

Simulating Diffusion with Conflicting Knowledge

Kenneth Joseph and Kathleen M. Carley

November, 2012
CMU-ISR-12-111

Institute for Software Research
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Kathleen.carley@cs.cmu.edu
Kjoseph@andrew.cmu.edu



Center for the Computational Analysis of Social and Organizational Systems
CASOS technical report.

This work was supported in part by the Air Force Office of Sponsored Research for examining the unintended consequence of action using multi-agent simulation FA95500910018, the Office of Naval Research through a MURI N000140811186 on adversarial reasoning, and the Army Research Institute for Network analytics W91WAW07C0063. Additional support was provided by the center for Computational Analysis of Social and Organizational Systems and the Institute for Software Research at Carnegie Mellon University. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Air Force Office of Sponsored Research, the Office of Naval Research, the Army Research Institute, or the U.S. government. We are grateful to Geoffrey Morgan, Jonathan Morgan and Michael Martin for their assistance on the ideas presented in this work.

Keywords: Agent-based modeling, Diffusion Processes, Decision-making.

Abstract

A simulation model is developed and used to better understand the diffusion of conflicting knowledge in a social network. Within the simulation, correct and incorrect knowledge diffuse simultaneously, however, simulation agents are not explicitly made aware that contradictory information exists. Instead, they may become aware of this by obtaining both correct and incorrect knowledge of the same piece of information. The agent must then determine what knowledge to accept. We implement various models of how people react to conflicting knowledge, which we refer to as intrapersonal conflict resolution strategies (ICRS). We analyze these ICRS with respect to the speed and breadth of correct and incorrect knowledge spread and how often agents find themselves holding conflicting knowledge. We also consider the effect of varying the odds of incorrect knowledge existing in the system, the network structure and the differentials in trust between agents. We find evidence that our model is consistent with many real-world notions of diffusion. Consequently, we make tentative conclusions about the effect of our various experimental conditions on the spread of conflicting knowledge in a social network. We note that conflicting knowledge spreads more slowly and persists longer than knowledge which cannot be conflicted.

Table of Contents

I. Introduction	1
II. Related Works	2
III. Experimental Description	5
A. Simulation model overview	5
B. Experimental Design	6
1) The Basic Intrapersonal Conflict Resolution Strategy (ICRS):.....	8
2) Basic ICRS Modification 1- Changing the set of acceptable alters:	9
3) Basic ICRS Modification 2- telling alters to “recheck” their information upon conflict: 10	
4) Basic ICRS Modification 3- using the concept of uncertainty:.....	10
C. Network Generation	12
IV. Results	12
A. Noise and Density	12
B. Network Size and Deviation of Trust.....	15
C. Topology	18
D. Decision Threshold	19
E. Set of Alters from which agents accept information	20
F. Decision Enhancements	21
V. Conclusion	23

1. Introduction

Information-sharing devices, in particular cellular phones, have evolved to the point where owners now have almost uninterrupted access to an endless supply of media (e.g. text messaging, video calling, Twitter) which allow for the rapid dispersion of almost any kind of information to their social contacts. While benefits have been reaped from the advent of these technologies, it does hold that because the barriers to communicating any information have been lowered, people are more likely than ever to spread information that has not been vetted, and which is potentially incorrect across their social network.

In this work, we attempt to understand the diffusion of knowledge in situations where a set of people observing information in a noisy environment obtain correct or incorrect knowledge and begin to spread it across a social network. For example, perhaps a small subset of people observe an explosion at a nearby building – with high likelihood, such people spread this information across their respective networks. And, also with high likelihood, some will incorrectly state which building the explosion was at, while others will exaggerate the amount of damage, and so on. Here, we attempt to model how such information spreads, and what people do when, for example, someone is told by some people that the explosion occurred at Building A, and by others that it occurred at Building B.

As with all computational models, we find it necessary to simplify the actions taken by people when they are faced with conflicting knowledge and must choose what to retain. We attempt to generalize this process to a series of intrapersonal conflict resolution strategies (ICRS). The ICRS implemented are not meant to optimize the amount of correct information the system retains, a goal generally realized in similar problems in artificial intelligence, machine learning and information fusion literature (for example, [3-6]). Furthermore, these ICRS are not meant to perfectly mimic the way in which humans update their beliefs in the face of conflicting information, as others have discussed [7][8]. Rather, drawing on literature from a range of disciplines, we implement a series of simplistic ICRS that allow us to investigate how we might retain general notions of how humans deal with conflicting knowledge with the lowest level of complexity possible, therefore allowing our simulations to scale to large social networks.

We focus on three basic concepts to represent how people in the simulation (simulation “agents”) might respond in situations where they may receive incorrect knowledge. First, we vary the number of other agents (or “alters”) from which an agent accepts knowledge. We then consider whether or not an agent might, when faced with differing opinions from a set of alters, ask them to “recheck” what they have told him. Finally, we consider how agents might deal with the uncertainty attached to their current knowledge. We observe the effects of varying an agent’s ICRS under various levels of redundant access to information (modeled with network density), various levels of trust between agents and the size and topology of the social network.

Our results suggest that the models we test in many cases relate well to real-world notions of how conflicting knowledge spreads in a network, and therefore show promise as a basis for a more generalized model of such spread in the future.

Because we feel our model well-represents general notions of how conflicting knowledge diffuses in the real world, we make tentative conclusions about the effects of various micro and macro-level experimental conditions on the diffusion of conflicting knowledge. We analyze these effects by considering the rate at which correct knowledge spreads, the percentage of agents receiving knowledge that hold the correct knowledge, and the raw counts of agents receiving both correct and incorrect knowledge. We find that information spread when no

conflicting knowledge exists in the system may be different in several important ways when compared to the spread of conflicting knowledge. As one might expect, in systems with conflicting knowledge, we find that tradeoffs exist between the speed at which correct knowledge spreads and the percentage of agents who obtain correct knowledge.

The rest of the paper is structured as follows: Section II gives a concise overview of previous work. Section III describes our experimental set ups. Section IV gives results from our experiment and Section V concludes the paper with remarks about the limitations of our work and a further discussion of results.

2. Related Works

A social network is defined as a set of individuals and the social relations that exist between them. Information diffusion is the process by which information (or almost equivalently, knowledge or facts) spreads throughout a social network. In general, information diffusion studies are concerned with one of two fundamental problems: either understanding which specific agents are most influential in spreading information [9] [10], or understanding general properties of how information flows through a social network [11] [12].

Studies on which agents are the most influential in a network have shown that the position of agents in a network is critical to their ability to spread information quickly [9] [13]. Such studies have also found that personal traits of the agent impact their role in the network [14]. These results suggest that an influential individual's position in the network, as well as their decisions about who to send information to, which information to believe, and how they send information all may strongly impact how efficiently it flows throughout the network.

Studies on the properties of information diffusion have shown that various aspects of the network within which diffusion occurs affect the rate of flow. Such factors include density [15], the number of agents with a disproportionately high number of connections to others [16] [17] and, consequently, the topology of the network [18]. Obviously, such findings suggest that the structure of the network has a compelling effect on how information flows through it. We use two network metrics, topology and density, to understand how differences in network structure affect the flow of correct and incorrect information. In addition, we modify the number of agents in the network, as size is generally understood to be correlated with various network-level metrics [17].

One obvious limitation of the information diffusion studies discussed above is their focus on the flow of information in general, ignoring any notion of the content of the information or the context in which it is sent. Consequently, questions focusing on the diffusion of conflicting knowledge in a network are rarely addressed. A notable exception is the work of [19], who describes a tool which allows users to specify conflicting opinions within a social network. The tool then predicts the most likely outcome of a scenario where agents must choose to side with one or more of these opinions (e.g. which candidate to vote for in an election).

Our work differs from the work of [19] in several ways. First, our study deals only with the spread of knowledge, as opposed to the spread of opinion, which as defined could depend on several bits of knowledge (or equivalently, pieces of knowledge) [20]. In addition, we are interested in understanding how the way in which agents make decisions affects the flow of information, as opposed to considering what opinion is most likely to be held by the majority of agents. Finally, we use simulation to generate results, whereas [19] treat the issue as an optimization problem. Consequently, our focus is much more heavily tied to understanding the flow of information.

One interesting question that we expected our work to address is the difference in flow between conflicting and “unconflicted knowledge”, which refers to the generalized notion of knowledge used in most diffusion studies. Literature on the diffusion of behaviors presents an interesting hypothesis for the difference between the spread of conflicted and unconflicted knowledge. Centola [21] shows that behavior spreads best in networks with many clusters, while noting that information tends to spread best when agents are more uniformly connected throughout a network. The reason behind this, it is argued, is that while people often need to be influenced by multiple others in order to accept a behavior, the same is not true of more infectious objects such as knowledge. However, while an agent may not need to obtain knowledge from multiple alters to accept it initially, an agent likely must be persuaded by several others to accept knowledge which conflicts with knowledge he currently holds. In this sense, it is possible that networks with higher densities may have lower levels of incorrect knowledge, provided an agent is more likely to obtain correct knowledge than he is incorrect. The work of [2] supports this claim, suggesting that networks which are more complex are better able to withstand information distortion in the form of misinformation.

Further support, outside the realm of networks, exists for the notion that redundant access to information may have an effect on outcomes variables dependent on information flow. Redundancy of information access has been shown, under certain conditions, to increase classification accuracy of web documents [22], an agent's trust in a piece of information [23][24], the likelihood of winning a game where an increased pool of knowledge increases odds of success [25] and accuracy on classification problems using voters from Mechanical Turk [26].

