An Explanation for the Efficiency of Scale Invariant Dynamics of Information Fusion in Large Teams

Robin Glinton, Paul Scerri, Katia Sycara
Robotics Institute
Carnegie Mellon University
rglinton,pscerri,katia@cs.cmu.edu

Abstract – *Large heterogeneous teams will often be in situations where sensor data that is uncertain and conflicting is shared across a peer-to-peer network. Not every team member will have direct access to sensors and team members will be influenced mostly by teammates with whom they communicate directly. Simple models of a large team sharing beliefs to reach conclusions about the world show that the dynamics of such belief sharing systems are characterized by* information cascades*. These are ripples of belief changes through the system caused by a single additional sensor reading. Glinton et al. [1] showed that such a system will exhibit qualitatively different dynamics sensitive to ranges over system parameters. In addition they showed that In one particular range, the system exhibits behavior known as* scale-invariant dynamics *which was found empirically to correspond to dramatically more accurate conclusions being reached by team members. In this paper we provide an analytical explanation for the performance of scale invariant dynamics by leveraging signal processing concepts. We show that scale invariant dynamics behave as an adaptive information filter with a response that automatically adjusts to the accuracy of sensor inputs. This adaptation causes the performance gain.*

Keywords: Self-organization, Complex systems, Large Scale Information Fusion

1 Introduction

[1] Large heterogeneous teams will often be in situations where sensor data that is uncertain and conflicting is shared across a peer-to-peer network. Not every team member will have direct access to sensors and team members will be influenced mostly by teammates with whom they communicate directly. The effective sharing and use of uncertain information is key to the success of large heterogeneous teams in complex environments because without a correct understanding of the environment it is not possible to appropriately plan and act. Typically, noisy information is collected by some portion of the team and shared via the social and/or physical networks connecting members of the team [2]. Each team member will use incoming, uncertain information and the beliefs of those around them to develop their own beliefs about relevant facts. However, the volume of incoming data relative to bandwidth constraints, will often make it impossible for agents to communicate all the received information. Each agent must filter and abstract the information, communicating only its conclusions. Example applications of such systems include large scale disaster relief, environmental monitoring, military crisis response etc. [3].

Before such teams are deployed in domains where there are significant costs for bad behavior, it is important to understand and, if necessary, mitigate any system-wide phenomena that occur during belief propagation. Understanding the dynamics of the system and linking this understanding to overall system performance is difficult since network-based belief propagation in large heterogeneous teams exhibits complex emergent behaviors [4]. Previous attempts to describe the information dynamics of complex systems includes describing propagation of fads [5, 4], rumors [6] and gossip[7] through social networks. The key difference between this work and previous work is that in previous work a single type of information spread whereas here we can have conflicting data that fundamentally changes the dynamics. Moreover, we are able to use agents to predict and control system dynamics in order to guide the team to areas of optimized performance.

To analyze the dynamics, we follow Glinton et al. [1] and model a team as being connected via a network with some team members having direct access to sensors and others relying solely on neighbors in the network to inform their beliefs. Each agent uses inference over communications from direct neighbors and sensor data to maintain belief about the environment. This model is attractive because the level of abstraction of the model allows for investigation of team level phenomena decoupled from the noise of high fidelity models or the real-world, allowing for repeatability and systematic varying of parameters. In this model the dynamics of information exchange are dominated by large *cascades*

[1]This research has been sponsored in part by AFOSR FA95500810356.

of changes of belief that can occur from a single new sensor reading after many previous sensor readings led to no significant change in beliefs of team members.

Glinton et al. found empirically that when there was an exponential frequency distribution of cascade sizes, a situation known as *scale invariant dynamics*, there were dramatically fewer instances of propagation of incorrect beliefs. In fact, teams exhibiting belief propagation with scale invariant dynamics can be as much as 50 times more reliable at reaching correct conclusions than teams whose dynamics did not exhibit scale invariance. Moreover, they observed that in the range where the system exhibits scale invariant dynamics, system convergence is much faster than in the other system ranges. In this paper we provide an analytical explanation for the phenomenon that the performance of teams that exhibited scale invariance was much higher than those that did not.

