The Impact of Changes in Network Structure on Diffusion of Warnings

Cindy Hui* Malik Magdon-Ismail† William A. Wallace‡ Mark Goldberg§

Abstract

Diffusion occurs in various contexts and generally involves a network of entities that interact in some way. Through these interactions, some property, e.g. information, idea, innovation, disease, etc., is diffused through the network. The flow of these information, ideas, etc. may in turn have an effect on how the entities interact. The network becomes dynamic as a result of the diffusion process. This paper presents a general model of diffusive processes in dynamic networks. The model is based on a small number of parameterized diffusion axioms. We use the model to explore how social network structure, population inhomogeneity, and seed set selection affects the diffusion process. Simulation experiments were performed on three simulated networks: grid; Erdos-Renyi; scale free; and one LiveJournal blog comment network. The context of the experiments reflect on an evacuation scenario where a warning message is diffused through the network and the goals are to propagate the message and perform evacuation. The network dynamics observed in this study are the results of the diffusion process, where nodes may leave the network some time after receiving a warning message. The simulation results indicate that inhomogeneities in the population help with the spread of the warning and leads to largest evacuation. The network structure and the seeding mechanism used in delivering the initial broadcast also have drastic impact on the efficiency of the diffusion.

Keywords: diffusion, model, simulation, dynamic networks

1 Introduction.

Diffusion occurs in various contexts and generally involves a network of entities that interact in some way. These interactions could be disease spreading, collaboration, innovation adoption, or verbal or written communications depending on the circumstances. Through these interactions, some property, e.g. information, idea, innovation, disease, etc., is diffused through the network. These networks can consist of individuals, groups, organizations, or entities like computers.

The flow of these information, ideas, etc. may in turn have an effect on the entities within the network as well as the network itself. For instance, through the diffusion of a product review, individuals may become curious and browse for addition reviews on the product, adopt or purchase the product, and/or form user groups to discuss the product.

In this paper, we present a general model of diffusion in dynamic networks. The dynamic nature of the network is driven by the local network dynamics and also in part by the diffusion that occurs through the network. The local dynamics determine the interactions between individual nodes in the network and how the diffusion may spread. The network may change due to the diffusion that occurs. Edges and nodes may appear and disappear with time. Individuals can make new friends, join social groups and introduce additional edges in a social network. On the contrary, individuals might leave the network, e.g. evacuate the network in the case of an emergency or disastrous event. The general diffusion model can accommodate various scenarios and diffusion contexts.

We use the model to explore how the structure of the network, population inhomogeneity, and seed set selection affects the diffusion process. In particular, we illustrate these concepts in the context of evacuation warnings. The network in this case represents a social network of households. We quantify the relations between household nodes based on a notion of trust, where population inhomogeneity is the product of varying trust levels between specified social groups. The property being diffused is a warning message. As warning messages propagate through the network, individual households may seek additional information, spread the information, or take an action, i.e. perform evacuation. This context demonstrates one form of network dynamics where nodes may leave the network and remove their interaction edges. The network dynamics observed in this study are the results of the diffusion process. It may reveal disruptions in the flow of information, identify areas in the social network that assist in propagating the evacuation warning, as well as identify parts of the social network that fail to evacuate.

*Dept of Decision Sciences and Engineering Systems, RPI huic@rpi.edu

†Dept of Computer Science, RPI magdon@cs.rpi.edu

‡Dept of Decision Sciences and Engineering Systems, RPI wallaw@rpi.edu

§Dept of Computer Science, RPI goldberg@cs.rpi.edu

2 Background.

Modeling information flowing through social networks is an active research area. The spread of infectious diseases and the spread of infectious ideas have common characteristics in terms of their diffusion process. For this reason, many diffusion models for studying the spread of ideas were developed based on models from epidemiology.

Many of the epidemiology models are derived from the Susceptible/Infected/Removed (SIR) model, which was formulated by Lowell Reed and Wade Hampton Frost in the 1920s [29]. The SIR model divides the population into three possible categories (susceptible, infected, and removed) that reflect the status of the individuals. Susceptible are individuals who are not infected but may become infected when they gain contact with an infected individual. Infective are individuals who are carrying the disease and have the potential to spread it. Removed are individuals who have either recovered from the disease or died, and cannot spread the disease. The model assigns a disease transmission probability based on a given average rate of contact, and assumes that all individuals are equally likely to become infected. Mathematical models can then be used to infer population average parameters such as contact rates and duration of infectious periods.

More realistic models must relax homogeneity, for example invoking a social structure for contact based spread, [29, 24, 30, 31, 28, 16, 3, 15, 9], or varying disease transmission probability [29]. Some aggregate statistics, such as mean outbreak size may be obtained analytically. In general, these models can depict the disease spreading process by tracking the average number of infected individuals and identifying individuals who are prone to become infected with the disease. Some theoretical results on optimal infection and immunization have been studied in [22, 23].