Of particular interest to our study is the work of [26], who consider how best to use a constrained set of human labelers whose abilities are initially unknown to perform a standard machine learning classification problem. First, [26] discuss the existence of a cost in adding additional human labelers, showing that redundancy of labeling can be both very cost efficient and cost inefficient, depending on how it is implemented. The cost in their study is monetary, but correlates to the cost of reducing incorrect information at the expense of quickly spreading correct information. We see similar trade-off issues in our results. Secondly, [26] note that as the number of labelers giving an answer for a specific classification problem increases, conflict is naturally more likely to occur. In a similar vein, [27] shows that high levels of task redundancy (the number of people doing the same thing) in an organization may decrease performance, in part because it leads to conflicting results. Based on this, we postulated that in a network with highly redundant access to information, knowledge will spread quickly in a stable environment, but much slower in an environment that produces a significant amount of conflicting knowledge. Finally, we make use of work in [26] on the concept of a probability-distribution based approach to how certain their classifier should be about a particular problem. We utilize this method for computing uncertainty when we represent how “uncertain” an agent is about the knowledge he holds.

While introducing the notion of uncertainty into agent decision-making in this way is somewhat novel, the concept of decision-making with conflicting knowledge at the agent and small group level has been well-studied. Early results seemed to show that conflict resulted in poor organizational performance [28] [29]. However, later results suggest that conflict, when focused solely on the task at hand and not based on other issues (such as differences in race) has been shown to lead to increased organizational performance [28] [29].

One further question answered in the literature is how highly an agent regards knowledge they already know versus new information they receive from others. In general, it is known that

agents are reluctant to change their knowledge once they have made up their minds [30]. Exceptions to this rule do exist, however, in that people who are perceived as experts are repeatedly consulted for information and may therefore be better able to change the opinions of others [30-32]. Our experiment disregards the concept of an expert in favor of simplicity, and also simplifies the idea of an agent valuing their current knowledge more than new opinions by giving past knowledge twice the weight of new knowledge coming from his alters.

As is clear, we make several simplifying assumptions about the manner in which agents seek, select, retain and pass knowledge. These decisions are amplified by the manner in which we model the diffusion of information. Two standard models exist with regards to simulating the diffusion of information through a static social network [33]. The Linear Threshold Model [11] states that agents will obtain the object being diffused only after a certain “threshold” of their neighbors obtain it. The Independent Cascade Model [12] states that, once infected, an agent has one chance to give the object being diffused to each of his alters with some probability of success.

We use the Independent Cascade Model as the baseline model of diffusion in our simulations. Specifically, we use a modified form of the model [12] where an agent has one chance to interact, but will only interact with at most one of his alters. Such a model is a simplification of the model used in Construct, a social simulation tool [30]. The limitation of this model is that the idea of a broadcast, where agents send information to multiple alters at once, is absent. This limits the number of times that agents receive opinions from multiple others on the same turn, meaning agents in our simulation rarely receive conflicting opinions from multiple alters simultaneously.

One further point of our model to be introduced is how likely it is that any two agents will communicate with each other at any given time. Agents which have more knowledge in common have been shown to be more likely to interact. Such a phenomenon, called relative expertise [20], is an extension of the concept of homophily, which states that the more similar two people are, the more likely they are to interact [35].

A concept similar to the idea of “social proximity” described above is the concept of how much one agent “trusts” another. An extensive literature exists on the meaning of trust, though there appears to be no general agreement on a definition [23]. However, two recent surveys of literature on trust [23] [24] agree on three important theories that have emerged from work on the subject. Both reviews agree that one reasonable definition of trust is “the measurable belief of Agent A in Agent B to do Service Y”. Our experiment uses a form of this definition- in particular we define trust as “the measurable belief of Agent A that Agent B is giving him the correct knowledge of information Y”. Secondly, both works agree that trust is multidimensional- it extends beyond a continuous value between two agents. Indeed, Guha et al. [36] show that trust is not a static value in a network, but rather that trust is dynamic and can actually diffuse through the network itself. As trust is defined as a static construct in our work, we accept this as a limitation, but refer below to work which similarly treats trust as a static, continuous value between two agents. Finally, as mentioned, both [23] and [24] agree that confirmation from multiple alters of information increases an agent's willingness to trust that the information is correct. We use this as the basis for the notion of agent decision-making based on thresholds.

Our usage of trust as a static, continuous value between two agents draws from past work. Hui et al. [37] consider how varying the amount of trust between two groups of agents can affect how information flows. They study this by varying the “trust differential” across two groups of agents in a simulation, changing how much each agent values information given by

members of the other group. Hui et al. [37] find that larger differentials of trust result in increased levels of diffusion. We test a significant alteration to this conclusion in our experiment by considering the effect of changing trust differentials across dyadic connections, ignoring the notion of groups. Consequently, while we find that increasing the trust differential in the network actually decreases the amount of diffusion, we do not claim that our results invalidate the findings of [37].

In keeping our model of how an agent makes a decision and retains information simplistic, we readily note that our work is much more theoretical in nature than other simulations of agent decision-making in uncertain environments. These efforts have considered situations in which agents interact with machines [38], how to assist agents who are overloaded with information [38], how agents make decisions when under different levels of anxiety [39], and how to incorporate methods from the field of information fusion into group-level decision making [40]. While our decision models are greatly simplified from these efforts, we believe that our focus on micro and macro-level properties of the social network and the flow of information within it allow us to develop new ideas about how conflicting knowledge spreads.

3. Experimental Description

3.1. Simulation model overview

We use an agent-based dynamic-network model to simulate the diffusion of conflicting knowledge in a social network. While noting that up until this point we have used the terms knowledge, information and opinion somewhat interchangeably, we now operationalize the passing of conflicting knowledge by introducing four rigid definitions that fit with the typical social network view of information diffusion.

1) *Agent*- An agent can be thought of as a person in a simulation. Agents are connected by social links, and only those agents which are connected may interact with each other. Each agent may or may not have knowledge of each piece of information in the system. These agents resolve situations where they obtain conflicting knowledge of a piece of information based on their ICRS, and can also pass their knowledge in the form of an opinion.

2) *Information*- Information is represented in our simulation as a set of nodes in the network that emit knowledge when it is requested by an agent. Each information “node” in the network represents its own unique piece of information independent of all others. We measure each of the six outcome variables for each information “source” (or node). How these information nodes are generated is discussed in detail in Section C. Note that only agents who have links to an information source can request new knowledge from it. Upon request, the agent will receive either correct or incorrect knowledge with some probability (set by experimental conditions). Agents that are not connected to an information node can only obtain knowledge of that information from other agents.

3) *Knowledge* - An agent has knowledge of each piece of information in the network. An agent’s knowledge of a piece of information can be *correct*, *incorrect*, or *unknown*. An agent decides what knowledge he has of each piece of information based on what his alters tell him and any previous knowledge he possesses.

4) *Opinion* – an opinion is how one agent expresses to another agent what his knowledge is. Agents pass their opinion on a piece of information based on their knowledge, and can therefore pass an opinion of *correct* or *incorrect*. These opinions are used by the receiving agent to determine whether or not his knowledge of a piece of information should change.

Our simulation utilizes a turn-based system, where at each turn of the simulation each piece of information diffuses through the network of agents. The simulation ends when no piece of information can be passed within the network, which may occur relatively quickly, as an agent will only attempt to interact with other agents during the turn immediately after his knowledge changes (e.g. from *unknown* to *incorrect*).

During such a turn, the agent will pick at most one other agent to interact with, according to a well-defined CDF generated by using the concept of relative expertise [20].

5) *Interaction likelihood*- the interaction likelihood, IL , of two agents x and y which have a social link (i.e. the pair $\langle x, y \rangle$ exists in the set of edges \mathbf{E}) is computed across all information nodes i as:

$$IL_{\langle x, y \rangle \in \mathbf{E}}(x, y) = \frac{\sum_{i \in I} KnowledgeInCommon(x, y, i) + 1}{1.5 * (\sum_{z \langle x, z \rangle \in \mathbf{E}} \sum_{i \in I} KnowledgeInCommon(x, z, i) + 1)} \quad (1)$$

Equation (1) states that an agent's odds of interacting with an alter (an agents to whom he is linked) increase with the amount of knowledge the two share, as proportional to the amount of knowledge that agents shares with all other agents. If an agent shares no knowledge with any alter, the odds of him interacting with each are the same. Note that we multiply the denominator of the interaction likelihood by 1.5 to allow for the possibility that an agent does not interact with any other agents when it is his turn to interact (this occurs with probability 1/3).

3.2. Experimental Design

We measure the impact of varying each agent's ICRS, network density, topology and size, random perturbations of agent trust, and the levels of incorrect knowledge in the system using seven different outcome variables. The outcome variables are given in Table 1 below- we use the mean values of each variable across all information nodes. Unless conditions where no noise in the system are explicitly shown or discussed, these means are *compiled only over the set of conditions where noise exists in the system*.

Table 1. A list of the outcome variables recorded for each piece of information simulated, along with a description.

Outcome Variable	Description
Total Correct	number of agents having correct knowledge
Total Incorrect	number of agents having incorrect knowledge
Total Receiving	number of agents having any knowledge
Simulation Steps	number of simulation steps before no more opinions are passed
Total Conflicts	the number of "conflicts"
Correct Spread Rate	average number of agents obtaining <i>correct</i> information per simulation step
Percent Correct	The number of agents having correct knowledge over the total number of agents <i>who obtained knowledge</i>

Table 2. An outline of the basic experimental design.

Variables	Number of Conditions	Values
Network Topology	2	Scale Free, Erdős-Rényi
Network density	4	.01, .03, .05, .07
Number of agents	3	50,150,250
Percentage of facts from information sources that are incorrect (noise level)	3	0%, 20%, 40%
Standard Deviation of Trust	3	0, .15, .3
Decision Threshold for agents	5	.5, .6, .7, .8, .9
Alters from which agents will accept information	4	All, All Trusted, Only Most Trusted, Mixed
Agent able to “recheck” information?	3	Yes, No, Mixed
Agent Use of Uncertainty	$3 + (4*3*3) = 39$	“Pass Uncertainty Along” (Yes, No and Mixed) “Use Uncertainty Threshold” (--“Check threshold before changing belief” (Yes, No and Mixed) --“Tell others to recheck based on uncertainty threshold”- (Yes, No and Mixed) Threshold levels- 0, 0.15, 0.3, 0.45
Overall Design	$2 \times 4 \times 3 \times 3 \times 3 \times 5 \times 4 \times 3 \times 39$ = (approx.) 505,440	

The full set of experimental conditions is listed in Table 2. Network density serves as a proxy for the notion of how redundant access to information is in the network. Note, of course, that increasing network density has other effects as well, most importantly increasing the speed of information spread in the network at large. When analyzing our results we use exact densities, as the network generation algorithms utilized produce networks that vary slightly around our experimental values.

