}) \approx 0.002$. Therefore, for the simulations on each graph, we initialize the adoption threshold to 0.001, meaning that nodes need at least one of their neighbors to have adopted a trait in order to adopt it themselves. After each successful cascade ($> 90\%$ of nodes in

the graph adopt the new state), the adoption threshold is increased by 0.001, meaning that, relative to the previous simulation, nodes require at most one additional neighbor to have adopted a trait in order to adopt the trait themselves. The simulation is then repeated until we reach the adoption threshold at which cascades fail. We say the value of the critical threshold (T_c) is the value of the adoption threshold at the last successful cascade (0.001 below the threshold at the first failed cascade). We repeat this procedure for each of the network topologies described above.

Finally, we repeat this set of experiments using each of the models of variable influence described above. In these simulations, the contributions of each neighbor towards the adoption threshold are weighted by the influence score of that neighbor.

III. Analysis of simulations

To compare the relationships between network topology, critical threshold, and influence, we plot 3-dimensional graphs for each of our models with h , r , and T_c on the axes. This analysis, however, only describes cascades in terms of success or failure.

To explore the behavior of cascades in greater detail, we created heatmaps to represent the size - in terms of percent of nodes infected - of both the last successful cascade (adoption threshold = T_c) and the first failed cascade (threshold = $T_c + 0.001$) on each network topology.

IV. Validation on real world data

To test the relevance of our models to real world cascades, we construct a three-tiered network using data from the Twitter network. We start with a source user and retrieve two levels of neighbors. First, we retrieve the primary neighbors; the followers of the source user. We then retrieve the secondary neighbors; the followers of the primary neighbors. Since each of our users has around 100 followers (this is the average number of followers for Twitter users, according to Lerman et. al, 2010), this should

leave us with a tree-like graph with $\approx 10,000$ nodes/users and $\approx 1,000,000$ connections.

A limitation of the current study is that Twitter has harsh limits on the number of queries one can make to an api endpoint (15 queries every 15 minutes). As a result, we have not yet completed building the tree described above. We hope to have it completed and analyzed for the poster session on Friday.

Once the graph is complete, we will simulate cascades using each of the three models of variable influence described above. To apply Model I, under which every node exerts equal influence over every other node, we seed the graph with a neighborhood comprised of the source user and all of her followers. We then simulate "cascades" at variable adoption thresholds and track how many secondary followers we expect to be infected. For Models II and III, where the influence a node exerts on a target node is a function of the distance between those two nodes, the distance was determined using the time zones of the users. The distances between each pair of time zones are specified in Table 1 below. With these distances in hand and the influences of primary followers over secondary followers weighted appropriately, the seeding and simulations are performed as described for Model I. The proportion of users infected by each simulated cascade will be recorded. Using the protocol described above, we will also use these simulations to calculate a critical threshold (where $> 90\%$ of nodes in the network are infected) for each of the three models of variable influence on the graph we created from the Twitter data.

To compare the results of this cascade to real world cascades originating from the source user, we extracted the source user's 100 most recent tweets. We also extracted the number of times that each of these 100 tweets was retweeted counts for these tweets. We assume that the average number of retweets across these tweets is the expected size of an information cascade originating from this user. Note that the 100 most recent tweets of the source user include tweets that the source user herself has retweeted. We chose to include

these retweets for two reasons. First, 'personal' tweets (personal here does not imply private) may be less likely to reach users beyond the source user's original friend group. Tweets that did not originate with the source user are less likely to be personal, and therefore are more likely to be retweeted by her followers and be more interesting to analyze. Second, we assume that, in many cases the users in our graph retweeted each story because they encountered it through the source user. It is uncertain how valid this assumption is; the source user and her followers could have discovered and chosen to retweet the tweet independently. Additionally, the source user could be following one of her followers, and so the tweet could have originated in a lower tier of the graph. However, it seems that we should not completely discount this number.

