

THE IMPORTANCE OF LOCAL CLUSTERS FOR THE DIFFUSION OF OPINIONS AND BELIEFS IN INTERPERSONAL COMMUNICATION NETWORKS

JÜRGEN PFEFFER

*Institute for Software Research, School of Computer Science, Carnegie Mellon University
Pittsburgh, PA 15213, United States of America
jppfeffer@cs.cmu.edu
<http://www.casos.cs.cmu.edu/>*

KATHLEEN M. CARLEY

*Institute for Software Research, School of Computer Science, Carnegie Mellon University
Pittsburgh, PA 15213, United States of America
kathleen.carley@cs.cmu.edu
<http://www.casos.cs.cmu.edu/>*

Opinions and beliefs are essential ingredients in the diffusion of innovation. We present a framework to model and simulate diffusion processes of opinions and beliefs in interpersonal communication networks. We introduce an algorithm to create stylized networks with attributes of real world interpersonal communication networks. We also introduce a simple, but expandable model for simulating the dynamics of the diffusion processes of opinions and beliefs. We apply network multi-agent simulations to show the importance of local clusters of connected agents for keeping opinions and beliefs endemic in a social system. We also argue that due to the structure of interpersonal networks, these local clusters have the capability to propagate opinions and ideas into the entire system.

Keywords: diffusion; interpersonal networks; multi-agent simulation; opinions; beliefs; local clusters.

1. Introduction

In his fundamental work *Diffusion of Innovation*, Rogers [1995] defines *diffusion* as “a process in which, over time, information regarding an innovation is communicated over specific channels between members of a social system.” Rogers described the *social system* in which innovation is diffused as a *network model*. Social network analysis [Wasserman and Faust, (1995)] studies networks consisting of humans and their interactions. Rogers and Kincaid [1981] and Valente [(1995); (forthcoming)] focused on the role of networks and the communication flowing through networks in the context of diffusion processes. In literature [Newman *et al.*, (2006); Newman, (2003); Abrahamson

and Rosenkopf, (1997)] diffusion in networks covers primarily two areas. First, researchers analyze the propagation of diseases and epidemics [Friedman and Aral, (2001); Heckathorn *et al.*, (1999)]. Second, the diffusion of information, news and rumors [Galam, (2002); Galam, (2003); Malarz, (2006)] as well as computer viruses [Lloyd and May, (2001); Pastor-Satorras and Vespignani, (2001)] are analyzed, but often by using the definitions and notations of epidemic research.

In Rogers [1995], the failed diffusion project of changing cultural behavior of inhabitants of a Peruvian town [Wellin, (1955)] is discussed. In analyzing why the people did not adopt boiling the water to improve their health, Rogers detected conflicting cultural *beliefs* as the main reason. Beliefs are strongly connected to the diffusion of technological innovations as well as the adoption of new behavior. Guerin and Guerin [1994] enumerated “beliefs and opinions towards the technology” as one of the constraints to the adoption of innovations. Valente [1995] showed similar results for health technologies. A successful market launch for new technologies is very dependent on the customer’s opinions and beliefs about whether the new products fit their needs and are superior to other products. Similar to the spread of an innovation, the diffusion of opinions and beliefs is different to the spread of information. Opinions and beliefs have to somehow *stick* into people’s brains [Brodie, (1996); Lynch, (1998)] while information or rumors can spread through an entire social system, even if a lot of people do not believe in its truth [Galam, (2003)]. The diffusion of an opinion or a belief is therefore a story of persuasion and decision rather than knowledge. Ryan and Gross [1950] showed that the diffusion of beliefs associated with hybrid corn seed presaged and were necessary for the use of the new technology. In the context of diffusion of technology, we can see the diffusion of opinions and ideas as a forerunner of the technological propagation.

When looking at the actual diffusion of opinions and beliefs as well as innovations, personal networks between humans are getting important because people have higher trust in their friends and neighbors than in mass media [Bohlen and Beal, (1957)]. Even though mass media, especially newspapers, were expected to shape the beliefs of the masses [Lippman, (1922)], these mass-communication technologies are not sufficient for cultural change [Carley, (1995)]. The importance of interpersonal communication and the two levels of communication, from mass media to individuals and between individuals, are the result of many studies [Katz and Lazarsfeld, (1955); Rogers, (1995)]. Carley *et al.* [2009] described a simulation model which combines the spread of information by media and the spreading and altering of beliefs in social network model. In this article, however, we focus on human-to-human networks without the influence of mass media.

Real world interpersonal communication networks are dynamic, complex, and non-linear systems. To study such systems, multi-agent simulation is a widely used method [Gilbert, (2007)]. The idea of a large number of interacting agents fits perfectly to the network model of nodes, which are connected by edges. Before we are able to run simulation experiments, network models have to be created. Different models and methods have been developed [Erdos and Renyi (1959); Watts and Strogatz (1998); Barabási and Albert (1999)]. These network models can be used to study the diffusion processes of interpersonal networks. Different network models have different structure [Pfeffer and Carley, (2011)], however, and the structural characteristics have large

influence to diffusion patterns. Therefore, the algorithmically created networks have to fit the attributes of real world interpersonal networks [Hamill and Gilbert, (2009)].

In a nutshell, to analyze the diffusion processes of opinions and beliefs in interpersonal networks, we need an adequate network model of real world interpersonal networks and a congruent diffusion model to describe the propagation of opinions and beliefs in these networks. With these models, we have the ability to study which structural patterns influence which kind of diffusion positively. This is important in the context of *diffusion intervention*. For example, technology manager and marketers are interested in how to better start diffusion campaigns. Political players need assessment in detecting which protests can lead to uprisings, and which cannot because of unstable underlying networks. Siebers et al. [2010] stated the lack of “frameworks or methodologies to guide researchers and analysts through the agent-based modeling and simulation process.” This is what we are offering with this article. We are going to present a framework to model diffusion processes of opinions and beliefs in interpersonal communication networks. In section 2 we elaborate the network model for the simulation experiments. We discuss the structural attributes of networks, which are formed by personal communication between the people of a real world social system and we offer an algorithm to create stylized networks having these attributes as well as ways to calibrate this algorithm using real world data. In section 3, we discuss differences between the diffusion of infectious diseases, information, and opinions and beliefs and present the dynamic model for our simulation experiments. In section 4, we simulate these models using multi-agent simulations. We show that locally connected clusters of infected agents play an important role for the stabilization of new opinions and beliefs and their propagation through entire social systems. Section 5 discusses the results and interpretations of the simulation experiments and gives guidelines for increasing the chance of successful diffusion processes of opinions and beliefs in interpersonal communication networks. Finally, in section 6, outlooks and future work are presented.