The percentage of knowledge emitted from information sources that are incorrect represents the amount of noise in the information gathering environment. Indeed, throughout the rest of the report, we use the term *noise* to represent this percentage. Note that we test the condition where no noise exists in the system- this allows us to compare the spread of conflicted and unconflicted knowledge in our model. The standard deviation of trust is representative of the variance of trust at the dyadic level in the network – a standard deviation of zero means that each agent trusts all alters, while any deviation represents a random seeding of trusts from a normal distribution (with the standard deviation listed) across each pair of connected agents.

At the agent level, we modify the ICRS that agents may utilize. We present a basic model of how an agent reacts when he is confronted with conflicting knowledge, and then introduce three further levels of complexity, with each level building upon the previous one.

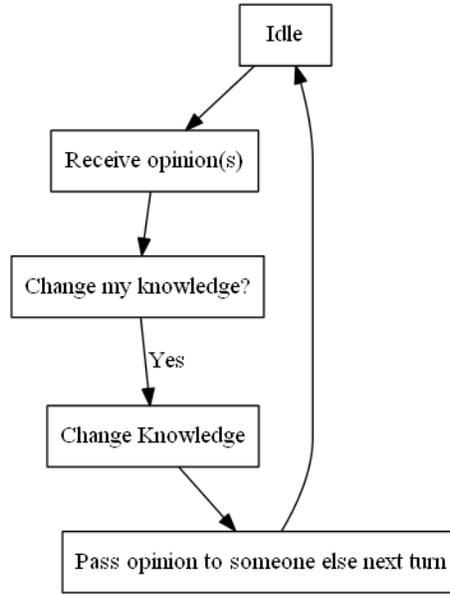


Figure 1. The Basic ICRS. Each agent in all simulations was seeded with this decision model, though where appropriate agents were allowed further capabilities, as described in later sections

3.2.1 *The Basic Intrapersonal Conflict Resolution Strategy (ICRS):*

Figure 1 above represents the basic ICRS used by each agent. On any turn, an agent may receive zero or more opinions about zero or more pieces of information. For each piece of information, upon receiving one or more opinions, the agent will decide whether or not to change his knowledge by utilizing the procedure above. Note that while the agent utilizes this strategy in processing all information, he will trivially accept any information which is not conflicted. In the basic ICRS, the agent decides what knowledge to retain using a two step process. First, he sums the correct and incorrect opinions he has received in combination with any knowledge he currently has. Then, he determines if what he now knows, either correctly or incorrectly, crosses some threshold (deemed his *decision threshold*) of his total knowledge. We use four different threshold values to understand the effect of how selective agents are about accepting new knowledge. If the percentage of neither correct nor incorrect knowledge crosses the decision threshold, the agent comes into what we call a state of *conflict*. Throughout the rest of the report, we explain the result of an agent evaluating his knowledge against his decision threshold as whether or not he *passed his decision threshold*. Note that an agent processes information he has received during the same turn on which he receives it. In this sense, there is no “processing” period for an agent to spend thinking about what knowledge to accept. Such a concept is no doubt reasonable, and may be explored in future work.

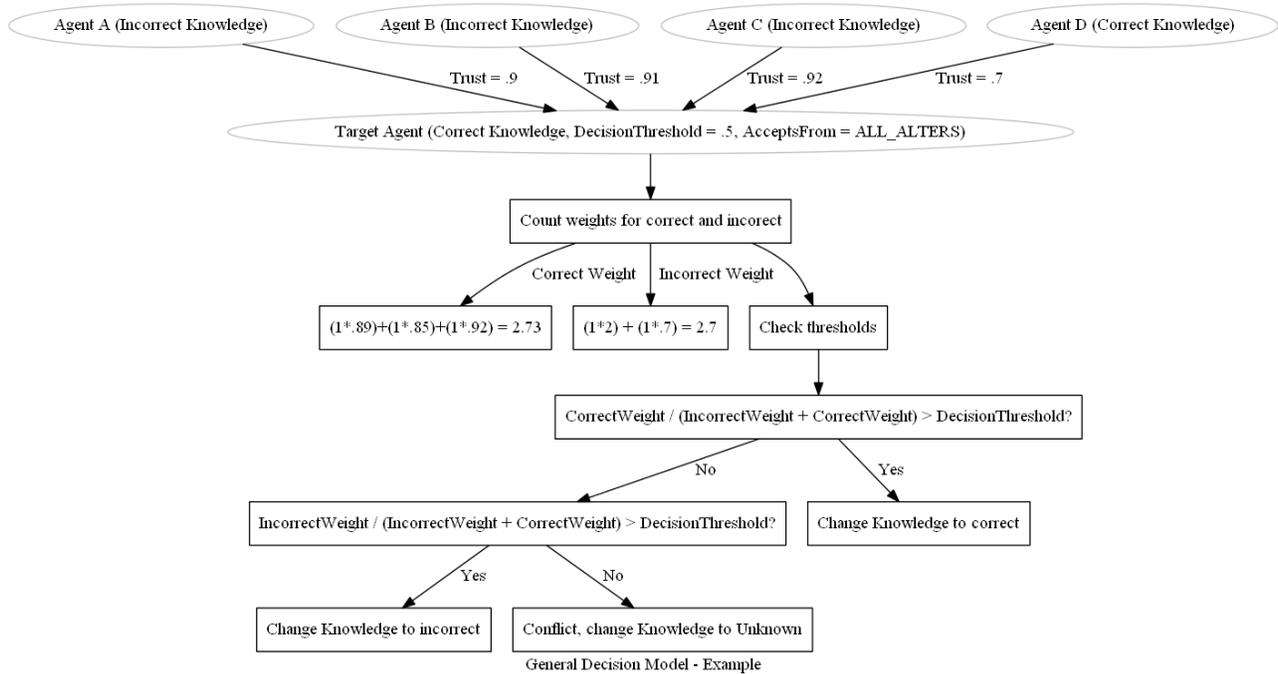


Figure 2 An example of how an agent using the Basic ICRS resolves a situation where he must decide between two conflicting viewpoints of the same piece of information.

Figure 2 gives an example of an agent using the general decision model. As you can see, although it is obvious to us what information is correct, the agent actually changes from having the correct knowledge to incorrect knowledge of this piece of information. Also note in Figure 2 that we explicitly state the target agent’s decision threshold and from which alters he accepts information (the “AcceptsFrom” parameter).

3.2.2 *Basic ICRS Modification 1- Changing the set of acceptable alters:*

We consider three different sets of alters an agent might accept opinions from when making a decision. These sets are listed in Table 3 – in the experiment, we consider networks where all agents accept from the same subset of alters, as well as networks where these strategies are mixed.

Table 3 A list of the different experimental conditions for what set of alters an agent will accept information from. Note that, because trust is a random variable, these sets amount to a random subset of alters.

Set of alters from which agents accept information (Abbreviation)	Description	Percentage of alters that each agent will accept information from
All (A)	Consider opinions passed by any alter	100%
All Trusted (AT)	Consider only opinions from alters trusted equal to or more than the mean	~50%
Most Trusted (MT)	Consider only information from most trusted alter	$\sim 1 / (\text{density} * \text{number of agents in the network})$

3.2.3 Basic ICRS Modification 2- telling alters to “recheck” their information upon conflict:

The second level of complexity tested and added to the basic ICRS model is that an agent may, when faced with a conflict, request that the alters he received opinions from most recently recheck their knowledge of the piece of information in conflict. In doing so, the agent in conflict begins a recursive process back towards the actual information node, where each agent told to recheck his knowledge tells the agents from whom he received that knowledge most recently to recheck their information, and so on, back to the information source. Note, however, that this process is subject to the same rules of interaction given in Equation (1), and therefore it is unlikely that all conflicts trace all the way back to an agent connected to the information source.

There are two other important points to note in our implementation of this process. First, when agents are told to “recheck” their information, their knowledge of that information changes to *unknown*. This knowledge can then be changed by any new opinions received – the agent does not necessarily have to wait for opinions from the agents he tells to recheck. Secondly, an agent who is already in the process of rechecking his knowledge will not do it again until his knowledge has changed from the *unknown* state.

3.2.4 Basic ICRS Modification 3- using the concept of uncertainty:

Thus far, the assumption has been made that an agent who barely passes a decision threshold and an agent who unanimously passes a decision threshold are equally happy to adjust their knowledge of a piece of information. In real life, of course, this is hardly the case – for example, we are much more certain about what we know when we are unaware of any alternatives. This naturally extends to the idea of adding some notion of “uncertainty” in the way agents make their decisions. We operationalize the amount of uncertainty that an agent has by following [26] and computing uncertainty as a continuous value on the range [0, 0.5] using a regularized incomplete beta function. The calculation used is shown below:

$$I_{DecisionThresh}(CorWeight, IncorWeight) = \frac{1}{beta(CorWeight, IncorWeight)} * \int_0^{DecisionThresh} t^{CorWeight-1} (1-t)^{IncorWeight-1} dt \quad (2)$$

$$uncertainty = \min(I_{DecisionThresh}(CorWeight, IncorWeight), 1 - I_{DecisionThresh}(CorWeight, IncorWeight)) \quad (3)$$

Once obtaining this uncertainty value for each agent, it can be used in various ways. We choose to operationalize how agents use the uncertainty they have about a given piece of information in two different ways. The first, which we term *passing uncertainty*, occurs when agents express their uncertainty by multiplying any opinion they pass by the inverse of their uncertainty. For example, if Agent A in Figure 2 were to have an uncertainty value of .2, his opinion would be counted by the target agent as $opinionValue = 1 * trust * (\frac{1}{1-.2})$. This usage of uncertainty represents the fact that although humans may be convinced to change our own minds about a piece of information in the face of conflicting knowledge, we may express our uncertainty about it when passing it to others.