We will attempt to validate the methods used in this analysis using real world information cascades in the following ways. First, we will compare the sizes of the simulated cascades on our Twitter graph to the average retweet count of the source user's tweets. This will allow us to assess the validity of our simulation procedure (if, on a real world graph, it can generate cascades similar to those seen in the real world). Second, for each model of variable influence, we will compare the critical threshold calculated on our Twitter graph to that calculated on our model graphs. This will allow us to check the validity of our graph topologies (if, on the model graphs, cascades behave similarly to those seen on real world graphs). Third, what we would most like to be able to do as a third step is to determine which of our models of variable influence most closely represents how information cascades behave in the real world. To do this, we would like to compare the critical threshold of the true retweet cascade to the simulated critical thresholds under each model of variable influence. Unfortunately, it is impossible to determine critical thresholds from data regarding real world cascades. An information cascade either succeeds or fails; it is impossible to tell under what conditions the same cascade would

behave differently. Another method would be to compare the sizes of real world cascades to that of our simulated cascades. This is limited by several factors. First, Models I and III produce very similarly sized cascades (as does Model II except on very specific graph topologies, when intermediate cascades can occur). As a result, it would be very difficult to distinguish these models using this method. Second, and perhaps even more seriously, in the real world, most cascades infect far fewer than 90%, or even 50%, of individuals in the network. Lerman et. al 2010 found that, in their sample of 140,000 active Twitter users, the average retweet count was 400. Furthermore, only 15 stories were retweeted more than 6,000 times. Therefore, even for tweets that clearly would be considered to have incited successful retweet cascades, the proportion of infected nodes in the network is only 0.04. This suggests that, to understand information cascades on real world networks, we need to rethink how we determine a successful cascade.

III. RESULTS AND DISCUSSION

I. T_c as a function of h and r

Model I - In general, the relationship between h , r , and T_c is saddle-shaped. T_c is the highest when neighborhoods are small and there is a low degree of inter-neighborhood connectivity ($h = 0$, $r = 1$, $T_c \approx 0.25$). As high T_c means that cascades are more likely to occur, these data suggest that, under the model of influence where all nodes exert equal influence on one another, cascades are most likely to occur when neighborhoods are small and there is limited connectivity between them. This finding is counterintuitive, but consistent with the results found in Centola 2007. In complex information cascades, long-range links are unlikely to infect distant regions of the graph. This is because nodes far away from the originally seeded neighborhood are less likely to be connected to multiple infected nodes. This phenomenon is especially true in our model, where edges are added between nodes in dif-

ferent neighborhoods as a decreasing function of distance. Thus, increasing the number of long-range links is likely to increase the degree of a node without increasing the number of infected neighbors it is connected to. This dilutes the impact of edges to nearby neighbors that are infected, making cascades less likely. However, inter-neighborhood edges are essential for cascades to spread from one neighborhood to another. This suggests that there will be some small or intermediate number of inter-neighborhood edges that will be sufficient to spread cascades, but small enough that the dilution of short-range edges will be limited. At small neighborhood sizes ($h \leq 1$), increasing r inhibits cascades, causing T_c first to drop rapidly, then flatten off between $T_c = 0.10$ and $T_c = 0.05$. At intermediate neighborhood sizes ($2 \leq h \leq 3$), increasing r displays no unidirectional effect with respect to T_c . At large neighborhood sizes ($h \geq 4$), increasing r slightly promotes cascades, causing T_c to increase slightly, then flatten off around $T_c = 0.125$. So, when all nodes exert equal influence over a target node, increasing neighborhood size appears to make cascades on the network more robust to increases in r .