In general, the existing diffusion models in the literature focus on two types of approaches, cascade models and threshold models. The most basic models are the Independent Cascade Model and the Linear Threshold Model. The cascade models are similar to the models of the spread of epidemic diseases [28, 29, 22, 27]. In the Independent Cascade Model, each node gets a chance to influence each of its inactive neighbors with a given probability of success. If the transmission is successful, the neighbor will become active at the next time step. In general, this process continues until there are no more possible transmissions.

In the Linear Threshold Model, an individual is infected based how many of their neighbors are already infected. There is a weight on the edge between two nodes, which defines a measure of influence. Each node has a threshold value, which is drawn randomly from some specified probability distribution. This threshold determines how many neighboring nodes have to be activated before the node itself becomes active. If the sum of the weights of all active neighbors exceeds the threshold, then the node will become active [37] [12]. The cascade and threshold approaches form a basis for many diffusion models and extensions to these models have been made to study different diffusion processes [22, 26, 7, 11].

Information diffusion is a long-standing research area [33], with work in online communities becoming a very active topic recently, due to diffusion of innovations, viral marketing and the spread of computer viruses [20, 13, 17, 19, 26, 27, 32, 33, 34, 36, 37]. Most of the research uses a static network which has limitations since it does not consider the changes that may occur over time. For this reason, dynamic networks are becoming more popular in the recent research literature [11, 8, 6, 4, 1, 2, 25, 14] which study evolving communication graphs conditioned on a static social group structure.

The model of diffusive processes in dynamic networks described in this paper is motivated by the existing diffusion models. The key concepts found in the SIR models used in epidemiology and the standard threshold and cascade models are reflected in the axiomatic framework. The proposed diffusion model is a general framework and these particular models can be incorporated as special cases. In general, we define node states and each node can change from one state to another as they become infected with information. The state of the node depends on their perception of the value of information, and their trust in information sources and propagators. The nodes in the model that are in an infected and contagious state attempt to propagate information by activating links to their neighbors.

3 Diffusion Model.

The diffusion process occurs on a network whose nodes represent individuals or organizations and edges represent interactions between nodes. This network may be a directed or undirected graph. The network may also be dynamic where edges and nodes may appear and disappear with time. There may be local dynamics such as the changes at the individual node or edge level, e.g., changes in thresholds or node states, and group dynamics, e.g., social group evolution. The network structure may change as a result of the diffusion process. When diffusion occurs over a social network, the dynamics of the social network (social group evolution) determine who is interacting at each time step (e.g. [10]), which in turn determines how the diffusion may spread at that particular time step.

The dynamic social network comprises of three layers. The first is the physical network, in which two nodes are connected if they have the potential to interact. The second layer is the social network, in which two nodes are connected if they socially interact. The third layer is the layer of actual interactions which is driven by the social groups interactions in the second layer. This third layer is the dynamic interaction graph which we traditionally observe. The diffusion occurs in the third layer, using the actual realized contacts which are determined by the social group structure in the second layer. By observing a diffusion process spreading, we may infer through reverse engineering the diffusion process parameters as well as the social group dynamics [5].

This abstract model can be applied to various types of networks in different contexts, such as emergency evacuation in disastrous events or spread of ideas in a blogs. For evacuation in a remote region like Hambantota, Sri Lanka which was hit by the tsunami of 2004, the social network can be based on a neighbor network, e.g. grid-like and the evacuation warning propagates through the network. In the Blogosphere, the social network can be formed from the social groups of bloggers and the diffusion occurs over the realized communications.

3.1 Parameters. For simplicity, the entities that are diffused through the networks are referred to as messages in this paper, but in general these entities can be a disease, an idea, a rumor, a warning message, etc. Nodes may propagate a message to their neighbors in the social graph. In the model, a message is a vector of source-value pairs. Multiple types of sources may exist for a message, each with its corresponding perceived information value. The perceived value of the message may be different for each source, i.e. trustworthiness of source.

The weight on each interaction edge represents a trust value between two nodes. This value is used to quantify the social relations based on a notion of trust in information and information sources [21]. Note that in other contexts the weight on the edge may represent other concepts. For example, in the spread of diseases, the weight on the edge can reflect on how infectious of contagious the disease may be.

Each node in the network has configurable attributes. The properties of each node are updated over time as interactions occur and messages are propagated.

3.2 Diffusion Axioms. The model for diffusion on dynamic graphs is based on four axioms: Information Loss Axiom, Source Union Axiom, Information Fusion Axiom, and Threshold Utility Axiom. These axioms define the diffusion process by specifying:

1. How the value of a message deteriorates as it is propagated,
2. How the nodes update their properties based on their interactions, and
3. How the nodes handle multiple messages from different neighbors with varying sources and information values.

AXIOM 3.1. *Information Loss Axiom.*
If (S, V) is a source-value pair at node i which is propagated to node j then the source-value pair at node j is $(S, \alpha(i, j)V)$, where $0 \leq \alpha(i, j) < 1$ is the propagation loss from i to j. $\alpha(i, j)$ quantifies the trust relationship between nodes i and j.

When a message is passed from one node to another, the information value of the message is non-increasing. The information value of the message at the receiver node is a function of the social relationship between the sender and the receiver and not just a function of distance. The social relationship may be asymmetric, i.e. the trust weights on the edge may be different depending on the direction of the information flow.