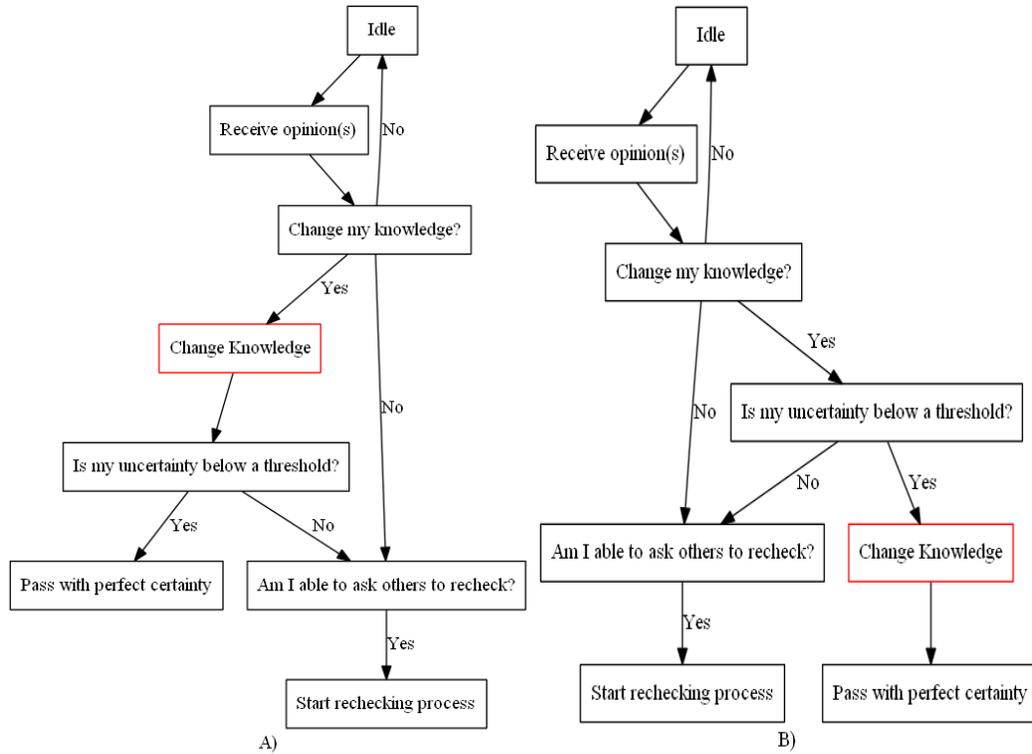


Figure 3 A) Left, The ICRS for agents who change their knowledge and then use a threshold based on the level of uncertainty they have of their knowledge in order to decide whether or not to pass it to an alter on the following turn. B) Right, Agents who first check their uncertainty threshold before changing their knowledge (and consequently before deciding whether or not to pass their knowledge on the next turn)

The second way in which agents may use the concept of uncertainty is by adding another layer of thresholding to their decision-making process, this time based on the amount of uncertainty they have. Figure 3a and 3b show two manners in which agents might use this idea of a threshold for uncertainty. In Figure 3a, agents *first* change their knowledge of a piece of information based on a decision threshold, *then* decide whether or not to pass along their information by checking their *uncertainty threshold*. In Figure 3b, agents only change their knowledge (and consequently only pass along their knowledge) *after* checking whether or not they pass both their decision and uncertainty thresholds.

In both Figures 3a and 3b, we also test the condition where an agent, even if he passes his decision threshold, might tell his alters to recheck their knowledge if he does not pass his uncertainty threshold.

3.3. Network Generation

The final networks used for propagation in the study were undirected, bimodal networks consisting of two node classes – information nodes and agent nodes, and two types of links – agent to information, and agent to agent. For the purposes of our study, we ignore any connections between information nodes, treating each information node in the system as an independent piece of information. Future work will consider the existence of dependencies and relationships between pieces of information in a system.

The network generation procedure is completed as follows. First, we generate a one-mode, stylized network of $1.25 * X$ nodes, where X is 50, 150, or 250 according to our experimental conditions. These networks are generated using algorithms which produce stylized networks adhering to the properties of one of two topologies (Scale Free or Erdős-Rényi), and have a specific overall density from our experimental conditions. After obtaining this uni-modal agent network, we randomly selected 25% of the agent nodes in the network to be changed to information nodes, giving us a bimodal network that still holds the properties of the topology we intend to study.

4. RESULTS

4.1. Noise and Density

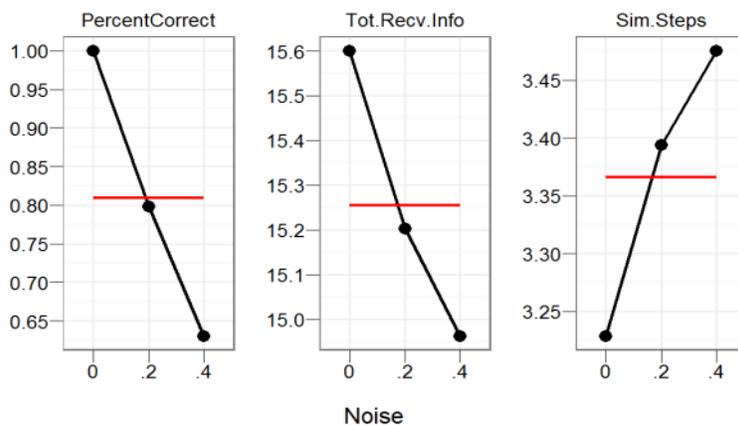


Figure 4 Noise in the system versus select outcome variables. Note that in this and all similar images, the red bar indicates average across conditions

Figure 4 shows that there is little difference between noise level and the percentage of agents having any knowledge which have the correct knowledge at the end of the simulation. Indeed, at a noise level of 0.2, the odds of an agent having the correct knowledge were actually slightly lower than the amount of noise in the system. This suggests that simple, typical actions taken by humans in a social network to resolve conflicts of knowledge may have little effect, for better or worse, on the quality of information across the network at large.

This suggests a clear danger of rapid spreading of incorrect knowledge in noisy information gathering environments where agents are not conscious or concerned over how they deal with conflicting knowledge. This necessarily implies the need for future studies which suggest better methods at both the intrapersonal and interpersonal levels for how to deal with and share knowledge in systems with conflicting information. As stated previously, we only attempt to model and understand typical human behavior – questions of how this behavior can be optimized in a large-scale network should now be posed.

Figure 4 also shows that the amount of noise does affect diffusion in a non-trivial way. In particular, noise is negatively correlated with the total number of agents receiving any knowledge, and positively correlated with the number of turns across which information persists (before no more agents can pass it). Such a finding fits the general notion that when people disagree on something, they tend to spend more time discussing the options and determining what knowledge to accept, leading to the information persisting longer in the network. On the contrary, the time spent debating what to decide typically means that the rate at which knowledge spreads decreases. These findings support the hypothesis presented in Section II, which presumed that conflicting information would spread differently through a network than “unconflicted” information.

While such a finding supports our hypothesis, we qualify it by noting that agents in our simulation who received conflicting knowledge immediately retracted to a state of “unknown” knowledge, and hence any conflicts would trivially decrease the number of agents having knowledge at the end of the simulation. On the other hand, we would have expected that as the number of simulation steps increased, information would have more time to spread and consequently should have meant more information was passed. Therefore, though the manner in which we set up the simulation had some effect on the confirmation of our hypothesis, there does exist balancing effects that model notions of how conflicting knowledge spreads in the real world.

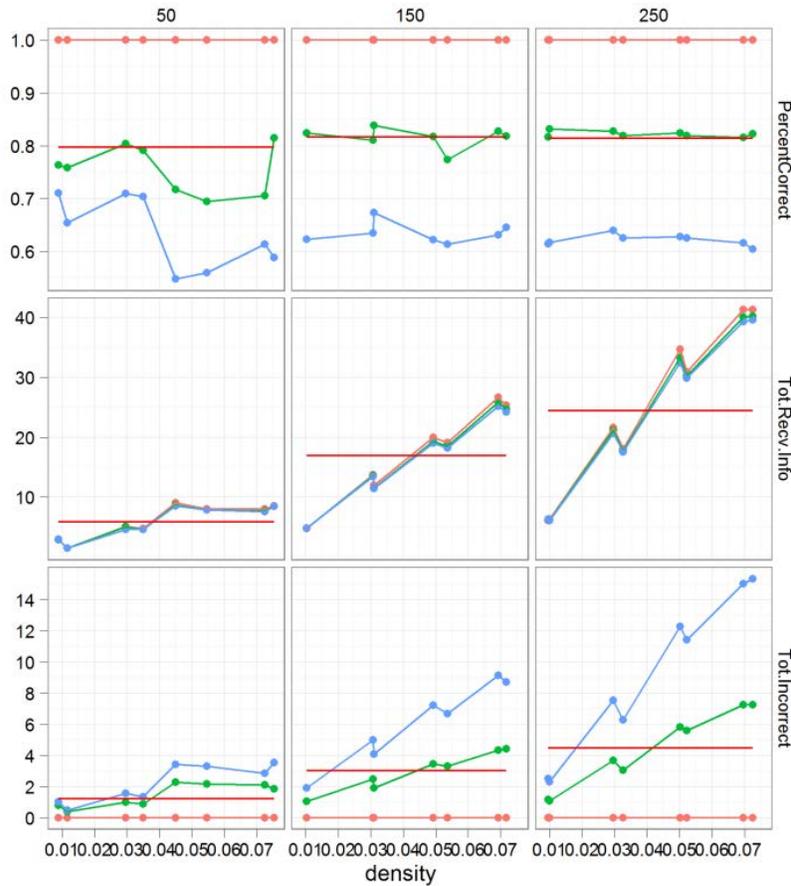


Figure 5 Density versus select outcome variables at different network sizes. The numbers on top indicate the number of agents at the different network sizes

In addition to our supposition in Section II that information would spread differently at different levels of noise, we further claimed that an increase in density would serve to both decrease the level of incorrect knowledge in the system, and to increase the disparity across noise levels of how many agents received knowledge in total. As Figure 5 shows, there is no evidence of the first claim, as we actually see a trend towards the opposite. Furthermore, while we do see evidence of the latter- the points are farther apart at larger densities across the larger two network sizes-the differences are slight.