Model II - Similarly to Model I, in Model II, T_c is greatest when both h and r are low ($h = 0$, $r = 2$, $T_c = 0.175$). However, under this model of influence, where nodes exert greater influence over nodes that are closer to them, the shape of the relationship between h , r , and T_c is much more plane-like than the saddle-shaped curve seen in Model I. When inter-neighborhood connectivity is small ($r \leq 3$), increasing h inhibits cascades (T_c decreases). Furthermore, when neighborhood size is intermediate or large ($h \geq 2$), increasing r promotes cascades (T_c increases). This relationship does not hold for small neighborhood size ($h \leq 1$), in which case increasing r is not associated with a visible effect on T_c . In general, cascades occur more readily under Model I than under Model II. In Model II, nearby nodes are weighted more heavily than more distant nodes. Thus, under this model, the nodes that are weighted the most heavily

are the nodes within the neighborhood of the target node. So, if a target node exists in an uninfected neighborhood, all of its most influential neighbors are also uninfected. It will therefore be more difficult to infect new neighborhoods under Model II than under Model I. However, when neighborhoods are small and inter-neighborhood connectivity is high, cascades occur more readily under Model II. Recall that, under Model I, increasing neighborhood size increases the robustness of cascades to increasing r . Under Model II, nodes exert greater influence over nodes close to them. This counteracts the effects of dilution; it increases the robustness of cascades to increasing r and promotes cascades, even on graphs with small neighborhoods.

Model III - Under Model III, where nodes that are farther apart exert more influence over one another than nodes that are close together, the relationship between h , r , and T_c is similar in shape to that seen in Model I. However, at small and intermediate neighborhood sizes ($h \leq 3$) and low inter-neighborhood connectivity ($r \leq 3$), when r is increased, cascades are inhibited and T_c rapidly decreases. Under Model I, increasing r at intermediate neighborhood sizes shows no clear, unidirectional effect on T_c . This suggests that cascades under Model III are less robust against increases in r than are those under Model I. T_c is frequently higher under Model III than Model I (cascades occur more easily). This can be understood in the following way. Nodes within a neighborhood will, by definition, be closer together than nodes that are in different neighborhoods. Consider a scenario in which a target node is located in a currently uninfected neighborhood. Many of the close neighbors of this target node will also be within the uninfected neighborhood. However, it is possible that several far away neighbors, in different neighborhoods, will be infected. As a result, giving greater weight to far away neighbors increases the target node's chances of adopting the new state. This may explain the ease of cascades under Model III relative to those under Models I and II.

II. Proportion of nodes infected in first failed cascade

Under Models I and III, the first failed cascade ($< 90\%$ of nodes infected; adoption threshold = $T_c + 0.01$) infects between 0.1% and 6.25% of the nodes in the graph (Figure 4). This corresponds to a single infected neighborhood (the original seed), or, at most, two infected neighborhoods. By contrast, under Model II, there exist graph topologies where up to 50% of nodes in the network become infected (e.g. $h = 5$, $r \geq 8$, Figure 4). By definition, under all three models and at every graph topology, when the adoption threshold is at or below the critical threshold, more than 90% of nodes in the network become infected.

Cascades under Models I and III, therefore, display an "all or nothing" behavior; either very few nodes are infected or nearly the entire network is infected.

Under Model I, a source node exerts equal influence over all of its neighbors. There are fewer edges between a source neighborhood and neighborhoods that are farther away than neighborhoods that are closer together, so it is most likely that the seeded neighborhood will first infect the neighborhood closest to it. This neighborhood also has few edges to neighborhoods at greater distances, however, with both of the first two neighborhoods infected, there are up to twice as many randomly selected inter-neighborhood edges. Infecting the second neighborhood at distance 1 increases the probability of infecting a further away neighborhood. Similarly, infecting the third neighborhood increases the likelihood of infecting the fourth, and so on, hence the "all or nothing" behavior. Under Model III, nodes further from the target node exert more influence than nearby nodes. Because the distance between two nodes is determined by the number of steps to the nearest common ancestor of the pair, the nodes that exert the greatest influence over a target node are the nodes in the opposite half of the binary tree. Thus, if the adoption threshold is below the critical threshold, once a neighborhood is seeded, this neighborhood