AXIOM 3.2. *Source Union Axiom.*
If multiple nodes propagate a message to a node j , then the source set at j after propagation is the union of the source set which was already at j with the union of the source set arriving from the multiple nodes.

AXIOM 3.3. *Information Fusion Axiom.*

a. *If a source S_i appears in multiple incoming messages with values $V_i^1, V_i^2, \ldots$, the information from this particular source, V_i^*, is fused into the single source-value pair (S_i, V_i^*), where $\max_k V_i^k \leq V_i^* \leq \sum_k V_i^k$. The value V_i^k corresponds to the information value of source i at node k.*

b. *Suppose that node k has source set $(S_1^k, S_2^k, \ldots)$ with information values $(V_1^k, V_2^k, \ldots)$. The fused information value at the node is at least the $\max_i V_i^k$ and at most $\sum_i V_i^k$.*

There are two components to consider. First part is how to fuse information from the same source appearing in multiple messages. When a source is found in multiple messages, the combined information value for the source at the receiver node is at least the maximum of the information values for the source over all the messages and at most the sum of all the information values of the source.

The second part of the information fusion law is similar but outlines how the information from different sources at a node fuses to give an information value for that node. In general, a weighted linear combination could specify the exact nature of the fusion.

In our implementation, we use a convex combination of the sum and max according to a parameter λ. Exactly where V^* should sit in this range will depend on the nature of the diffusion, for example gossips (which spread fast) will have V^* closer to the sum. Thus, the fused information value is at least the maximum of the information values (having more information cannot hurt) and at most the sum.

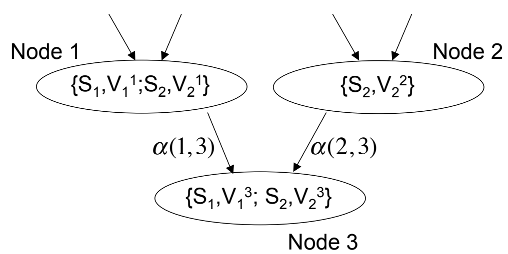

Figure 1: Message propagation

Figure 1 illustrates the information fusion process. $node1$ and $node2$ propagate a message to $node3$. The source value set at $node1$ is $\{(S_1, S_2), (V_1^1, V_2^1)\}$, and the source value set at $node2$ is $\{(S_2), (V_2^2)\}$. $node3$ is the receiver. After the message is propagated, $node3$ contains the source value set $\{(S_1, S_2), (V_1^*, V_2^*)\}$. The information value of source S_2 at $node3$ becomes $max(\alpha(1,3)V_2^1, \alpha(2,3)V_2^2) \leq V_2^3 \leq \alpha(1,3)V_2^1 + \alpha(2,3)V_2^2$. Subsequently, the information fused value of $node3$ is at least $max(V_1^3, V_2^3)$ and at most $V_1^3 + V_2^3$.

Axiom 3.4. *Threshold Utility Axiom.*
Each node has two defined threshold levels, a lower bound and an upper bound, which determine the boundaries for when the node will acknowledge the message and/or take an action. After computing the fused information value, the state of the user is determined based on whether the information value exceeds certain thresholds.

After computing the new source sets and information values, the receiver node will combine the values into a single information fused value. Table 1 summarizes the possible node states along with its corresponding behaviors in the context of evacuation warnings. The lower bound threshold lies between the disbelieved and uninformed states, while the upper bound threshold lies between the undecided and believed states. For warnings, we can assume that all nodes are initially uninformed and have not received any warning messages. When nodes fall in the undecided state, they will engage in information seeking behavior and query their neighbors in the network. The node will attempt to contact each of their neighbors. If the communication is successful, the neighbor will then send their source value set to the node. In general, the threshold levels may reflect the utility of the message and the resource requirements or potential risks associated with being in a certain states and corresponding behaviors can be defined to fit the context.

State	Description
Uninformed	Node has not received any messages and takes no action.
Disbelieved	Node has received a message but does not believe the message. Node takes no action.
Undecided	Node has received the message and is uncertain of what to do. Node will query its neighbors in the network.
Believed	Node has received the message and believes the value of the message. Node will perform an action, e.g. spread the message to its neighbors.
Evacuated	Node is no longer in the network.

Table 1: Node States for Evacuation Warnings

When the utility of the message is high, the lower bound threshold should be relatively low, since the individual node is more likely to acknowledge the message. However,if a state requires resources to be put at risk, then the threshold should be higher for that state. For example, if there are high costs and consequences associated with being in a Believed state, the upper bound threshold should be somewhat high. On the other hand, if the utility of the message is rather low, both thresholds would be relatively low. In this case, individual nodes may be more willing to acknowledge a message or take an action, i.e propagate the message, since there are low costs and consequences to the action.