2. Modeling the Structure of Interpersonal Communication Networks

The *members of a social system* and their connections, which are used for *communication*, are essential elements in Rogers’ [1995] definition of the diffusion of innovation. These diffusion elements can be modeled using network analytical models [Valente (1995)]. Social networks [Wasserman and Faust, (1995)] consist of a set N of agents and a set E of connections (edges) between these agents. Because the network model is independent from its content, the entities of a network are often called nodes. Nodes are connected by edges representing corresponding interactions between nodes. As nodes in our networks represent humans, we use the term agent to describe a single entity; the terms edges, connections, and links are used interchangeably. The edges in interpersonal communication networks are constructed by human-to-human interaction, which can be face-to-face, e-mail, or telephone communication. We explicitly exclude mass media communication from the considerations of this article.

Social network analysis created a vast amount of methods [Wasserman and Faust, (1995)] and theories [Freeman, (2004)]. To describe the structure of interpersonal communication networks and their dynamics we need the following definitions. Two agents a and b in a network are directly connected if there exists an edge $e(a,b)$; these

nodes are also called neighbors. Two nodes a and d are indirectly connected if there is a group of edges and nodes forming a *path* between a and d , e.g. the edges $e(a,b)$, $e(b,c)$, and $e(c,d)$. If there is no shorter path between two nodes, an indirect connection is called the *shortest path*. A single focal agent with all its neighbors is called an ego network.

Following the basic definition of networks, we have to describe the set of agents and the set of edges before we are able to analyze a network. In interpersonal communication networks, the agents are humans within a specific social system, e.g. a company, a university, or a city. These agents are normally well describable. A large volume of statistical data is available for almost any defined social system. In addition, representative surveys can be used to better describe attributes and behavior of the agents of a social system. When it comes to describing the links between the agents, the challenges are much harder. The connections of social systems with dozens or hundreds of agents can be observed or surveyed by asking the agents with whom they interact. When we try to analyze larger social systems, e.g. cities or countries, these approaches are ineligible. In these cases, networks are algorithmically created. The widest used network model is created by the *random network* algorithm from Erdos and Renyi [1959]. This algorithm takes a number of nodes and randomly connects pairs of nodes until the desired number of links is created. In the last five decades, numerous algorithms to create network models were presented. The small-world model by Watts and Strogatz [1998] and the power-law model by Barabási and Albert [1999] are the most popular ones. We call all artificial networks, which are not random networks [Erdos and Renyi, (1959)], stylized networks. These networks have different structural attributes. In the next sub-section we discuss attributes of interpersonal communication networks and present an algorithm to create networks with these attributes.

2.1. Interpersonal network models

The structure of a network is a very important parameter for the outcome of simulation experiments. Different random or stylized network models have different structural characteristics. Changing the network structure can result in different diffusion patterns. Therefore, the first question to discuss is, which networks to use for simulating diffusion processes of opinions and beliefs in interpersonal communication networks.

Hamill and Gilbert [2009] enumerated attributes of real world interpersonal networks. These attributes can be summarized into three main characteristics:

- (1) *Sparse networks with a right tailed distribution.* When looking at the number of connections of all nodes in a network, technical networks (e.g. servers in the world-wide-web) have a tendency to form scale-free power-law distributions [Barabási and Albert, (1999)]. This varies in case of networks based on interpersonal communication. Without the use of mass communication the number of people a single person can interact with is limited to a few hundred [Hill and Dunbar, (2003); Zhou *et al.* (2005)]. Consequently, large interpersonal communication networks are *sparse*. Nevertheless, the distribution of the number of neighbors is right tailed, because some agents have many more connections than the average.
- (2) *Local community clustering.* This property is known as transitivity, which describes the higher chance of a connection between two agents if these agents have a common

neighbor (“friend of a friend”) [Wassermann and Faust, (1995)]. Watts and Strogatz introduced the clustering coefficient to measure the amount of local density in a given network and also described an algorithm to generate stylized networks with high local clustering; the small-world networks [Watts and Strogatz, (1998)].

- (3) *Global closeness*. This aspect of human-to-human networks describes that the distances between all agents in such networks are small. Frigyes Karinthy [1929] illustrated the idea in one of his short novels that all the people of the world are somehow connected with a small number of steps. Four decades later, Stanley Milgram created the notion of *six degrees of separation* [Milgram (1967); Travers and Milgram (1969)] describing the same phenomenon.

Pfeffer and Carley [2011] showed that the widely used algorithms to create random and stylized networks do not meet the requirements for real world interpersonal networks claimed by Hamill and Gilbert [2009] and introduced an algorithm to create networks that cover all these attributes. Their algorithm combines different existing algorithms. The initial model is created by using the following social circle algorithm by Hamill and Gilbert [2009]. Nodes are randomly distributed in a two-dimensional space. The nodes are divided into three groups with different *social reach* describing the maximum *physical* distance in the two-dimensional space to establish a connection. Edges are now drawn by connecting all pairs of nodes (a,b) in case node a is within the social reach of node b and vice versa. Applying this first algorithmic part results in locally clustered networks with nodes connected that are physically near each other. We call these edges the *local connections*.

Pfeffer and Carley [2011] stated a lack of global closeness and links too evenly distributed between the nodes in networks created with the algorithm introduced above. To decrease the average path distances and to also increase the skewness in the networks, they suggested rewiring a small portion (~5 %) of the edges using the power rewiring function

$$f(d_k) = \begin{cases} 1 & \text{for } d_k \leq \bar{d} \\ (d_k - \bar{d} + 1)^x & \text{for } d_k > \bar{d} \end{cases} \quad (1)$$

to describe the number of times node k is in the urn to be selected as the target of a rewiring step based on the number of neighbors d_k compared to the average number of neighbors \bar{d} . These rewired edges connect local clusters on a global level independently from their *physical* distance. We call these edges the *global connections* of our networks.