is likely to infect every neighborhood in the opposite half of the tree. The newly infected half of the tree is equally far from the remaining neighborhoods of the first half of the tree (containing the seed neighborhood), and so those remaining neighborhoods are will be infected as well. This leads to the observed "all or nothing behavior." However, under Model II, cascades reach intermediate levels of success, infecting only a portion of the network. The proportions of infected nodes in the first failed cascade at each graph topology under Model II are values of $\frac{1}{2^i}$, where i is an integer between 0 and 10. This is because there are 2^i nodes at distance i from the source node. Under Model II, cascades are very likely to infect these 2^i nodes at small values of i . However, when i is small, there are few nearby neighborhoods to infect (1 neighborhood at distance 1, 2 at distance 2, 4 at distance 3, etc.). Once that small number of nodes in nearby neighborhoods has been infected, for the cascade to spread it must overcome a greater hurdle; it must infect the 2^i nodes at the next highest value of i . Thus, as more neighborhoods become infected, it will become more difficult to infect the remaining neighborhoods. Therefore, as the adoption threshold increases, first cascades fail to infect the farthest part of the graph (the opposite half), then they fail to infect the next farthest part, etc.

III. Validation - Cascades of Retweets on Twitter

As previously mentioned, data regarding retweet cascades on Twitter are forthcoming. The following are some basic statistics regarding the tweets we collected. Note the very small average retweet count of tweets originating from the seed user and the very large average retweet count of tweets that were retweeted by the user. This justifies our decision to include the source user's retweeted tweets in our analysis.

IV. FUTURE WORK - EXTENSIONS OF THE MODEL

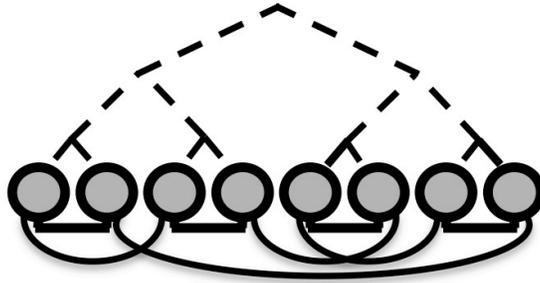
The results of the simulations on our model graphs are promising and provide valuable insight into how varying influence across nodes a network may impact how information spreads through a network. However, much work remains to be done in order to determine how variable influence operates to affect the dynamics of complex information cascades on real world networks. We have considered several variations of the model used in this analysis that may improve its relationship to real world data. Future analyses could benefit from addressing each of the extensions discussed below. First, our current model sets adoption thresholds to be the proportion of neighbors infected. It would also be possible to set the adoption thresholds to be the absolute number of neighbors infected. This would remove the dilution of edges to infected neighbors that we see in the current model with increasing r ; nodes would no longer be penalized for having large numbers of neighbors. At this point, it is unclear which choice is a better representation of real world information cascades. Second, we could implement a fourth model of variable influence. In the models of influence discussed in this analysis, the sole factor in determining the influence one node exerts on another is the distance between the influencing node and its target. However, it is also possible that the influence a node is able to exert on its neighbors is, at least partially, determined by the characteristics of that node. For example, if a person is particularly charismatic or particularly well-respected in his or her field, he or she may be more likely to influence all of his or her neighbors, regardless of how closely the two people are associated. The impact of this type of status on complex information cascades could be explored by randomly assigning an influence score to each node in the graph (See Figure 5, Model IV). In simulations, all edges originating from this node would be weighted by this influence score. It is not yet known whether individual characteristics, such as sta-

tus within a community, or relative characteristics, such as distance between individuals, play a more significant role in real world information cascades. It seems likely that both will be relevant in predicting how ideas will spread through networks. It is also unclear how best to calculate "distance" between individuals to predict how they will influence one another; geographic distance, number of shared communities (e.g. did two people go to the same high school, work at the same company, etc), number of shared interests (e.g. did two people comment on the same blog post or read the same journal article). To understand how our models of variable influence relate to real world networks, greater exploration of various distance metrics on real world cascade data is required. Third, in this analysis, we assume that each node in the graph has the same adoption threshold. This assumption is unlikely to bear out in real-world networks. Variable adoption thresholds is likely to impact cascade dynamics and is therefore worth further analysis. Fourth, the graphs generated in this analysis are symmetric. In real-world networks, however, the degree distributions follow power-laws. This means that there are likely to be some very highly connected individuals or neighborhoods and many smaller, more sparsely connected groups. Again, this alteration in network connectivity is likely to affect the dynamics of multiplex cascades. For example, if a cascade infects a highly connected individual or neighborhoods, it may very quickly spread to a large portion of the graph. To capture this, we could vary connectivity across neighborhoods (e.g. vary h from one neighborhood to another, Figure 5, Model V).