While it is quite possible that these trends do not actually exist, we also believe that our simulation might under-represent the effect of increasing density on both the level of incorrect knowledge and number of agents holding any knowledge at the end of the simulation. Both hypotheses are based on the notion that people in a community are constantly passing and receiving large amounts of knowledge. The more dense the community, we reasoned, the more knowledge that would be flowing, which would result in more conflicts and hence lower spread and less incorrect knowledge. However, what we neglected to realize is the way in which information is passed in our simulation does not function in exactly this manner. Specifically, because agents did not retain any history of previous opinions received aside from their current knowledge, and each agent could only speak with one other agent per turn, there was little chance that the amount of information agents received on any turn would increase heavily at the density increases we used in our experiment. Because real-world social networks rarely grow

above the densities we used, it would make little sense to experiment with this effect at higher levels.

This finding is indicative of the fact that for density to greatly affect disparity across noise levels with regards to our two hypotheses, either agents should retain some knowledge of previous events (allowing them to build up opinions they have received that might contrast with their current knowledge) or agents should be allowed to talk to multiple alters at each turn. Of the two, we find the second to be a slightly more compelling alternative. Allowing an agent to speak with multiple others on each turn relates well with the concept of a broadcast, where a person is able to relay information to several agents his ego network simultaneously. In today’s world, where communication is increasingly done on a multi-way scale (e.g. email, Tweets), it is possible that this may help to create a more real-world view of communication.

4.2. Network Size and Deviation of Trust

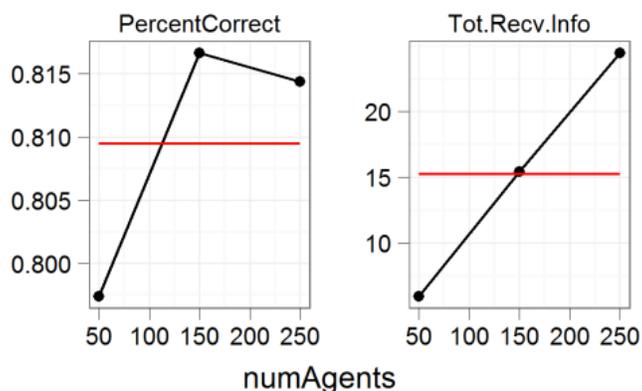


Figure 6 Network size (numAgents) versus select outcome variables

Figure 6 shows that smaller networks, on average, had a lower percentage of agents holding the correct knowledge of any given piece of information at the end of the simulation. Furthermore, we see that there was little difference, comparatively, between medium sized and larger networks. This non-linearity suggests that there may exist some “ideal” size for an organization facing noise in their information gathering environment. In addition, as Figure 7 below shows, different levels of noise may affect this ideal size – we see that smaller networks tend to do better at lower levels of noise, while larger networks were much more effective at moderate to high levels of noise. While interesting, we note that one reason why we see such results is that the percentage of agents who obtain correct knowledge is, of course, highly correlated with the total number of agents that obtained any knowledge.

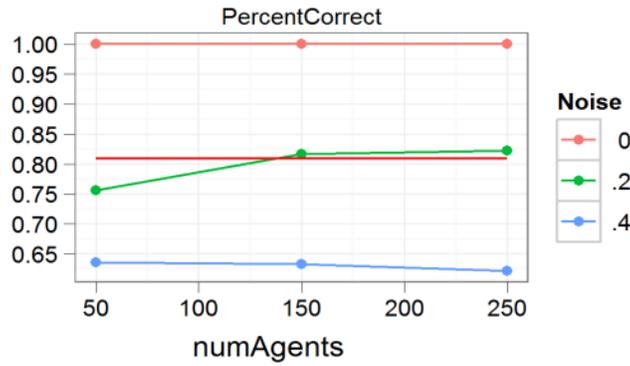


Figure 7 Network size (numAgents) versus percent of all agents holding knowledge who held the correct knowledge, split by amount of noise

While differences in the amount of total knowledge spread are obvious when considering network size, we were somewhat surprised to see that as deviations of trust in the network increased, the number of agents receiving any knowledge decreased. Expectedly, as Figure 8 shows, this resulted in the same correlation noted above, where percentage of agents holding correct knowledge is tied to the total number of agents receiving knowledge.

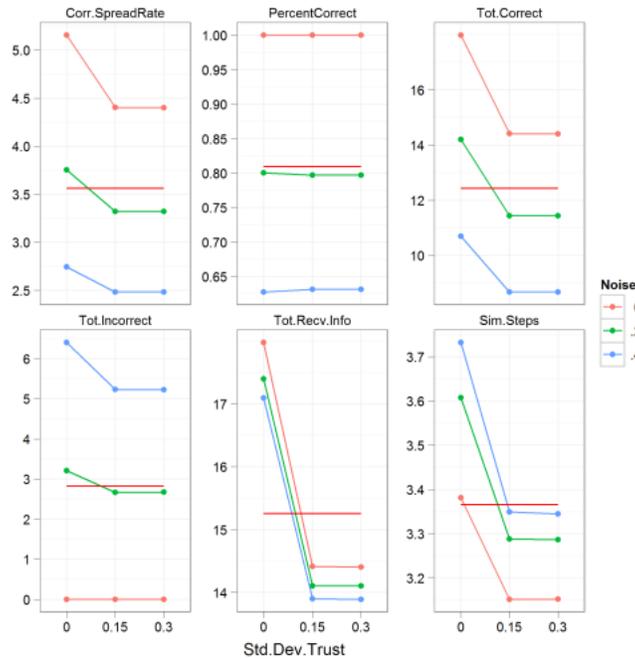


Figure 8 Standard Deviation of Trust, split by amount of Noise, across all outcome variables except the number of conflicts

As described earlier, trust was used by [37] with respect to agents within a group trusting each other more than members of other groups, not as a random variant across all links. While not directly applicable, we can qualify their finding with the fact that larger trust differentials may only increase the spread of information when trust is not randomly assigned. Indeed, our work represents a kind of worst-case scenario, where agents trust each other based on some characteristic wholly unrelated to the information being spread.

While the idea that agents within a group are more likely to trust each other is valid, in networks where communities have not yet emerged, or where the communication network matches forced organizational interactions instead of true social relationships, trust may indeed be closer to random than uniform across groups. For example, in military organizations, members of a platoon may tend to trust friends which are in other platoons more than those they are forced to operate with. Such quasi-random trust levels within a social network defined by organizational communication structures could easily cause incorrect knowledge to be valued more than correct knowledge, and hence in the worst case, this could significantly hurt information gathering performance. However, on average, as Figure 8 shows, this does not appear to be the case. This can of course be attributed to the large number of randomized trials we run – it would be interesting to see whether or not conditions could be contrived where such an effect does occur.

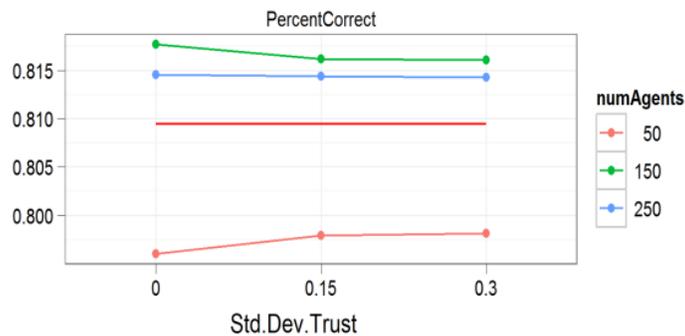


Figure 9 Standard Deviation of Trust, split by the network size (numAgents), against percent of all agents holding knowledge who held the correct knowledge, split by amount of noise

In noting that we see a similar interplay between rate of correct knowledge spread and breadth of total knowledge spread across both levels of trust deviations and network size, we consider the interaction between these two variables. As shown in Figure 9, in smaller networks, less information spread is better (with regards to percent correct), while in larger networks, a larger amount of information spread (no trust differentials) is desirable. This finding may point towards the dangers of, under conditions of noise in information gathering, having a large network where there is not enough communication to realize conflicts, or a small network where there is too much discussion, leading to the slow spreading of knowledge.

4.3. Topology

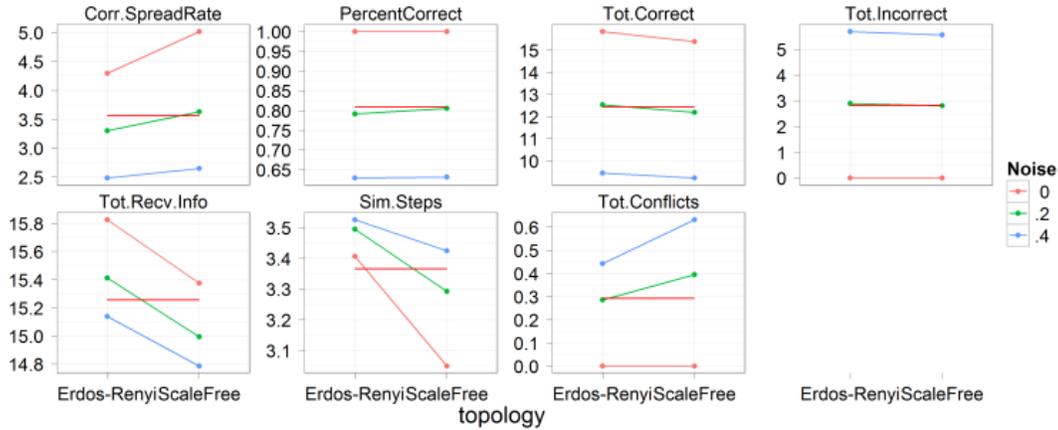


Figure 10 Topology of the network against all outcome metrics, split by level of noise

Figure 10 shows that across all metrics, we see what would be expected for diffusion using our modified Independent Cascade Model when no noise exists in the system. Note that if agents were to pass knowledge on every turn, or to multiple alters, we would surely have seen different results [17] [18]. In particular, we note that because “super-spreader” nodes could only interact with one other agent in our model, the duration over which information persisted in the network and the amount of knowledge spread are lower here than in other models.