To create stylized networks with the attributes of interpersonal communication networks [Pfeffer and Carley, (2011)], the basic ideas of three different network algorithms [Hamill and Gilbert, (2009); Barabási and Albert (1999); Watts and Strogatz, (1998)] are combined. We use networks created with this new algorithm for our simulation experiments in the following sections. Pfeffer and Carley [2011] also introduced a way to calibrate the structure of network models to better fit different real world social systems. The main idea of their method is to connect structural attributes of the macro level of entire networks with the micro level of ego networks. Ego networks can be collected easily using network surveys [Fischer, (1982); Burt, (1984)]. This data

can be used to calibrate the network model at the macro level because important characteristics (e.g. the distribution of the number of neighbors, the number of links in the network, or the amount of local clustering) can be projected between these two levels. We do not include empirical data to the models of this article and relinquish to calibrate our network models to keep the models simple. The interested reader is pointed to Pfeffer and Carley [2011] for further details.

3. Modeling the Dynamics of Opinion and Belief Propagation

In the previous section, we discussed the structure of interpersonal communication networks. In this section, we focus on the dynamics in these networks. Analyzing the diffusion of opinions and beliefs requires the concepts and methods of different research areas. First, the large set of definitions and terms developed by epidemic research helps us to describe different aspects of diffusion processes as well as different forms of diffusion. Second, to actually accomplish network diffusion experiments we need the concepts and methods of multi-agent simulation. Third, the *diffusion of innovation* (Rogers, [1995]) discusses different aspects of the propagation of new products or ideas. We use the *decision process* for adopting innovations introduced by Rogers [1995] and develop a *persuasion model* for opinion and belief propagation.

3.1. Modeling diffusion processes of opinions and beliefs

The terms beliefs, opinions, and attitudes are used interchangeably in many research projects. They all describe cognitive or psychological states of a person towards ideas, products, or other persons. Differences between these terms can be stated in dimension of intensity. Some, particularly psychologists, argue that beliefs are more strongly held than opinions or attitudes, and are therefore harder to change, but are also less based on facts. The strongest “version” of belief is a *sacred value*, which is not expected to change, and violations of which is expected to cause conflict [Atran and Axelrod, (2008)]. For simplicity, we do not handle different strength of opinions or beliefs in this article, but doing so would not change the models and considerations we provide.

To describe diffusion processes, concepts of epidemiology are often used. People use words like *infected*, *susceptible*, and *immune* to describe the different states of agents regarding to an *infection* with an opinion or a belief. Because these and other epidemiological terms are widely used, we want to take a closer look at these terms and their applicability to social^a diffusion processes. Moreover, we use an enumeration of these terms to cover all essential aspects of diffusion processes. In addition, we can also use the epidemic concepts to elaborate differences between different forms of social diffusion. Being aware of these differences is a precondition for more realistic diffusion models to get better results of simulation experiments. In table 1, we enumerated important epidemic concepts and their epidemic meaning and assessed the relevance for the diffusion of information as well as opinions and beliefs. We do not discuss diffusion processes of innovations here, but want to point the reader to the strong methodological

^a Social diffusion processes include the diffusion of information, innovation, opinions, and beliefs. The diffusion of information, opinions, and beliefs can also be named *immaterial* diffusion in opposite to the *material* diffusion of viruses, germs, and also goods. Innovation can be immaterial or material depending on infected persons change a behavior or start to use a specific product.

Table 1. Epidemic concepts and their applicability for the diffusion of information and for the diffusion of opinions and beliefs in interpersonal communication networks. (+) indicate concepts which can be used for social diffusions without change of meaning.

Epidemic Concept	Epidemic Meaning	Diffusion of Information	Diffusion of Opinions/Beliefs
Susceptible, Infected, Refractory, Immune	Groups to describe the agents status according to the infection	No refractory state	No refractory state
Latency/Incubation Period	Time from infection to disease outbreak	No relevance	+
Generation Time	Time from A's disease outbreak to the disease outbreak of those infected by A	No relevance	+
Alternate Host	Propagation happens not directly between agents	No distance constraints	No distance constraints
Infection	The moment a susceptible agent is infected through a contact with an infected agent	+	Persuasion as a process
Basic Reproduction Number	Average number of infected through a single agent	+	+
Predisposition, Relative Risk	The responsiveness of an agent to an infection	+	+
Contagion Index	The probability that an agent becomes infected after interacting with an infected agent, range 0-1	+	+
Incidence	Number of newly infected in a specific time period	+	+
Prevalence	The overall number of infected agents	+	+
Endemic	A disease remains permanently in the population	+	+

connections between the adoption of an innovation and the adoption of an opinion or belief (see section 3.3).

At first sight, many similarities between epidemics and social diffusion can be found. Both need persons for propagation and the agents of an observed system can be grouped identically into *immune*, *susceptible*, and *infected*. When looking more closely, important differences can be found, e.g. the *refractory* state, which marks temporarily immune agents after overcoming an infection, does not exist in social diffusion. This has implications for the selection of the diffusion model (see section 3.3). One substantial difference between epidemics and social diffusion is the possible distance between the infected and the susceptible person in the moment of infection. In case of epidemics, a direct or at least an indirect contact via an alternate host within a small physical distance is necessary. This is totally different in case of social epidemics where technical artifacts (e.g. phone, e-mail, etc) can serve as an alternate host to bridge persons in different cities or even continents. This difference between physically connected and socially connected agents is the essential argument for using network models instead of cellular automata [Gardner, (1970)] to describe social diffusion processes (see section 3.2).