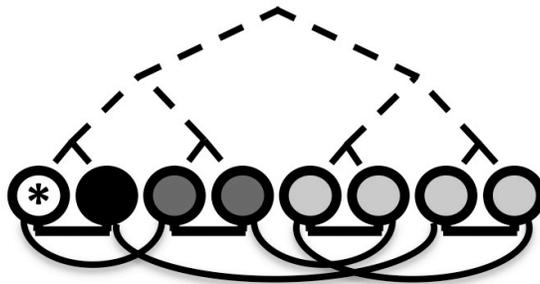
REFERENCES

- [1] Centola, D. (2010) The spread of behavior in an online social network. *Science* 329:1194-7.
- [2] Centola, D, Macy, MW, Eguiluz, VM. (2007). Cascade dynamics of multiplex propagation. *Physica A* 374: 449-56.
- [2] Granovetter, M. (1973). The strength of weak ties. *Am J Sociol* 78: 1360-1380
- [3] Granovetter, M. (1978). Threshold models of collective behavior. *Am J Sociol* 83: 1420-42.0
- [3] Kempe, D, Kleinberg, J, Tardos, E. (2003). Maximizing the spread of influence through a social network. SIGKDD.
- [3] Romero, DM, Meeder, B, Kleinberg, J. (2011) Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on Twitter. *Proc WWW*.
- [3] Leskovec, J, Huttenlocher, D, Kleinberg, J. (2010) Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on Twitter. *Proc WWW*.
- [3] Lerman, K and Ghosh, R. (2010) Information contagion: An empirical study of the spread of news on Digg and Twitter Social Networks.
- [3] Watts, DJ. (2001) A simple model of global cascades on random networks. *PNAS* 99: 5766-71.
- [3] Watts, DJ and Strogatz, SH (1998). Collective dynamics of "small-world" networks. *Nature* 393: 440-2

Model I – *Navigable, Small-World Network*



Model II – *Influence as a Function of Distance From Target*



Model III – *Randomly Distributed Influence*

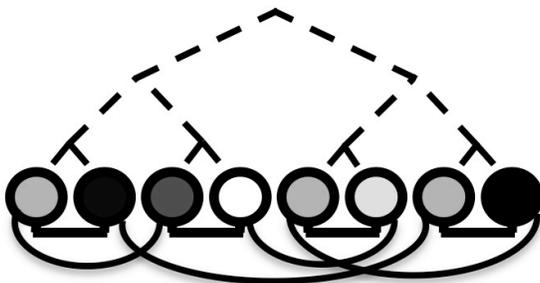


Figure 1 – *Representation of models.* Here networks are shown with tree height $H=3$, neighborhood height $h=1$, number of inter-neighborhood random edges $r=1$. The darker a node, the higher its status. In **Model II**, the target node is denoted by an asterisk.