Figure 10 also shows that with noise, the direction of change is the same as when no noise exists across all relevant metrics. However, we do see that increases in noise serve to significantly decrease the effect of topology on the number of turns information persisted in the network.

This is the result of conflicts occurring at “super-spreader” nodes, which, with the “recheck” condition, would elongate the amount of time the information persisted. This indicates that if we were to model information as starting at the outer nodes of a scale-free network and moving inwards, it is possible that “super-spreader” nodes would become not just hubs of information, but hubs of conflicting knowledge as well. As these nodes are centers of information spread, the way in which these “super-spreader” nodes deal with conflicting knowledge could have huge implications on the speed and breadth of correct knowledge spread.

4.4. Decision Threshold

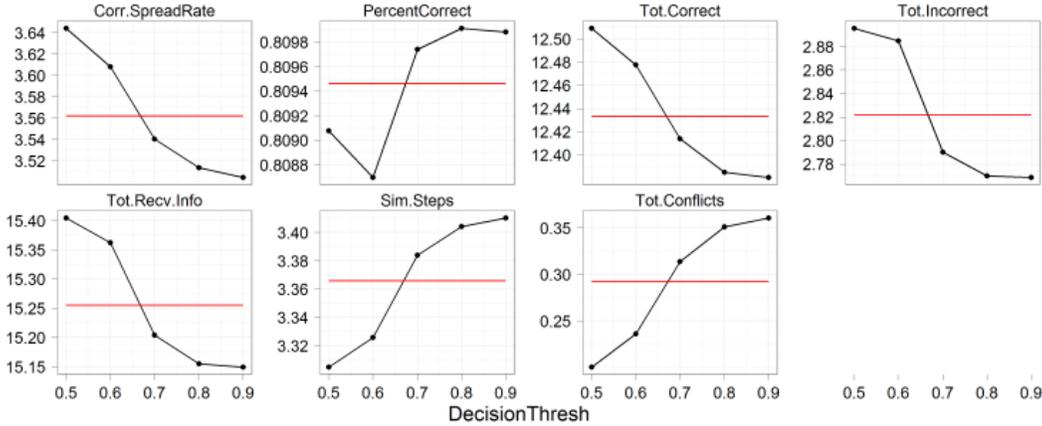


Figure 11 The threshold agents used to make decisions on what knowledge to accept versus all outcome metrics

As shown in Figure 11, the general trend for each of our six outcome variables shows expected outcomes across the decision threshold experimental variable. Specifically, as the decision threshold for an agent increases, the percentage of agents with correct knowledge increases as well. As discussed, this correlates with a decrease in the rate at which this knowledge spreads. In addition, we naturally see more conflicts at higher thresholds, leading to more agents “rechecking” knowledge, which in turn causes information to persist longer. The interesting point that jumps out from the graph above is the obvious hitch between the decision threshold of .6 and .7. The cause of this hitch is another artifact from our simulation model. As described, agents trust their own opinion twice as much as another agents. Consequently, when an agent with a threshold of .7 receives a new opinion from a single other agent, it will, on average, cause a conflict, meaning the agent will change its knowledge to *unknown*. In contrast, an agent will not enter a conflicting state at a threshold of 0.6.

In a sense, increasing the threshold from 0.6 to 0.7 can be thought of as magnifying the “network effect” – agents are more likely to come into conflict, return to a state of *unknown* knowledge and require new opinions from elsewhere in the network. Note, though, that the odds of an agent receiving incorrect information once are significantly higher than an agent receiving incorrect information twice. Consequently, we see here an additional factor, aside from the total number of agents receiving knowledge that affects the percentage of agents who have correct knowledge. Interestingly, however, is that this effect has an obvious ceiling - increasing the decision threshold beyond 0.7 does little to increase the percentage of agents receiving correct knowledge, while continuing to decrease the rate at which correct information spreads.

This finding has implications in future simulation studies concerning the spread of conflicting knowledge. Researchers simulating how agents deal with conflicting knowledge, if using a thresholding approach (instead of, e.g., a fuzzy approach), should be aware that the performance of their agents at any given threshold will be largely affected by how much information an agent receives, on average, over the course of a simulation turn (in the case where agents have no historical knowledge). The threshold selected will also be largely affected by how the agent treats his existing knowledge.

4.5. Set of Alters from which agents accept information

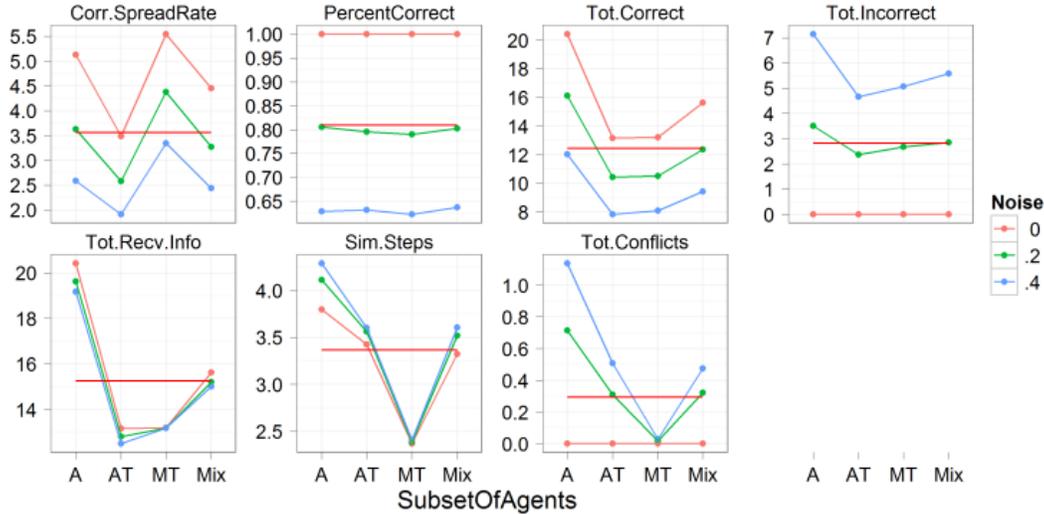


Figure 12 The set of alters from which agents would accept opinions condition, split by level of noise

As Figure 12 shows, across all noise levels, we see the same general patterns for each of our outcome variables for each agent type. In all cases, we observe that the number of agents receiving correct knowledge is greatest when an agent listens to opinions from all of his sources. This is due to the fact that, much like in the real world, agents are oblivious to who will accept their opinion when they send it. Because an agent has only one chance to send their opinion, it is of course the case that knowledge will spread more slowly when opinions passed are often ignored.

Given this, however, we were somewhat surprised that when agents received information from only a single alter, almost no more agents received correct knowledge than when the agents obtained knowledge from a random subset of their alters (the “all trusted” condition). This finding can be explained, in part, by noting that in smaller networks, average degree tended to be close to two, meaning these two conditions were essentially equivalent. However, we also note that agents using the “most trusted” ICRS had almost no conflicts on average, consequently allowing all knowledge to spread almost uninterrupted throughout the system. This corresponds well to real-life, where the fewer people someone accepts information from, the less likely they are to know two sides to a story.

Figure 12 also suggests that at low noise levels, networks have the highest percentage of agents with correct knowledge when all agents listen to all alters, whereas with high levels of noise networks performed slightly better when agents listened to only a subset of their alters. This general trend indicates the possibility that the less noise that exists in a system, the more open agents can be to opinions from anyone, while at higher levels of noise, agents may want to balance the number of agents they accept opinions from in order to balance their ability to realize knowledge is incorrect while keeping conflicts to a minimum.

4.6. Decision Enhancements

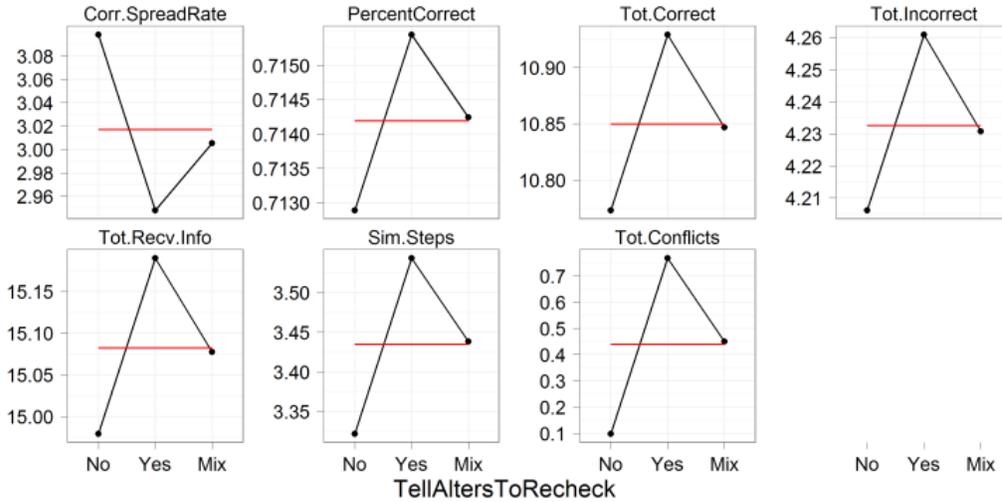


Figure 13 The “recheck” condition (TellAltersToRecheck) versus all outcome metrics

Figure 13 plots the results from implementing the “rechecking” enhancement, where agents who do not pass their decision threshold ask alters from whom they received an opinion most recently to “recheck” their information. As Figure 13 shows, our “rechecking” enhancement tends to increase the number of simulation steps an information piece persisted for and the percentage of agents obtaining correct knowledge. Furthermore, because more information flow incited more conflicts, the rate at which information spread decreased. Such a finding again supports the conclusions of [1], where conflicting knowledge leads to more debate, which in turn slows the spread of new knowledge.