4 Experiments.

The following set of experiments were performed to illustrate the concepts in our diffusion model. For simplicity, we model only the dynamics as a result of the diffusion and do not consider the dynamic nature of the network itself. In particular, we observe the effect of population inhomogeneity, network structure, and seed set selection on the diffusion process. The context of the experiments was motivated by the evacuation warnings scenario presented in [18]. In this scenario, there is a single information source who can initially connect to a specified number of nodes (seeds). The goal is to spread the warning message and have the individual nodes take action, in this case, evacuate. The assumption is that once an individual node enters the Believed state, it will spread the warning to its neighbor nodes and will leave the network after five time steps. The interest here is in demonstrating how those three factors affect the ultimate proportion of evacuated nodes.

4.1 Network Structures. We compare four social network structures, three of which are simulated networks and one which is a real blog network. The three types of simulated networks are the regular grid, the Erdos-Renyi $G(n, p)$ random network, and the scale free network. In the regular grid, most nodes have 4 neighbors. In the Erdos-Renyi network, nodes are linked randomly with an edge probability of $p = 0.00004$. The scale-free network is constructed using the Barabasi-Albert model for generating random scale-free networks using preferential attachment. The degree distribution of the resulting graph follows a power law of the form $P(k) \sim k^{-3}$. The blog network is a comment network where the two nodes are connected if one node wrote a comment on the other node's blog. This blog network was collected from LiveJournal. The network captures the comments that were made on the set of posts over the course of a two week time frame. More information on the blog network can be found in [10]. In these experiments, the edges in the networks are treated as undirected edges where messages may flow in either direction.

Network	Size	Density
Random	100,000	0.00004000
Grid	100,000	0.00003987
Scale-free	100,000	0.00003900
Blog	138,007	0.00004926

Table 2: Network Structures

We divide the population into two types of nodes (social groups) by randomly assigning each node to one of two groups A, B. In a more realistic model, nodes are more likely to communicate with other nodes belonging to the same group than with nodes of a different group.

4.2 Diffusion Scenarios. The following scenarios were modeled to incorporate population heterogeneity, in particular differences in the degree of trust between nodes. The trust values between nodes are dependent on the sender and receiver's social groups. Since the population is split into two social groups (two types of nodes), there are four types of links showing the direction of information flow between any two neighboring nodes in the network: (A to A), (A to B), (B to A), and (B to B). Each link represents the trust between the two nodes when information is transfered from the sender to the recipient. We define high trust links and low trust links. We also define trust differentials as the difference between the high and low trust links. In all the scenarios, we keep the average trust constant.

All Trust All. *All nodes have equal trust in other nodes.* In this scenario, the trust levels between all node types are the same, i.e. low trust links = high trust links. This represents a homogeneous network where everyone has the same trust in everyone else.

Like Trust Like. Like Trust Like. *Nodes have high trust in other nodes of the same type and low trust in nodes of different type.* In this scenario, individuals who belong to the same type have higher trust in each other and have lower trust in individuals of a different type. This represents a population where people have higher trust in others who are in the same group or similar to them. Such is the case, for example, in populations with ethnic groups.

All Trust A. *Links from nodes of type A have high trust and all other links have low trust.* In this scenario, recipients have higher trust in senders who belong to one specific type regardless of the recipient's node type. Recipient nodes have high trust in senders from group A and low trust in senders from group B. Such is the case, for example, in populations with status related groups, such as professionals and blue-collar workers, or educated and non-educated.

All Trust Prob Half. *All trust links have probability one-half.* This represents a population where individuals trust about one-half of the population but there is no structure in how this trust is distributed.

4.3 Seed Set Selection Strategies. In these experimental networks, there is one single information source that is randomly connected to 20% of the population. In the context of evacuation warnings, this information source represents an alerting system that can initially reach 20% of the population. This subset of nodes is referred to as the seed set. This set of nodes becomes infected with the information directly by the source at the beginning of the simulation. The information is then spread from the seed set through the network according to the dynamics of the network and the diffusive process defined by the experimental parameters. Depending on how the initial set of nodes in the seed set are selected, the diffusion dynamics observed would vary. We consider two different types of seeding mechanism: selecting a random seed set and selecting the seed set to be the high degree nodes. The set of nodes based on highest degrees can roughly correspond to influential or popular nodes of the network [35].

4.4 Experimental Parameters. In these experiments the only parameters of the model that are manipulated are: (i) the social network structure, i.e. the type of graph that is used, (ii) the trust scenario, i.e. the distribution of the trust values on the links, (iii) the trust differential, i.e. the difference in value between the high trust and the low trust edges, and (iv) the seeding mechanism. For each trust scenario, we examine two levels of trust differentials: 0.1 and 0.3, while keeping the average trust value of the links in the network at 0.75. The following parameters were kept constant. The node thresholds for all the nodes were the same. The lower bound and upper bounds values were 0.1 and 0.5, respectively. A lower bound threshold of 0.1 implies that nodes have a low tendency to disbelieve a warning message when they receive one. The range between the upper and lower threshold bounds suggests that nodes may fall in the undecided state before believing the message. The upper bound threshold of 0.5 infers that the nodes are relatively likely to believe the warning message. All nodes have high trust of 0.90 in the source. The probability of successful communication occurring on a link was 0.75. The source has a high information value of 0.95. In addition, the following assumptions were made

- The source sends the broadcast warning message at the first time step and will successfully reach all its initial connections.
- Once a node enters believer state, they will evacuate from the network after 5 time steps.
- When a source appears in multiple messages with different information values, the information is joined into a single value by taking the maximum of all the information values.
- In computing the information fused value at the node, we use the sum of the information values of all the sources present at the node with a maximum at 1.