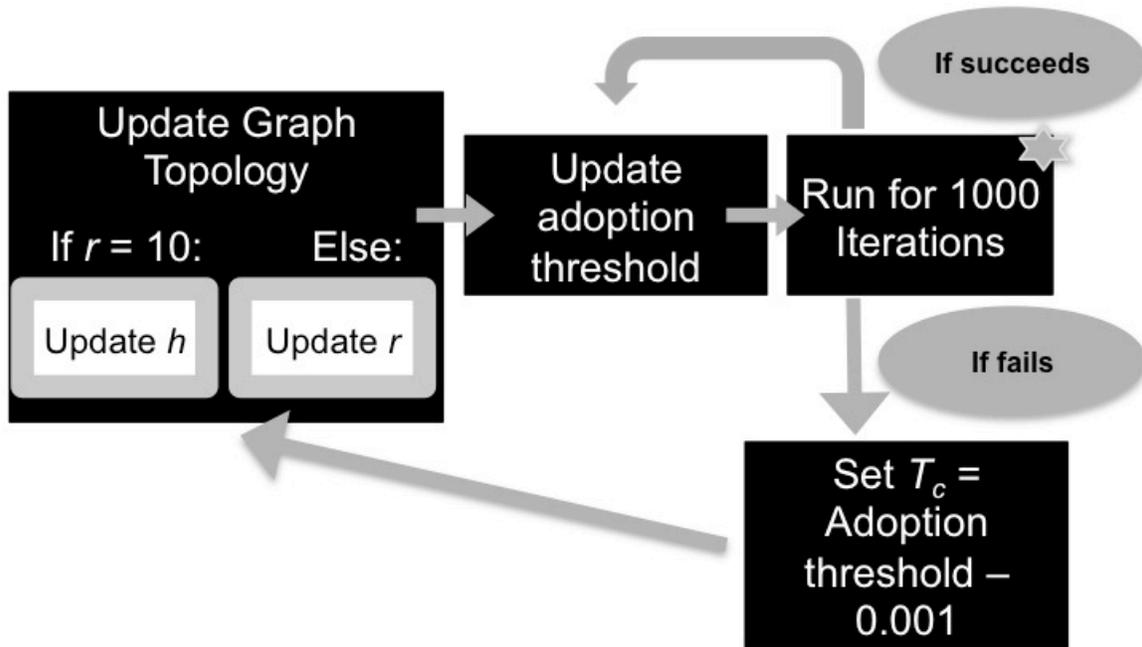


Figure 2 – *Flowchart representing calculation of critical thresholds.* Start with a specific graph topology [h initialized to 0, r initialized to 1]. Adoption threshold initialized to 0.001. Run simulation (start at star). When updating r , $r=r+1$. When updating h , $h=h+1$. When updating adoption threshold (T_a), $T_a=T_a+0.001$.