Due to the lack of *latency*, information can spread with almost no delay. Therefore, the *generation time* can be reduced to some seconds. Opinions and beliefs rely on persuasion processes of an agent (see next sub-section) before this agent has the ability to infect others. When it comes to opinions and beliefs the infected agent has to mentally act before an adoption is the starting point of an active propagation. In the context of

information propagation this has little importance; information can be passed on without any considerations. In epidemic infections, this process is automated and without intention of the infected person, which is an important *advantage*. In both social cases, people have the ability to intentionally *stop* social epidemics.

As mentioned above, the *infection* with an opinion or a belief is rather a process than a certain moment. Rogers [1995] describes a persuasion process with the need of social interactions between *peers* of the infected agent as part of the adoption process of an innovation. Adopting an opinion or belief needs the same considerations and interactions. Before a positive decision is made, arguments are discussed and alternative opinions or beliefs are balanced, especially when a new opinion or belief is in conflict with existing ones. We discuss more details of the *persuasion process* in section 3.3.

3.2. Simulating networks as multi-agent models

To analyze the dynamics of interpersonal communication networks, we accomplish simulation experiments. When discussing simulation, we use the very neutral simulation definition of the Association of German Engineers (VDI). The guideline 3633 [VDI, (1996)] defines simulation as “the reproduction of a system with its dynamic processes, using a model that lends itself to experimentation, to gain insights, which are transferable to reality.” Herbert Stachowiak [1973] defines a model in his general model theory as “a shortened copy of a real world system which stands not for its own but has to be interpreted.” It is, of course, impossible to completely reverse engineer the complex social behavior of human beings. The central question in the context of modeling social systems is to cover and describe those characteristics of a real world social system, which are relevant to the research questions. Miller and Page [2007] enumerate basic principles for modeling social systems. Their first principle is “keep the model simple.” The goal of the models and the simulations in this article is therefore, not to create an extensive copy of interpersonal communication networks, but to focus on a describable and realizable part to have the ability to gain conclusions for real world based research questions.

Different simulation paradigms focus on different description levels of the interested system. To model and simulate the interactions of a large number of agents in a social system, multi-agent simulation is a widely used method [Gilbert, (2007)]. Multi-agent models consist of agents that interact with each other. Modeling social systems, an agent can be a representation for a person, a company, an organization, or a country etc. In a network multi-agent model, the agents represent the nodes of the network while the edges connecting the nodes establish the agent’s interactions.

To illustrate a diffusion process using a multi-agent network model, we created a small random network [Erdos and Renyi, (1959)] with 47 nodes and 79 edges. Every node (a circle) in fig. 1 represents an agent of our model. The edges between the nodes show the communication connections of the agents. We started our simple diffusion simulation by *infecting* the larger agent in the middle of the network with specific information. If we assume that the information is handled in every simulation step from an infected agent to all his uninfected connected neighbors, the information spreads in four steps over the entire network. To better illustrate the time points of infection, we use a color gradient from black to white. This simple example illustrates the main difference between simulation experiments, which are based on network models, and those that are

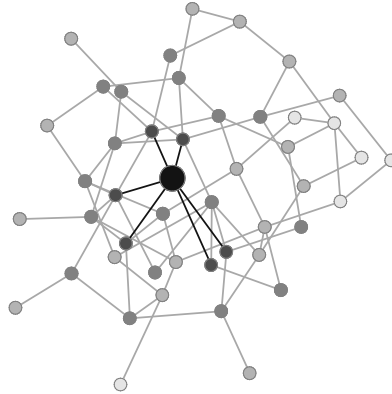


Fig. 1. Visualization of a simulation of a diffusion process in a random network.

based on cellular automata [Gardner, (1970)] or physical interactions of agents moving on a two- or three-dimensional grid [Resnick, (2001)]. In a network model, the connections between agents are important instead of their physical nearness. Looking at fig. 1, you can see that two of the physically adjacent agents of our initially infected agent are not infected until the second simulation step because of a *social distance*^b (or *path distance*) of 2.

We can also use this simple diffusion model to recapitalize the importance of fitting the structure of a network model to its underlying real world social system. If we create a couple of different models using the same 47 nodes visualized in fig. 1, but another random set of the same number of edges, the diffusion curve [Rogers, (1995)] would still be identical. If we create the edges using a different algorithm, e.g. the small-world algorithm [Watts and Strogatz, [1998]], however, the diffusion curve would be different. Conclusions based on simulation experiments can therefore only be drawn if the network model is a good fit of the underlying real world system.

3.3. Persuasion model

Most of the diffusion models of researchers from a great variety of academic fields are based on the so called “Reed-Frost Epidemic Model” [Abbey, (1952)]. Every agent in this model can be described at any given simulation step with one of the states enumerated in the first line of table 1. *Susceptible* agents can be contaminated when they get in contact with a virus or a belief. *Infected* agents carry the virus or the belief and have the ability to infect others. Agents in the *refractory* state are immune to an infection. Depending on the content of the diffusion, these states can be reached differently, e.g., a person healed from Tuberculosis can be re-infected immediately (SIS-model), other diseases have immunization periods (SIRS-model) or an agent is completely immunized for future infections after the infection dies down (SIR-model). In section 3.1, we argued that in the context of social diffusion, the refractory state is not used and normally, re-

^b The *social distance* or the *path distance* between two agents in an interpersonal network is the number of *hops* it takes between these nodes. If agent *a* is connected with agent *b*, agent *b* with agent *c*, and agent *c* with agent *d*; the social distance, e.g., from *a* to *c* is 3, from *b* to *c* is 1, and from *a* to *d* is 4.

infection with the same opinion or belief is possible. Consequently, an SIS based model is the appropriate representation to model social diffusion.

Rogers [1995] describes the decision to adopt an innovation as a process rather than a spontaneous reaction; e.g., in the hybrid corn seed study [Ryan and Gross, (1943)], the diffusion process needed several years. Even the adoption of opinions and beliefs normally take place in a shorter period of time, we assume a similar process. According to Rogers [1995], the adoption process consists of different steps. We modify these steps to describe the adoption process of opinions and beliefs:

- (1) *Knowledge*. The agent gets in contact with a new opinion or belief.
- (2) *Persuasion*. The agent starts to have a positive or negative attitude to a specific opinion or belief. Additional *infections* are needed to come to a decision.
- (3) *Decision*. The opinion or belief is accepted or rejected.
- (4) *Propagation*. The agent starts to actively propagate the opinion or belief. The agent therefore is able to infect other agents of the social system.
- (5) *Affirmation*. Positive feedback encourages the agent while negative feedback destabilizes the agent in his persuasion as well as in his propagation.