5 Results and Discussion.

Each simulation run lasts 30 time steps and is repeated 100 times. Based on these experiments we report the average proportion of evacuated nodes at the end of the simulations for each network type. The results of the experiments using the two seeding mechanisms are shown in Table 3 and Table 4.

5.1 Network Structures and Seeding Mechanisms. The results show that the network structure and the seeding mechanism used both have an impact on the proportion of evacuated nodes. When seeding randomly in the homogeneous *AllTrustAll* scenario, evacuation diffusion was most successful in the grid network since every node has about four neighbors. The diffusion was less successful in the scale free and blog networks, where there are a few high degree nodes with large number of connections but many lower degree nodes.

In these experiments, seeding based on highest degree nodes were shown to be more beneficial in propagating the evacuation message than using a random set of nodes. The effect of the seeding mechanism is also dependent on the network structure. In the grid network, there is little difference since most nodes share the same degree. In scale free networks, seeding using high degree nodes results in a drastic increase in the proportion of evacuated nodes. Seeding based on highest degree nodes also showed improvement in the random network and the blog network, although not as drastic.

5.2 Trust Scenarios and Trust Differentials. In most cases, the inhomogeneous trust scenarios resulted in larger proportions of nodes evacuating as compared to the homogeneous trust scenario. The scenarios that gave the best results were when nodes have high trust in other nodes of the same type (*LikeTrustLike*), and when nodes have high trust in group A (*AllTrustA*). The efficiency differed by less than 1%. The similarities in those values may have been caused by the fact that there were equal number of individual nodes in both social groups A and B. Nonetheless, the results indicate that inhomogeneities in values of trust based on social groups assists in the spread of the evacuation warnings. Particularly, the higher trust differentials resulted in larger proportions of nodes evacuating which suggests that large asymmetries in trust have a positive effect.

The $AllTrustProbHalf$ scenario gives an interesting comparison to the other scenarios. In this case the

Network	All Trust All	Like Trust Like		All Trust A		All Trust Prob Half	
		0.1	0.3	0.1	0.3	0.1	0.3
Grid	0.62991	0.76277	0.89298	0.76955	0.89491	0.54261	0.81611
Random	0.60282	0.75685	0.89175	0.75950	0.88922	0.52413	0.85138
ScaleFree	0.55837	0.77486	0.88678	0.78570	0.89064	0.48724	0.85438
Blog	0.58220	0.77784	0.83711	0.77512	0.83431	0.51249	0.81227

Table 3: Proportion of Evacuated Nodes in Each Trust Scenario (Infect Randomly)

Network	All Trust All	Like Trust Like		All Trust A		All Trust Prob Half	
		0.1	0.3	0.1	0.3	0.1	0.3
Grid	0.66927	0.79705	0.90741	0.79633	0.91213	0.58002	0.84066
Random	0.76392	0.85594	0.90170	0.85663	0.90310	0.67432	0.85770
ScaleFree	0.94527	0.97670	0.98088	0.97593	0.98088	0.84784	0.91613
Blog	0.82265	0.87204	0.88492	0.87147	0.88610	0.74046	0.81910

Table 4: Proportion of Evacuated Nodes in Each Trust Scenario (Infect High Degree)

specific trust values are not homogeneous but the overall trust landscape is homogeneous. The results suggest that trust inhomogeneity by itself may not be sufficient to ensure an effective spread. With a small trust differential, this assignment of random trust levels actually performed worse than the homogeneous $AllTrustAll$ scenario. With a large trust differential, this scenario performed significantly better. However, although the proportion of evacuated nodes was greater, they did not surpass the proportions from the other inhomogeneous scenarios. The $LikeTrustLike$ and $AllTrustA$ scenarios, when a large differential was used, still had the greater proportion of evacuated nodes. These observations suggest that in addition to varying trust values, the role of social groups also help in the diffusion process.

5.3 Additional Experiments with the Scale-Free and Blog Networks. We present a slightly modified configuration and compare the observations with the previous configuration. Instead of having one source connected to 20% of the population, there are two sources, each one connected to 10% of the population. For the two sources, one source will send a warning message with high information while the warning message from the second source will have a lower information value. This configuration can represent the scenario where there is conflicting or inconsistent information being broadcasted to parts of the population. In this case, individuals of group A will receive a message with high value (e.g. mandatory evacuation) while individuals of group B receive a message with a lower value (e.g. evacuation advisory).

This is modeled by using two source nodes, where one source is connected to 10,000 nodes from group A and has high information value (0.95) and a second source is connected to 10,000 nodes from group B with information value (0.70). If a node receives both the high-valued message and the lower-valued message, the information values are fused by taking the maximum of both messages. When selecting a random seed set, the 10,000 nodes are randomly selected from within each group. Similarly, when selecting a highest degree seed set, we also select the highest degree nodes from within each group.