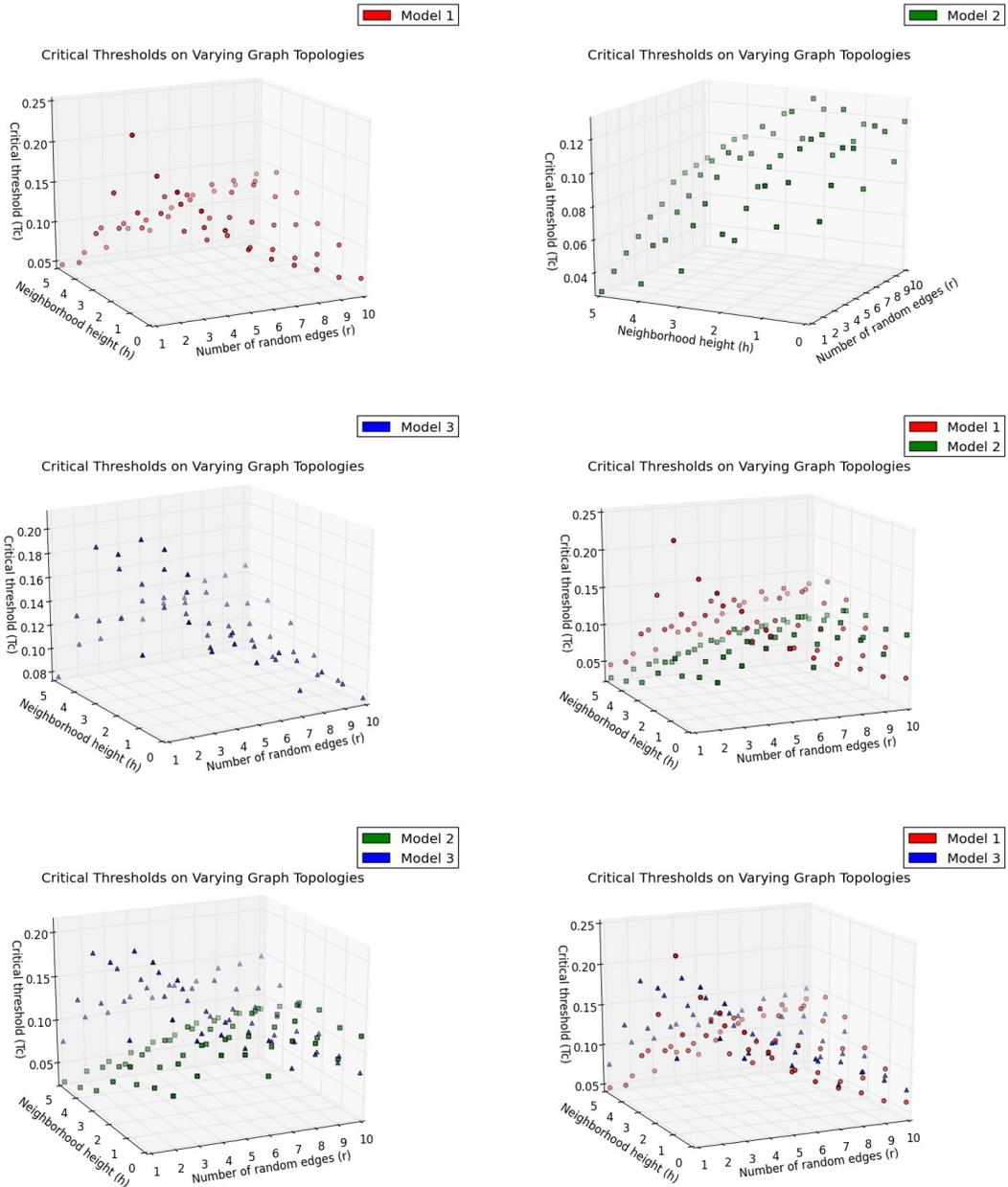


Figure 3 – Plots showing critical thresholds as a function of h and r for each of the three models of variable influence, and for each of the pairwise comparisons between the models.