During the process described in the enumeration above, repeated interactions with infected people are necessary to proceed with the different steps. Rogers [1995] describes the persuasion process as an affective course of actions. Information is collected actively through interaction with *peers* of the social network. Because opinions and beliefs carry the risk of being rejected from family, friends, and acquaintances, the persuasion process needs *positive feedback* from other agents having similar opinions or beliefs.

Based on these considerations, we extend the SIS diffusion model to a dynamic *persuasion model*. To incorporate the different persuasion levels, the infected state is divided into two sub-states depending on whether the agent has the ability to infect others or not. As we have no immune agents in our model, there are three possible states s for every agent i , susceptible (S), infected (I), and contagious (C). For an agent i , at least one interaction with an infected agent at simulation step t increases the positive persuasion p to an opinion or a belief with a value p^+ :

$$p_t(i) = p_{t-1}(i) + p^+ \quad \text{if } \exists e(i, x) \in E \wedge s(x) = C \quad (2)$$

In table 1, we introduced another epidemic concept, which becomes a parameter in our diffusion model. The contagion index c^* is the probability that a contact with an infected agent actually leads to a transmission of the infection, e.g. a contagion index of 0.5 results, on average, in an infection through every other interaction with an infected agent. In general the probability ϕ of an actual infection of node i is a function of the number of infected neighbors $d_I(i)$ and the contagion index:

$$\phi_t(i) = 1 - ((1 - c^*)^{d_{I_{t-1}}(i)}) \quad (3)$$

Human actors are faced with innumerable beliefs and opinions, which compete with each other [Brodie, (1996)]. Therefore, in our model, beliefs and opinions do not persist forever independent of reinforcement. Without reinforcement through repeated interactions with other infected agents a belief or an opinion will die out. This is logically equivalent with a decreasing persuasion level p^- at every simulation step:

$$p_t(i) = p_{t-1}(i) - p^- \quad (4)$$

If the *level* of persuasion $p(i)$ is higher than a specific *threshold value*, the agent becomes a *promoter* (contagious state) and starts to propagate the opinion or belief. If $p(i)$ drops under the promoter threshold, the agents stops with an active propagation, but is still infected. The three possible states of the model are therefore coded in the persuasion level variable:

$$s(i) = \begin{cases} S & \text{for } p(i) = 0 \\ I & \text{for } 0 < p(i) < t \\ C & \text{for } p(i) \geq t \end{cases} \quad (5)$$

4. Simulation Experiments

Using this persuasion model and the stylized networks introduced in section 2, we accomplished simulation experiments to tackle the question of which pattern supports the diffusion of opinions and beliefs. The simulation networks consist of 2,000 nodes and every node is connected with 12.1^c other nodes on average. For the dynamics of the persuasion model the parameters are set as follows. The *positive persuasion value* p^+ is set to 0.2; an agent who is (re-)infected increases his persuasion level with this value. We also define a *negative value* p^- per simulation step with 0.1; we assume an opinion or belief, which declines over time if there is no re-infection. The *promoter threshold* t is set to 1.0 and represents the persuasion level at which an agent starts to actively promote an opinion or belief. This threshold can also be seen as the decision moment in the adoption process introduced in section 3. For the starting state of the simulation model we randomly infect 10 agents. These agents have an initial persuasion level of $p(i) = 2.0$. This level is also the maximum persuasion value for the agents. We do not make differences between weak and strong links [Granovetter, (1973)]; in every simulation step an infection can use any connection to propagate with the same probability of $c^* = 0.5$ (*contagion index*). All agents in our model are susceptible (no immune agents) and an infection is possible at any time independently from previous infections (no refractory state).

Running the simulation experiment with different, randomly selected, 10 initially infected agents, reveals an interesting outcome. In the vast majority of the experiments (79 of 100) the infection completely dies out within 20 simulation steps. What happens? In the first iteration steps of the simulation experiments, the initially infected agents infect and re-infect their neighbors. These infections, however, do not suffice to pass the promoter threshold for newly infected agents. At the same time, the persuasion levels of the initially infected agents decrease. After 20 simulation steps there are no longer any infected agents. The opinion or belief is extinguished from the population.

Looking more closely at the minority of simulation experiments with an *endemic*^d infection, we are able to identify a common pattern in the distribution of the initially infected agents. In case of connected infected agents, these agents re-infect and therefore

^c This value is equivalent to the average size of people's leisure networks collected by one of the authors of this article in an ego-network study ($N=1,000$). According to the hierarchical social network model of Zhou *et al.* [2005] this number of neighbors represents the average size of the *sympathy group* of humans.

^d An infection is called *endemic* when at least some agents stay infected in a given system.

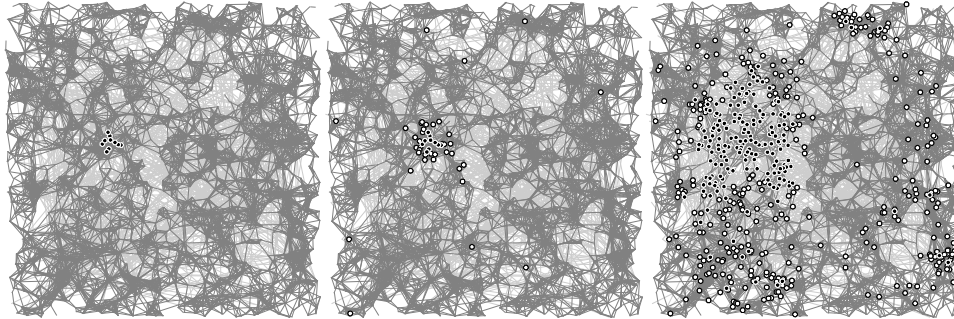


Fig. 2. Three states of a single simulation experiment of the persuasion model with *local initial infection*, initial state (left), after 5 simulation steps (middle), after 45 simulation steps (right).

stabilize each other. In addition, uninfected neighbors of infected agents have a higher probability of multiple infections within a short period of time, which is a precondition in our model to reach the promoter threshold.