These experiments were performed on the scale-free network and the blog network. Tables 5 and 6 shows the average proportion of evacuated nodes at the end of the simulations over 100 repetitions. As expected, when we modify the way the source is connected to the population and the information value of the warning message, the proportions of evacuated nodes decrease. The proportions of evacuated nodes are much lower compared to the previous configuration where there was one source with high information value connected to 20% of the total population. Figure 2 compares the results observed with this configuration with the previous experiments. Trust asymmetry appears to have a positive effect on the proportion of evacuated nodes and the proportions of evacuated nodes are greater for the scenarios with large trust differentials.

When randomly selecting the initial seeds, the $AllTrustA$ scenario yields the largest proportion of evacuated nodes. When highest degree seeds are used, the $LikeTrustLike$ scenario results in a larger proportion of evacuated nodes. The proportions of evacuated nodes for the $AllTrustProbHalf$ scenario did not drop as much from the previous configuration. This observation suggests that the distribution of the node types in the network have an effect on the diffusion process. In the current configuration, the node group types are randomly assigned throughout the network. In a more realistic model, nodes from the same social group are more likely to communicate with each other and form clusters in the network.

6 Conclusions.

The findings provide interesting observations on the effects of asymmetry in trust and of trust differentials. When nodes have high trust in other nodes of their same group, the evacuation is more efficient compared to the scenario where all nodes have equal trust between them. The proportions of evacuated nodes were also greater when nodes had high trust in a certain social group. In addition, increasing the trust differentials led to larger proportions of evacuated nodes. The results presented in this paper suggest that asymmetric values of trust based on social groups facilitate the efficiency of message diffusion.

The results also illustrate the value of social groups and communities in the diffusion of evacuation warnings. The findings suggest that trust inhomogeneity by itself may not necessarily ensure an effective spread. The diffusion was more efficient when the trust inhomogeneity was based on social groups than in an unstructured way. This conclusion may also apply to disease spreading in epidemiology where evacuation corresponds to death or hospitalization. In addition, the diffusion process and its effectiveness will be dependent on the network structure and the seeding mechanism used in delivering the initial broadcast message.

These findings reveal a number of concepts for further investigation and experimentation. Network structure is known to be important in determining the extent of spread. We can simulate other network structures to see how the network affects the diffusion process, in terms of edge weight homogeneity, connectivity, and density. The simulation experiments demonstrate one form of network dynamics which is created by information diffusion. The network itself can also have dynamics of its own which are not the result of the diffusion but of social phenomenon and other factors. This is a topic for future study. As part of that research, we need to investigate the effect of the seed set size on the final proportion of infected nodes and the diffusion efficiency.

The model can be further developed to incorporate other forms of changes in node states. In the current implementation, the states generally flow in the direction from uninformed to believed. Once the node is infected and contagious (believed), it propagates information before leaving the network. One extension to the current implementation is to enable nodes to go from an infected and contagious (believed) state back to an infected and not contagious state, or even an undecided state. In the context of evacuation warnings, this situation may occur if the node receives additional information that counters the evacuation, e.g. an abort message that retracts the evacuation order. The general structure of the model allows it to be flexible with various diffusion scenarios, by the specification of the parameters of the model axioms. Another approach would be to calibrate the model parameters using data from a specific application.

Acknowledgements

This material is based upon work partially supported by the U.S. National Science Foundation (NSF) under Grant Nos. IIS-0621303, IIS-0522672, IIS-0324947, CNS-0323324, NSF IIS-0634875 and by the U.S. Office of Naval Research (ONR) Contract N00014-06-1-0466 and by the U.S. Department of Homeland Security (DHS) through the Center for Dynamic Data Analysis for Homeland Security administered through ONR grant number N00014-07-1-0150 to Rutgers University.

References

[1] R. Albert and A.-L. Barabasi, *Statistical mechanics of complex networks*, Reviews of Modern Physics, 74 (2002).

[2] L. Backstrom, D. Huttenlocher, J. Kleinberg, and X. Lan, *Group formation in large social networks: membership, growth, and evolution*, in Proc. of the 12th ACM SIGKDD Int. Conference on Knowledge Discovery and Data Mining, Philadelphia, Pennsylvania, August 2006, KDD, ACM Press.

[3] M. Barthélemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani, *Velocity and hierarchical spread of epidemic outbreaks in scale-free networks*, Physical Review Letters, 92 (2004).

[4] T. Y. Berger-Wolf and J. Saia, *A framework for analysis of dynamic social networks*, in Proc. of the 12th ACM SIGKDD Int. Conference on Knowledge Discovery and Data Mining, Philadelphia, Pennsylvania, August 2006, ACM Press, pp. 523–528.

[5] H.-C. J. Chen, M. Goldberg, M. Magdon-Ismail, and W. A. Wallace, *Reverse engineering an agent-based hidden markov model for complex social systems*, International Journal of Neural Systems, 18 (2008), pp. 491–526.

[6] C. Cortes, D. Pregibon, and C. Volinsky, *Computational methods for dynamic graphs*, tech. rep., AT&T Shannon Labs, Florham Park, New Jersey, November 2003.