Figure 13 also shows that across all metrics, networks which have a mix of agents who “recheck” and others that do not lie somewhere in the middle of the other two conditions. If one is to consider the “recheck” mechanism as a way to represent a “cautious” agent in an organization, our findings hold reasonably close to real life in that networks in which everyone is cautious would probably have more correct knowledge but be slower-moving, while those which move with abandon pass information more rapidly but with lower rates of correctness, and organizations employing a diverse set of people fall somewhere in the middle.

Of course, this “recheck” mechanism is only one way to represent a “cautious” agent. Indeed, even within our experiment, we have other models that can be thought of as representing a cautious person. These other methods all centered on the notion of uncertainty. While interesting in concept, the effects of each of our usages of uncertainty was smaller by at least one order of magnitude than any other effect discussed up to this point. As such, we can only say that differences across conditions for when agents 1) did or did not “pass along” their uncertainty, 2) changed their knowledge before or after checking their uncertainty thresholds, and 3) told or did not tell others to “recheck” their knowledge based on their uncertainty values all had no obvious effects across any of our outcome metrics.

Further understanding of these results would be interesting in several ways. With regards to simulation studies, it is interesting to note that, although these uncertainty-based ICRS enhancements consider various levels of cautiousness at the agent-level, they do not have the same effect as the “recheck” enhancement. This elucidates the dangers of differences between semantics of an agent’s properties and operationalization. In particular, we note that the “recheck” enhancement affects both the way agents retain and pass information, while all of our

uncertainty-based enhancements affect only one or the other. This correlates well with findings discussed in works discussed in Section II that suggest the way an agent decides what information to keep and to send both can affect how influential they are in a network.

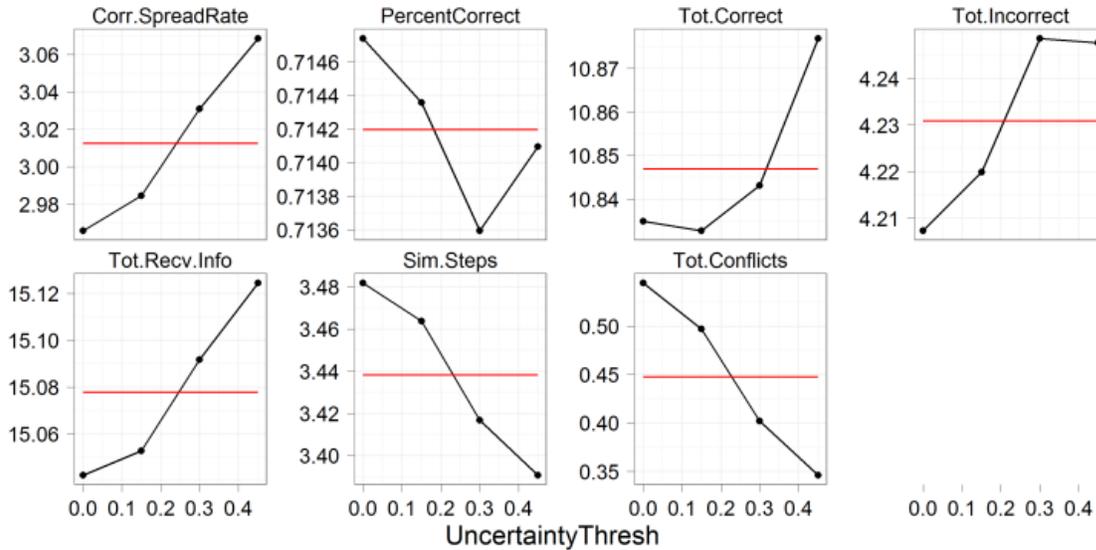


Figure 14 The threshold agents used to make a decision on whether or not to use one of the uncertainty-based enhancements to the Basic ICRS versus all outcome metrics

We must also consider the fact that, when summing the effects of all uncertainty enhancements based on thresholds, we do find results on a more relevant scale, and that have similar curiosities seen across decision thresholds. Figure 14 shows that an uncertainty threshold of 0.45 (corresponding to a decision threshold of 0.55) is the only threshold that is near or above average in the number of agents receiving correct information, the percent correct and the spread rate. In particular, there is a significant increase in the percent correct from an uncertainty threshold of 0.3 to an uncertainty threshold of 0.45, similar to those that we saw between decision thresholds of 0.6 and 0.7

It is of reasonable interest, but some skepticism, that we note the effect over the point at which an agent acts, and not his specific action, differ by an order of magnitude. While further research is necessary, modeling the concept of when an agent acts is much easier than modeling any possible action that an agent might take. If it is indeed the case that the point at which an agent acts is as relevant as the action an agent takes, simulation models at the agent level may be able to become much more expressive while still maintaining a low level of complexity.

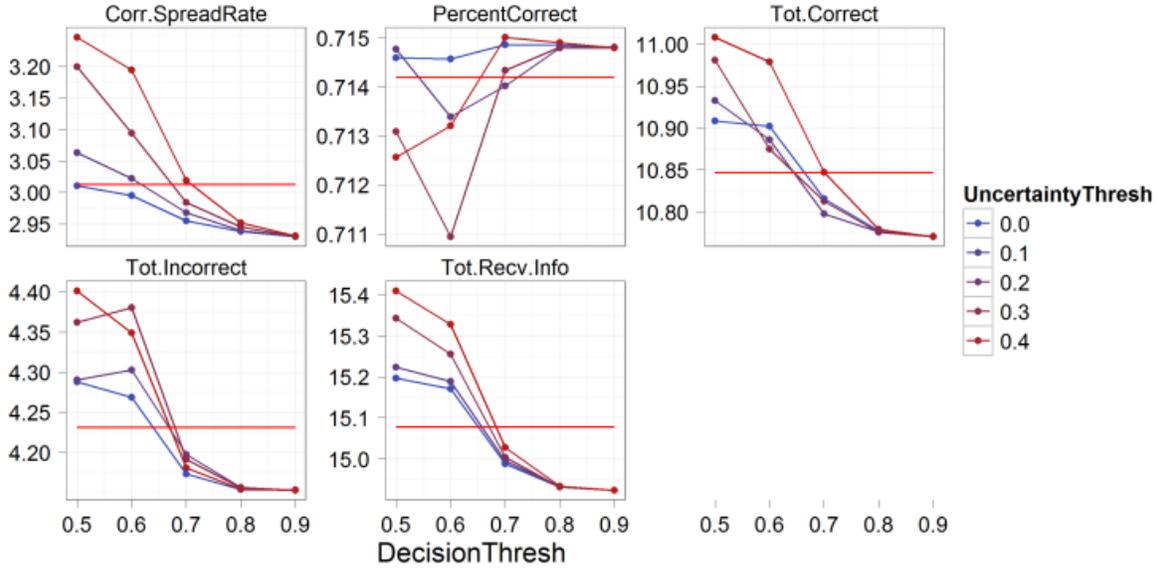


Figure 15 Decision threshold versus select outcome metrics, split by uncertainty threshold

As an example of the type of analysis that might result from considering only the point at which agents commit to an action, and not the action itself, Figure 15 plots our six metrics for each combination of uncertainty threshold and decision threshold. We see that the combination of a decision threshold of 0.7 and an uncertainty threshold of 0.45 consistently performs above average with regards to spread rate of correct information, percentage of agents having knowledge who have correct knowledge, and the total number of agents with correct knowledge. In addition, there exists an added benefit of having a lower than average number of incorrect agents in the system.

5. CONCLUSION

Overall, we find that our model fits well with several real world notions of how conflicting knowledge might spread through a network. Such findings must now be compared with empirical data collected from real world situations. Indeed, in several places we find that our model might under-represent certain factors of how conflicting knowledge spreads, but might over-represent others. The cause of these under and over-representations are artifacts from both our selection of outcome variables and our simulation model, and can and should be manipulated to fit with findings from real-world data.

Although we believe the conclusions we have made are valid and interesting, we first consider limitations before expounding upon findings. Most obviously, as we work with simulation, our results must be digested with the knowledge that we are using artificial data generated from a simplified view of information diffusion in the real world. We simplify knowledge to being either wholly correct or entirely false, which is not always the case. In addition, we note that the magnitude of some of the results we interpreted is small. While common statistical notions tell us that, because of the volume of tests that we run, magnitude is not as important as direction of change, whether or not the change is large enough to affect real world networks may be questionable. Our simplification also entails the incorrect notion that agents have no historical evidence from which to draw their decisions, which may be a misrepresentation of actual decision-making strategies [2]. We simplify the notion of trust to a random, static, bidirectional value, a form that is known to have little real-world use. Finally, we

simplify real-world social networks to two types of highly stylized structures which are unlikely to exist in their purest form in large-scale social networks, and treat social links as static, an idea which has been highly disputed in recent years [17].

Keeping in mind the given limitations, we can condense much of our findings into three points which have both practical and theoretical implications. First, we note that there is support for our hypothesis that conflicting knowledge in a system spreads differently than “unconflicted” knowledge. Chiefly, we note that the speed and breadth of knowledge spread in systems where knowledge is not conflicted is significantly higher than in systems where conflicting knowledge exists. We see that this difference may be exasperated at higher densities – on the other hand, of course, this difference may be negligible at lower densities. Furthermore, we note that “super-spreaders” in Scale-Free networks play an even more important part in conflicted knowledge spread than they do with “unconflicted” knowledge, as they must decide not only to whom they will pass knowledge but also what to send. We note that more modern views of diffusion, such as those which include the concept of broadcasting information simultaneously to multiple alters, and more complex models that allow agents to retain more intricate knowledge bases, may both find this effect to be even more pronounced. In addition, we find that, in our simulations, conflicting knowledge tends to persist longer in the network. Given these findings, we believe that future studies on information, knowledge and/or fact diffusion should be cautious in any assumptions made about the spread of information in an environment where it has the potential to be conflicted.