Based on these observations, we can claim the hypotheses that a simulation experiment with a completely connected set of initially infected agents is the optimal constellation for diffusion in our model. To check this assumption, we set up a second simulation experiment. Fig. 2 is the visual representation of a single simulation experiment with a local seeding strategy by using the same network than in the first simulation experiments with the random selection of initially infected nodes. Black colored nodes represent infected agents with a persuasion value above the promoter threshold, white nodes are infected, but below this level. For better readability of the network figures, we draw just these two groups of infected nodes; nodes with $p(i)=0$ are not drawn. The edges are drawn in two different ways, the locally clustered edges (compare network model in section 2) are drawn dark gray and the globally rewired edges are drawn in a lighter gray to conserve the community structure in the network visualizations.

In the left picture you can see that the 10 initially infected agents were selected in a locally connected cluster. The second picture in fig. 2 illustrates the infection state after 5 simulation steps; the 10 black colored initially infected promoters are surrounded by infected zones. Contrary to the previous experiments, the opinion or belief is still endemic after 20 simulation steps and keeping the simulation running, the right picture in fig. 2 shows the infection state after 45 simulation steps. The infection has spread out in concentric circles via local edges. In addition, more *distant* agents are also infected via the global edges. Connected to a stable infection center, infected agents in the new areas have a better chance to build up their own infected environment.

To analyze the impact of local connectivity to the chance that a belief or an opinion remains endemic within a population, we setup a series of simulation experiments with different average distances between the 10 initially infected nodes. These different distances are the result of different random sets of selected nodes. Carley and Newell [1994] state that when modeling social agents not all parameters need to be investigated in order to show that a model creates discernable social results. To be able to concentrate on this single independent variable (the average distance of the initially infected nodes) we freeze all model parameters at the values that were introduced in the first simulation

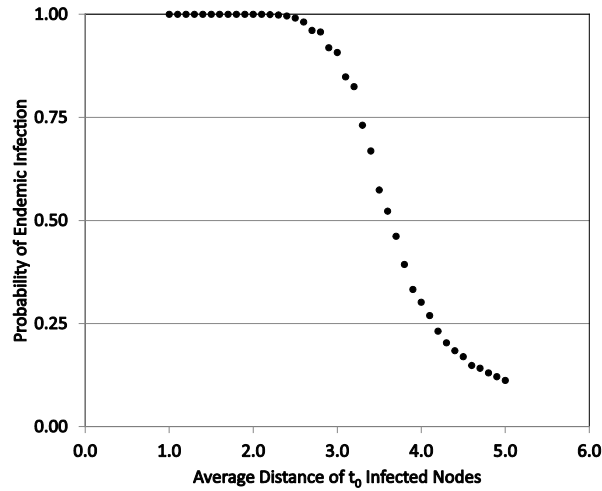


Fig. 3. Average distance of initially infected nodes and the probability that an infection stays endemic in the network after 20 simulation steps.

experiments. To also freeze all structural characteristics of the network we use a single network model for all simulations experiments.

Fig. 3 shows the result of 100,000 simulation runs. The average path distance of the 10 initially infected nodes is in the range from 1.0 to 5.0. The aggregation level for the path distance is 0.1. Therefore, every data point in fig. 3 representing an average path distance x of the initially infected nodes incorporates at least 500 simulation experiments within the range of $x \pm 0.05$. The y-axis is the probability that an infection with a specific average distance of the initially infected nodes remains endemic after 20 simulation steps.

The chart in fig. 3 shows an s-shaped negative connection between an increasing average path distance of the initially infected nodes and the probability that the infection remains endemic in the network after 20 simulation steps. The probability is almost 100 % for an average path distance of 2.5 or below. After that point the curve drops rapidly reaching a probability of 50 % at an average path distance of ~ 3.6 and 25 % at ~ 4.1 .

Discussion of the Simulation Experiments

For the simulation experiments, we were interested in structures that promote diffusion processes. In the previous sections we described a persuasion model to simulate diffusion processes of opinions and beliefs in interpersonal networks. These considerations and the results of the simulation experiments can be interpreted as follows. First, opinions and beliefs need a *critical mass* of infected people to survive. This is even more important in a social system with hundreds or thousands of competing opinions and beliefs. Second, opinions and beliefs need multi-infections and positive feedback through interpersonal communication to survive. Third, in case of multiple infected agents within a local community, the chance of a direct connection to other infected agents is higher for a single infected agent. This leads to re-infections, increases the persuasion level, and guarantees stable promoters. Fourth, in case of multiple infections within a local community, uninfected agents are more likely to get in contact with more than one infected agent. A repeated infection from various *social directions* increases the chance

of a sustainable infection. Fifth, *global* connections between agents carry the infection from an infected community into uninfected communities. Single infected agents in uninfected communities can be stabilized through these global connections.

These interpretations provide strong evidence that local clusters of infected agents are very important for a successful diffusion of opinions and beliefs. For applying these results into guidelines for assessing technology manager or policy makers in their objective to intentionally propagate opinions or beliefs, we can summarize our findings into three points:

- (a) *Start Local*. The agents that are targeted for an initial infection of an opinion or a belief should be selected from a geographical limited area, e.g., one city. If the structure of the network is known, agents belonging to the same cluster or agents who are at least somehow indirect connected with each other should be selected for an initial infection.
- (b) *Stay Local*: If it is possible to keep an infection endemic in a local area for a certain amount of time, the global connections carry the infection into other parts of the network. In other words: A large-scale fire needs steady seats of fire.
- (c) *Global Connecting*. In case the infected agents are not directly connected a priori, these connections should be established after the infection. Connecting *socially distributed*, but infected agents to a *global community* is important for stabilizing the infection.