Network	All Trust All	Like Trust Like		All Trust A		All Trust Prob Half	
		0.1	0.3	0.1	0.3	0.1	0.3
ScaleFree	0.43980	0.65928	0.82196	0.71313	0.83909	0.70058	0.86586
Blog	0.37017	0.58281	0.75307	0.66491	0.76504	0.66321	0.79593

Table 5: Proportion of Evacuated Nodes in Each Trust Scenario (Infect Randomly) in the Modified Configuration

Network	All Trust All	Like Trust Like		All Trust A		All Trust Prob Half	
		0.1	0.3	0.1	0.3	0.1	0.3
ScaleFree	0.69916	0.86983	0.89394	0.79235	0.80935	0.88696	0.91839
Blog	0.58643	0.73560	0.78487	0.69753	0.72730	0.75531	0.80822

Table 6: Proportion of Evacuated Nodes in Each Trust Scenario (Infect High Degree) in the Modified Configuration

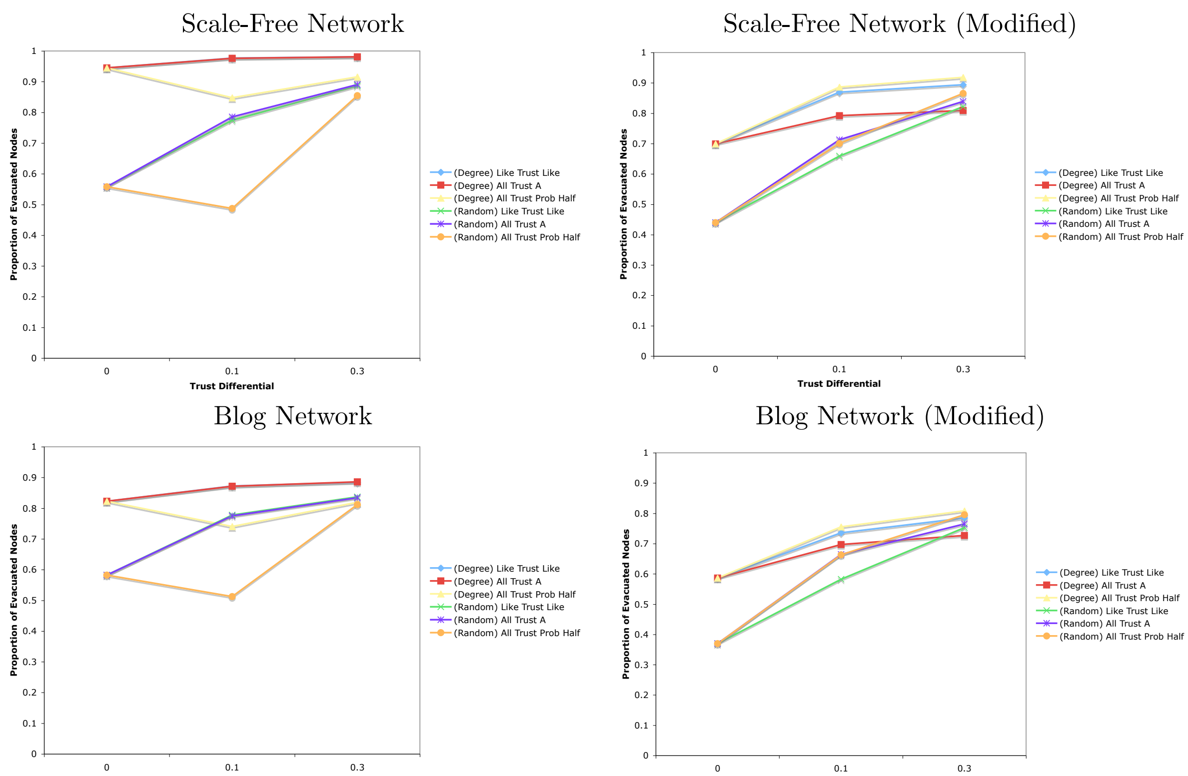

Figure 2: Comparison of the Proportion of Evacuated Nodes observed in the Scale-free and Blog Networks. The graphs on the left shows the results for the original configuration (with a single information source connected to 20% of the population). The graphs on the right shows the results for the modified configuration (with a high information value source connected to 10,000 nodes of group A and a lower information value source connected to 10,000 nodes of group B).

[7] S. Delre, W. Jager, and M. Janssen, *Diffusion dynamics in small-world networks with heterogeneous consumers*, Computational & Mathematical Organizational Theory, 13 (2006).

[8] P. Dodds and D. Watts, *A generalized model of social and biological contagion*, Journal of Theoretical Biology, 232 (2005), pp. 587–604.

[9] M. Girvan and M. Newman, *Community structure in social and biological networks*, Proc. National Academy of Science USA, 99 (2002), pp. 7821–7826.

[10] M. Goldberg, S. Kelley, M. Magdon-Ismail, K. Mertsalov, and W. A. Wallace, *Communication dynamics of blog networks*, in Proc. SIGKDD Workshop on Social Network Mining and Analysis, August 2008, p. ACM Digital Library.