Second, under the simulation assumptions we used, and the decision strategies implemented, it is evident that with respect to the breadth versus rate of correct knowledge spread in the face of conflicting knowledge, one may not be able to “have one’s cake and eat it too”. Obvious trade-offs existed between the speed at which correct knowledge spreads and the percentage of agents that received the correct knowledge. While unsurprising, this empirical finding is obviously at odds with traditional diffusion studies, which suggest that the faster information spreads, the better. In addition, it suggests that organizations which obtain information from noisy sources may require more elaborate decision-making schemes or network structures, or at least be cognizant of the fact that this trade-off exists. These tradeoffs have real-world applications- as Kittur [1] showed, in organizations such as Wikipedia, where truth is paramount, new information will consequently tend to spread at a slower rate. Conversely, when the goal is to spread information as quickly as possible, organizations may want their agents to pass on whatever information they receive without considering whether or not it conflicts with their previous knowledge. This may be the case in situations where agents can “recheck” information without requesting it from their alters (e.g. using a web search) at a later time. This may also be true in situations where there exists a need to rapidly notify people that a certain situation is pending, even if details are hazy or incorrect. Such may be the case with disaster relief efforts, where the important fact may be that “a tsunami is coming,” and not so much that “a tsunami is coming in 46 minutes”.

Finally, we observe that more complex ICRS tended to add little to the diversity of our outcomes. This can be seen, for example, in our results for uncertainty usage, where we found that, although the implementation and run complexity of the actions taken by agents was high, their effect was exceedingly slight, if existent at all. Indeed, we found that simple changes to agent’s ICRS produced relevant changes, and that future simulation studies on this topic might consider other simplistic modifications over ones that are unnecessarily complex.

6. References

- [1] A. Kittur, B. Suh, B., B.A. Pendleton, "He Says, She Says: Conflict and Coordination in Wikipedia," in CHI 2007: Proceedings of the ACM Conference on Human-factors in Computing Systems, 2007.
- [2] K.M. Carley and Z. Lin, "A Theoretical Study of Organizational Performance under Information Distortion," *Management Science*, vol. 47:7, pp. 976-997, July 1997.
- [3] S. Benferhat, D. Dubois, and H. Prade, "How to infer from inconsistent beliefs without revising?," in Proc. of the 14th International Joint Conference on Artificial Intelligence, IJCAI'95, pp. 1449–1455, 1995.
- [4] X. Yin , J. Han and P.S. Yu, "Truth discovery with multiple conflicting information providers on the web," in Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining, August 2007
- [5] J. Cantwell, "Resolving Conflicting Information," *Journal of Logic, Language and Information*, vol. 7:2, pp. 191-220, April 1998.
- [6] A. Hunter and L. Weiru, "A Context-Dependent Algorithm for Merging Uncertain Information in Possibility Theory," *Systems, Man and Cybernetics, Part A: Systems and Humans*, IEEE Transactions on , vol.38:6, pp.1385-1397, Nov. 2008
- [7] T.A. Cameron, "Updating Subjective Risks in the Presence of Conflicting Information: An Application to Climate Change," *Journal of Risk and Uncertainty*, vol. 30:1, pp. 69-97, Jan. 2005
- [8] K. Carley, "On the Persistence of Beliefs," Working Paper, Department of Social and Decision Sciences, Carnegie Mellon University, 1990.
- [9] D. Kempe, J. Kleinberg and E. Tardos. "Maximizing the spread of influence through a social network," in KDD '03: Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 137-146, 2003.
- [10] P. Domingos, and M. Richardson "Mining the Network Value of Customers," in Proceedings of the Seventh International Conference on Knowledge Discovery and Data Mining pp. 57-66, 2001.
- [11] M. Granovetter, "Threshold Models Of Collective Behavior." *The American Journal of Sociology*, vol. 81:6 pp. 1420-1443, May 1978
- [12] 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*, vol. 12:3, pp. 209-221, 2001.
- [13] S.P. Borgatti, "The Key Player Problem," in *Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers*, pp.241-252, 2003.
- [14] J. Diesner, T. Frantz and K.M. Carley, "Communication Networks from the Enron Email Corpus 'It's Always About the People. Enron is no Different'," *Computational & Mathematical Organization Theory*, vol. 11:3, pp. 201-228, Oct. 2005
- [15] V. Buskens and K. Yamaguchi, "A New Model for Information Diffusion in Heterogeneous Social Networks," *Sociological Methodology* vol. 29:1, pp. 281-325, 1999.
- [16] B. Tadic and S. Thurner, "Information super-diffusion on structured networks," *Physica A* 332, pp. 566, 1999
- [17] K. Joseph and K.M. Carley, "Issues Associated with Sampling and Aggregating Interaction Data," Submitted. Accessible at: <http://www.casos.cs.cmu.edu/publications/papers.php>

- [18] A.-L. Barabási, R. Albert, and H. Jeong, "Scale-free characteristics of random networks: the topology of the world wide web," *Physica A*, 281: pp. 69–77, 1999.
- [19] M. Broecheler, P. Shakarian and V.S. Subrahmanian , "A Scalable Framework for Modeling Competitive Diffusion in Social Networks," in *Social Computing (SocialCom)*, 2010 IEEE Second International Conference on , vol., no., pp.295-302, 2010
- [20] C. Schreiber, S. Singh, and K.M. Carley, "Construct - A multi-agent network model for the coevolution of agents and socio-cultural environments". Carnegie Mellon University, School of Computer Science, Institute for Software Research, International. Technical Report, CMU-ISRI-04-109, 2004.
- [21] D. Centola, "The Spread of Behavior in an Online Social Network Experiment, " *Science* vol. 3, September 2010.
- [22] D. Downey, E. Etzioni, E., and S. Soderl, "A probabilistic model of redundancy in information extraction." In *Proceedings of the 19th Annual International Joint Conferences on Artificial Intelligence*, pp. 1034-1041, 2005
- [23] D. Artz and Y. Gil, "A survey of trust in computer science and the Semantic Web". *Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 5:2, pp. 58-71, June 2007.
- [24] K. Kelton, K. Fleischmann and W. Wallace, "Trust in digital information, " *J. Am. Soc. Inf. Sci. Technol.* vol. 59:3 pp. 363-374, February 2008.
- [25] P.R. Laughlin, S.W. VanderStoep, and A.B. Hollingshead, "Collective Versus Individual Induction, Recognition of Truth, Rejection of Error, and Collective Information Processing," *Journal of Personality and Social Psychology* vol. 61:1, pp.50-67, 1991
- [26] P.G. Ipeirotis, F. Provost, V.S. Shen, and J. Wang, "Repeated Labeling Using Multiple Noisy Labelers," *NYU Working Paper No. CEDER-10-03, September 2010.*
- [27] K.M. Carley , "Coordinating the success: trading information redundancy for task simplicity," *System Sciences, Proceedings of the Twenty-Third Annual Hawaii International Conference on , vol.4, no., pp.261-270, Jan 1990*
- [28] A.C. Amason, "Distinguishing the effects of functional and dysfunctional conflict on strategic decision making: Resolving a paradox for top management teams," *Academy of Management Journal*, vol. 39, pp. 123-148, 1996.
- [29] K.A. Jehn, G.B. Northcraft, and M.A. Neale, "Why Differences Make a Difference: A Field Study of Diversity, Conflict, and Performance in Workgroups," *Administrative Science Quarterly* vol. 44:4, pp. 741-763, December 1999
- [30] K. M. Carley, M. K. Martin and B.Hirshman, "The Etiology of Social Change," *Topics in Cognitive Science*, vol. 1.4, pp. 621-650, 2009.
- [31] J.H. Gross, "Cues and heuristics on Capitol Hill: Relational decision-making in the United States Senate". Doctoral Thesis / Dissertation, 2010, 213 Pages
- [32] Mostaghimi, M.; , "Bayesian estimation of a decision using information theory," *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on , vol.27, no.4, pp.506-517, Jul 1997*
- [33] E.M. Rogers, E. M. Diffusion of Innovations. New York: Free, Print, 2003.
- [34] J. Howison, A. Wiggins, and K. Crowston, Validity issues in the use of social network analysis with digital trace data. Working paper. <http://flosshub.org/content/validity-issues-use-social-network-analysis-digital-trace-data>

- [35] P.F. Lazarsfeld and R.K. Merton, "Friendship as Social Process: A Substantive and Methodological Analysis." In Monroe Berger et al., eds., *Freedom and Control in Modern Society*, 1964.
- [36] R. Guha, R. Kumar, P. Raghavan, and A. Tomkins. "Propagation of trust and distrust," in *Proceedings of the 13th international conference on World Wide Web (WWW '04)*: 2004.
- [37] C. Hui, M. Goldberg, M. Magdon-Ismail, and W. A. Wallace, "Agent-based simulation of the diffusion of warnings". In *Proceedings of the 2010 Spring Simulation Multiconference (SpringSim '10)*. Article 9 , 8 pages, 2010.
- [38] K. Perusich and M.D. McNeese, "Using Fuzzy Cognitive Maps for Knowledge Management in a Conflict Environment," *Systems, Man, and Cybernetics, Part C: Applications and Reviews*, IEEE Transactions on , vol.36, no.6, pp.810-821, Nov. 2006
- [39] R.R. Yager and S. Kikuchi, "On the role of anxiety in decisions under possibilistic uncertainty," *Systems, Man, and Cybernetics, Part B: Cybernetics*, IEEE Transactions on , vol.34, no.2, pp.1224-1234, April 2004
- [40] P. Guo, D.Z. Zeng and H. Shishido, "Group decision with inconsistent knowledge," *Systems, Man and Cybernetics, Part A: Systems and Humans*, IEEE Transactions on , vol.32, no.6, pp. 670- 679, Nov 2002