At this point we want to add graph theoretical considerations to reveal another conclusion. Assuming a random or stylized network or a real world interpersonal network, if we randomly select a group of nodes, the average social distance between these nodes is similar than the average distance between all possible nodes in the given network. This is identical to the average distance from uninfected to infected agents. The higher the social distance between infected nodes, the more unlikely are direct or indirect re-infections between them. The dominant factor for the average social distance between all agents in an interpersonal network is the size of the network. The average distance is logarithmically connected to the network size, i.e., ten times the number of nodes increases the average distances by one [Newman *et al.*, (2006); Pfeffer and Carley, (2011)]. Large cities with millions of residents therefore have larger average social distances than smaller cities with some hundred thousand residents.

Following the conclusions presented in the previous paragraphs leads to possible implications for starting points of successful diffusions of opinions and beliefs. Smaller cities are better suited than large ones because the network structures in smaller cities support re-infections of infected agents as well as multiple infections of uninfected agents. In addition, from a marketing perspective, larger cities tend to have a larger glut on the market. Therefore, the number of competing opinions and beliefs is higher. From a policy perspective, it is interesting to connect our findings, e.g. to the uprisings in the Northern African countries and the Middle East during the Arab Spring 2011. Protests and rebellions often started in smaller cities before they spread over the entire country. We do not claim that the city size is the most important factor for the success of marketing campaigns or the formation of revolutions, but we will spend more efforts to this hypothesis in future research projects.

5. Conclusions and Outlook

In this article we elaborated a framework to model diffusion processes of opinions and beliefs in interpersonal communication networks. We enumerated the attributes of these networks and recapitulated the stylized network model from Pfeffer and Carley [2011]. We introduced a simple dynamic persuasion model and accomplished simulation experiments to study the diffusion of opinions and beliefs. A way to improve this model is to add empirical data. Pfeffer and Carley [2011] showed connections between the micro and the macro level of interpersonal communication models, which can be used to calibrate the network model with the results of social surveys. Doing so, the network model and also the diffusion model can be adjusted in particular to the actual diffusion questions. This is an important step because a network based on interpersonal communication about technical products looks different from the network consisting of the same agents, but discussing political questions. Different networks structure results in different diffusion patterns.

Independent from the details of a specific diffusion process, we were interested in general network structures that support successful diffusions. We were able to show that, when assuming the adoption of opinions and beliefs as an iterative process rather than a spontaneous event, local clusters of infected agents are essential to keep an infection in interpersonal communication networks endemic. Using the global network connections, the infection can spread out to other clusters in the second step. On the other hand, in case of an increasing average path distance between the initially infected nodes the probability of an infection to stay endemic drops rapidly. This importance of local clusters was also revealed in other studies (e.g. Helbing *et al.* [2009]). Future work has to focus on detailed aspects of these clusters, e.g. size and density, using altered network models.

The diffusion model we used for our simulation experiments in this article is a single topic model, i.e., just one opinion or belief is propagated. In real word systems, normally more than one, and often contrary, opinions and beliefs are disputed in interpersonal communication networks. These opinions and beliefs form another network on its own describing the patronages and enmities between them. When agents interact, the change of their persuasion level is dominated by the relation of their opinions or beliefs. Some interactions between agents result in mutual promotion while the persuasion level is detracting in other cases because competing opinions and beliefs restrain each other in their diffusion processes.

In the context of diffusion intervention, we offered guidelines for assessing innovation manager and policy maker to use the outcomes of our models and simulation experiments to improve their success when managing diffusion processes of opinions or beliefs of technologies, or the diffusion of political ideas. Using a persuasion model of competing opinions and beliefs, structural strategies for supporting or repealing specific market beliefs or political opinions can be developed in future studies.

Acknowledgments

This research was supported in part by ONR MURI N000140811186. The views and conclusions contained in this paper are those of the authors and should not be interpreted

as representing the official policies, either expressed or implied, of the Office of Naval Research or the United States Government. We are grateful to Peter Fleissner from the Vienna University of Technology for his interactions on a previous version of this paper. We are thankful to Thomas Zorbach from vm-people in Berlin and to Michael Zerr from Karlshochschule International University in Karlsruhe for discussing marketing and management related diffusion questions with us.

References

1. Abbey, H. (1952). An Examination of the Reed Frost Theory of Epidemics. *Human Biology*, **24**: 201-233.
2. Abrahamson, E., and Rosenkopf, L. (1997). Social Network Effects on the Extent Of Innovation Diffusion: A Computer Simulation. *Organisation Science*, **8** (3): 289-309.
3. Atran, S., and Axelrod, R. (2008). Reframing Sacred Values. *Negotiation Journal*, **2008** (7).
4. Barabási, A. L., and Albert, R. (1999). Emergence of Scaling in Random Networks. *Science*, **286**: 509-512.
5. Bohlen, J. M., and Beal, G. M. (1957). The Diffusion Process. *Special Report No. 18 (Agriculture Extension Service Iowa State College)*, **1**: 56-77.
6. Brodie, R. (1996). *Virus of the Mind: The New Science of the Meme*, Hay House, Carlsbad, CA, USA.
7. Burt, R. S. (1984). Network items and the General Social Survey. *Social Networks*, **6**: 293-339.
8. Carley, K. M. (1995). Communication Technologies and Their Effect on Cultural Homogeneity, Consensus and the Diffusion of New Ideas. *Sociological Perspectives*, **38** (4): 547-571.
9. Carley, K. M., Martin, M. K., and Hirshman, B. (2009). The Etiology of Social Change. *Topics in Cognitive Science*, **1** (4): 621-650.
10. Carley, K. M., Newell, A. (1994). The Nature of the Social Agent. *Journal of Mathematical Sociology*, **19** (4): 221-262.
11. Erdos, P., and Renyi, A. (1959). On random graphs. *Publ. Math., Debrecen*, **6**: 290-297.
12. Fischer, C. (1982). *To Dwell Among Friends*, Chicago University Press, Chicago, IL, USA.
13. Freeman, L. C. (2004). *The Development of Social Network Analysis: A Study in the Sociology of Science*, Empirical Press, North Charleston, SC, USA.
14. Friedman, S. R., and Aral, S. (2001). Social Networks, Risk-Potential Networks, Health, and Disease. *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, **78**.
15. Galam, S. (2002). Minority opinion spreading in random geometry. *The European Physical Journal B*, **25** (4): 403-406.
16. Galam, S. (2003). Modelling Rumors: The No Plane Pentagon French Hoax Case. *Physica A*, **320**: 571-580.
17. Gardner, M. (1970). The fantastic combinations of John Conway's new solitaire game "life". *Scientific American*, **223** (10): 120-123.
18. Gilbert, N. (2007). *Agent-Based Models*, Sage Pubn Inc, Thousand Oaks, CA, USA.
19. Granovetter, M. S. (1973). The Strength of Weak Ties. *American Journal of So-ciology*, **78** (6): 1360-1680.
20. Guerin, L. J., and Guerin, T. F. (1994). Constraints to the adoption of innovations in agricultural research and environmental management: a review. *Australian Journal of Experimental Agriculture*, **34**: 549-571.
21. Hamill, L., and Gilbert, N. (2009). Social Circles: A Simple Structure for Agent-Based Social Network Models. *Journal of Artificial Societies and Social Simulation*, **12** (2).
22. Heckathorn, D. D., Broadhead, R. S., Anthony, D. L., and Weakliem, D. L. (1999). Aids and Social Networks: Hiv Prevention Through Network Mobilization. *Sociological Focus*, **32** (2).