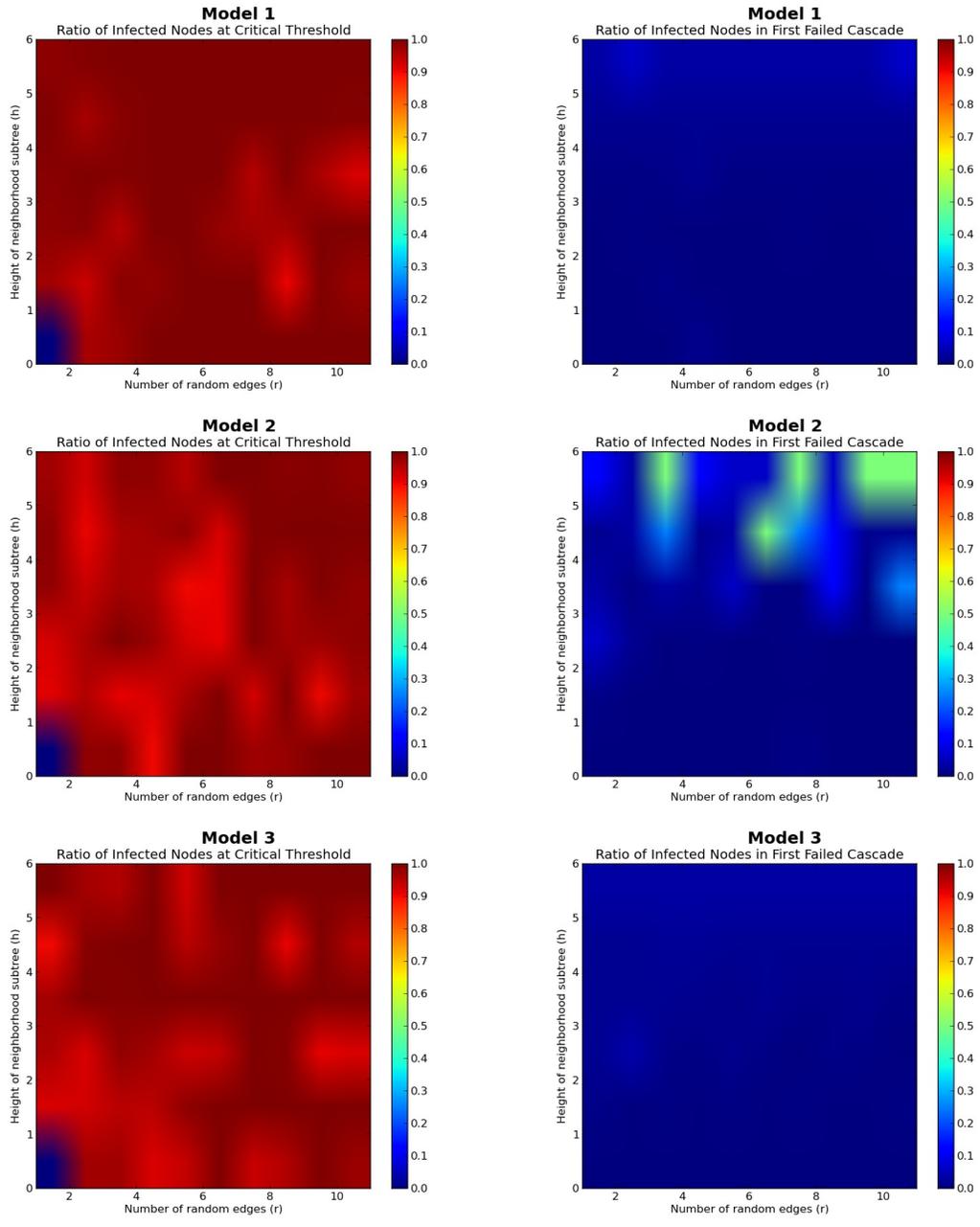
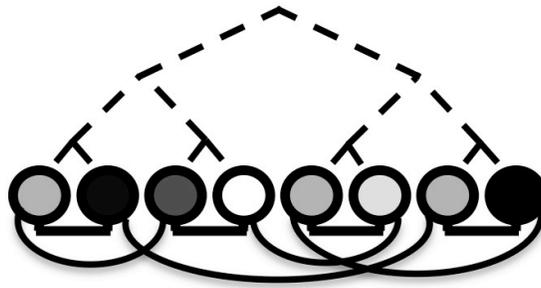


Figure 4 – Heatmaps showing the cascade ratio - in terms of percent of nodes infected - of both the last successful cascade and the first failed cascade on each network topology.

Model IV – Influence as a Function of Individual Node Characteristics



Model V – Varying Network Connectivity Across Graphs

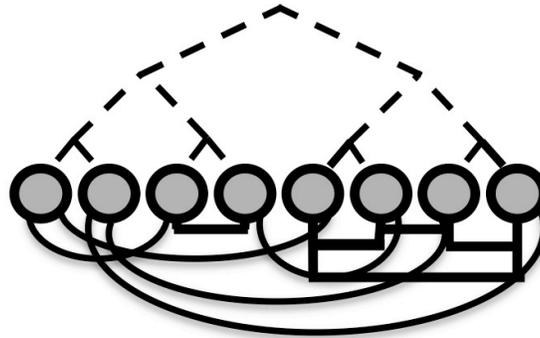


Figure 5 – Representation of potential future models. Here networks are shown with tree height $H=3$ and number of inter-neighborhood random edges $r=1$. In **Model IV** (top), status is randomly distributed across nodes. The darker a node, the higher its status. In **Model V**, neighborhood height $h=0$ for the leftmost neighborhood, 1 in the middle, and 2 on the right. In this representation, all nodes have equal status, but this model could be combined with any of the models of status described in this analysis.