[11] J. Goldenberg, B. Libai, and E. Muller, *Talk of the network: a complex systems look at the underlying process of word-of-mouth*, Marketing Letters, 12 (2001), pp. 211–223.

[12] M. Granovetter, *Threshold models of collective behavior*, American Journal of Sociology, 83 (1978), pp. 1420–1443.

[13] D. Gruhl, R. Guha, D. Liben-Nowell, and A. Tomkins, *Information diffusion through blogspace*, in WWW '04: Proceedings of the 13th international conference on World Wide Web, New York, NY, USA, 2004, ACM Press, pp. 491–501.

[14] Habiba, T. Y. Berger-Wolf, Y. Yu, and J. Saia, *Finding spread blockers in dynamic networks*, in Proceedings of the 2nd Workshop on Social Network Mining and Analysis at KDD (SNA-KDD), Las Vegas, NV, August 2008.

[15] P. Haggett, *Hybrid alternative models of an epidemic diffusion process*, Economic Geography, 52 (1976), pp. 136–146.

[16] H. W. Hethcote, *The mathematics of infectious diseases*, SIAM Review, 42 (2000), pp. 599–653.

[17] S. Hill, F. Provost, and C. Volinsky, *Network-Based Marketing: Identifying Likely Adopters via Consumer Networks*, ArXiv Mathematics e-prints, (2006).

[18] C. Hui, M. Goldberg, M. Magdon-Ismail, and W. A. Wallace, *Micro-simulation of diffusion on warnings*, in Proceedings of the 5th International Conference on Information Systems for Crisis Response and Management ISCRAM2008, F. Fiedrich and B. V. de Walle, eds., 2008, pp. 424–430.

[19] J. L. Iribarren and E. Moro, *Information diffusion epidemics in social networks*, ArXiv e-prints, 706 (2007).

[20] A. Java, P. Kolari, T. Finin, and T. Oates, *Modeling the Spread of Influence on the Blogosphere*, tech. rep., University of Maryland, Baltimore County, March 2006.

[21] K. Kelton, K. R. Fleischmann, and W. A. Wallace, *Trust in digital information*, J. Amer. Society for Information Science and Technology, 59 (2008), pp. 363–374.

[22] D. Kempe, J. Kleinberg, and É. Tardos, *Maximizing the spread of influence through a social network*, in In Proc. of the Ninth ACM SIGKDD Int. Conference on Knowledge Discovery and Data Mining, Washington, DC, 2003, ACM Press, pp. 137–146.

[23] ———, *Influential nodes in a diffusion model for social networks*, in In Proc. of the ICALP 2005, Lisboa, Portugal, 2005.

[24] M. Kuperman and G. Abramson, *Small world effect in an epidemiological model*, Physical Review Letters, 86 (2001), pp. 2909–2912.

[25] M. Lahiri, A. Maiya, R. Sulo, Habiba, and T. Y. Berger-Wolf, *The impact of structural changes on predictions of diffusion in networks*, ICDM Workshop on Analysis of Dynamic Networks,, (2008).

[26] J. Leskovec, L. A. Adamic, and B. A. Huberman, *The dynamics of viral marketing*, in EC '06: Proceedings of the 7th ACM conference on Electronic commerce, New York, NY, USA, 2006, ACM Press, pp. 228–237.

[27] J. Leskovec, A. Singh, and J. Kleinberg, *Patterns of influence in a recommendation network*, in Proc. Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD), 2006.

[28] C. Moore and M. E. J. Newman, *Epidemics and percolation in small-world networks*, Physical Review E, 61 (2000), p. 5678.

[29] M. Newman, *The spread of epidemic disease on networks*, Physical Review E, 66 (2002).

[30] R. Pastor-Satorras and A. Vespignani, *Epidemic spreading in scale-free networks*, Physical Review Letters, 86 (2001), pp. 3200–3203.

[31] L. Sander, C. Warren, I. Sokolov, C. Simon, and J. Koopman, *Percolation on heterogeneous networks as a model for epidemics*, Mathematical Biosciences, 180 (2002), pp. 293–305.

[32] D. Strang and S. A. Soule, *Diffusion in organizations and social movements: From hybrid corn to poison pills*, Annual Review of Sociology, 24 (1998), pp. 265–290.

[33] T. Valente, *Network Models of the Diffusion of Innovations*, Hampton Press, 1995.

[34] X. Wan and J. Yang, *Learning information diffusion process on the web*, in WWW '07: Proceedings of the 16th international conference on World Wide Web, New York, NY, USA, 2007, ACM Press, pp. 1173–1174.

[35] D. J. Watts and P. S. Dodds, *Influentials, networks, and public opinion formation*, Journal of Consumer Research: An Interdisciplinary Quarterly, 34 (2007), pp. 441–458.

[36] F. Wu, B. A. Huberman, L. A. Adamic, and J. R. Tyler, *Information flow in social groups*, Physica A, 337 (2004), pp. 327–335.

[37] H. P. Young, *The diffusion of innovations in social networks*, Economics Working Paper Archive 437, The Johns Hopkins University,Department of Economics, May 2000. available at http://ideas.repec.org/p/jhu/papers/437.html.