22. Helbing, D., Yu, W., and Rauhut, H. (2009). Self-organization and emergence in social systems. Modeling the coevolution of social environments and cooperative behavior. *The Journal of Mathematical Sociology*, **35** (1): 1-39.
23. Hill, R. A., and Dunbar, R. I. M. (2003). Social network size in humans. *Hum. Nature*, **14**: 53-72.
24. Karinthy, F. (1929). Chains. *Everything is Different*, ed. F. Karinthy, Atheneum. Press, Budapest, Hungary.
25. Katz, E., and Lazarsfeld, P. F. (1955). *Personal Influence: The Part Played by People in the Flow of Mass Communications*, Free Press, New York, NY, USA.
26. Lippmann, W. (1922). *Public Opinion*, Macmillan, New York, NY, USA.
27. Lloyd, A. L., and May, R. M. (2001). How Viruses Spread Among Computer and People. *Science*, **292**: 1316-1317.
28. Lynch, A. (1998). *Thought Contagion, How Belief Spreads Through Society: The New Science of Memes*, Basic Books, New York, NY, USA.
29. Malarz, K. (2006). Gossip in Random Networks. *Acta Physica Polonica Series B*, **37** (11): 3049-3058.
30. Milgram, S. (1967). The Small World Problem. *Psychology Today*, **1967** (5): 60-67.
31. Miller, J. H., and Page, S. E. (2007). *Complex Adaptive System – An Introduction to Computational Models of Social Life*, Princeton University Press, Princeton, NJ, USA.
32. Newman, M. E. J. (2003). The Structure and Function of Complex Networks. *SIAM Review*, **45** (2): 167-256.
33. Newman, M., Barabási, A. L., and Watts, D. J. (2006). *The Structure and Dynamics of Networks*, Princeton University Press, Princeton, NJ, USA.
34. Pastor-Satorras, R., and Vespignani, A. (2001). Epidemic Spreading in Scale-Free Networks. *Physical Review Letters*, **86**: 3200-3203.
35. Pfeffer, J., and Carley, K. M. (2011). Modeling and Calibrating Real World Interpersonal Networks. *Proceedings of IEEE NSW*, **2011**: 9-16.
36. Resnick, M. (2001). *Turtles, Termites, and Traffic Jams*, 7th Edition, The MIT Press, Cambridge, MA, USA.
37. Rogers, E. M. (1995). *Diffusion of Innovations*, 4th Edition, Free Press, New York, NY, USA.
38. Rogers, E. M., and Kincaid, D. L. (1981). *Communication networks: Toward a new paradigm for research*, Free Press, New York, NY, USA.
39. Ryan, B., and Gross, N. C. (1943). The Diffusion of Hybrid Seed Corn in Two Iowa-Communities. *Rural Sociology*, **8**: 15-24.
40. Ryan, B., and Gross, N. C. (1950). Acceptance and diffusion of hybrid corn seed in two Iowa communities. *Iowa Agricultural Experiment Station, Research Bulletin*, **372**.
41. Siebers, P. O., Macal, C. M., Garnett, J., Buxton, D. (2010). Discrete-Event Simulation is Dead, Long Live Agent-Based Simulation. *Journal of Simulation*, **4** (3): 204-210.
42. Stachowiak, H. (1973). *Allgemeine Modelltheorie*, Springer-Verlag, Vienna, Austria.
43. Travers, J., and Milgram, S. (1969). An Experimental Study of the Small World Problem. *Sociometry*, **32** (4): 425-443.
44. Valente, T. W. (1995). *Network Models of the Diffusion of Innovations*, Hampton Press, New York, NY, USA.
45. Valente, T. W. (forthcoming). Models and methods for innovation diffusion. *Models and Methods in Social Network Analysis*, eds. P. J. Carrington, J. Scott, S. Wasserman, Cambridge University Press, Cambridge, UK.
46. VDI - The Association of German Engineers (Ed.) (1996). Simulation von Logistik-, Materialfluß- und Produktionssystemen – Begriffsdefinitionen, *VDI-Guideline 3633*.
47. Wasserman, S., and Faust, K. (1995). *Social Network Analysis, Methods and Applications*, Cambridge University Press, Cambridge, UK.
48. Watts, D., and Strogatz, S. (1998). Collective dynamics of small world networks. *Nature*, **393**: 440-442.

49. Welling, E. (1955). Water Boiling in a Peruvian Town. *Health, Culture, and Community*, ed. B. D. Paul, Russell Sage Foundation, New York, NY, USA.
50. Zhou, W. X., Sornette, D., Hill, R. A., and Dunbar, R. I. M. (2005). Discrete hierarchical organization of social group sizes. *Proc. Biol. Sci.*, **272** (1561): 439-444.