Our analysis is predicated on a signal processing interpretation of the agent system as an information filter which transforms system level inputs to system level outputs. In signal processing theory the output or response of a filter given an input signal is obtained by first determining what is known as the impulse response of the system. This is the response of the filter to a sudden brief change in the input. The output is then determined by linear combinations of the impulse responses to discrete samples of the input signal. This mathematical operation is called a convolution. We conducted our analysis by first mapping the properties of the agent system to the analagous properties of an information filter. First we used percolation theory to formalize the notion of a system response for a large distributed system. Next we developed a probability distribution over the response of the system as a function of sensor input and the distribution over the cascade sizes that result. This probability distribution is analagous to the convolution operation with the cascade distribution acting as an impulse response of the system.

Our analysis showed that the high performance of scale invariant dynamics for large-scale information fusion is due to the fact that scale invariant dynamics act as an impulse response with an infinite *memory*. The memory of an impulse response is a measure of the number of past inputs that influence the current response of the system in the convolution of the input signal and the impulse response. This property of the scale invariant distribution allows the response of the system to adapt to the accuracy of sensor input, responding slowly when accuracy is low allowing the system to collect more information, and quickly when accuracy is high. The adaptive response also explains the efficiency of message passing with the system passing only as many messages as the sensor accuracy dictates. We confirmed the results of our analysis empirically showing that the response of the system does adapt to the sensor accuracy when the team utilizes scale invariant dynamics.

2 Model

In this section, we formally describe the underlying model used in the remainder of the paper. The model is intended to be the simplest model that can capture the complex dynamics of uncertain information being shared by a cooperative team. A cooperative team of agents, $A = \{a_1, \ldots, a_{|A|}\}$ are connected by a network, $G = (A, E)$ where E is the set of links in G which connect the agents in A. An agent a_i may only communicate directly with another agent $a_j \in N_{a_i}$ if $\exists e_{i,j} \in E$ where we refer to the set N_{a_i} as its *neighbors*. The average number of neighbors that the agents in G have is defined as $< d >$ where $< d >= \frac{\sum_i |N_{a_i}|}{|A|}$.

Sensors, $S = \{s_1, \ldots, s_{|S|}\}$ provide noisy observations to the team. Only one agent can directly see the output of each sensor. The sensors return binary observations about some fact b from the set $\{$*true*,*false*$\}$. In this paper the correct value of the fact is always *true*. We refer to the probability that a sensor s will return a correct observation as its reliability r_s. The reliability of a sensor is known to the agent that receives observations from it. In the remainder of this paper, unless otherwise specified, $|A| = 1000$, $|S| = |A|/20$ and $r_s = 0.55 \forall s$. That is, most agents must deal with relatively noisy data and do not have direct access to the sensors. For example in military intelligence where only a few intelligence analysts might have direct access to data from sensors like unmanned aerial vehicles.

A key assumption of the model is that it is infeasible for agents to communicate actual sensor observations to one another and that they may only communicate whether they currently believe the fact to be *true*, *false* or if they are undecided, *unknown*. Although restricting agents to communicating only their conclusions is purely an abstraction to make working with and understanding the model easier, we believe that there are many real world domains where it is infeasible to communicate actual sensor readings. For example, sensor data might be video or audio recordings than are expensive to share on a large network and require significant effort and skill to interpret, or sensor data might be secret, or physical specimens that cannot be shared. If there are large numbers of sensor readings, restricted communication channels and many facts that a large number of agents need to come to conclusions about, we expect it to be infeasible to send most types of raw sensor data.

Each agent a_i uses either an observation received from a sensor or conclusions about b communicated by neighbors to form a belief $P_{a_i}(b \rightarrow true)$ about b. A new observation is incorporated into the current belief to form a new belief $P'_{a_i}(b \rightarrow true)$ using Equation 1 an expression of Bayes' Rule.

$$P'(b \rightarrow true) = \frac{P(b \rightarrow true) * cp}{P(b \rightarrow false) * (1 - cp) + P(b \rightarrow true) * cp} \quad (1)$$

Where $cp = P(b \rightarrow true/o \rightarrow true)$ and o is an observation. Each observation from a sensor is treated as an independent observation. Only the last communication from any neighbor is treated as an observation. Observations from different neighbors are treated as independent in the application of Bayes' Rule. The treatment of observations of neighbors as independent is not correct, since they may have come to their conclusions based on the same data. Hence, agents relying on neighbors to reach a conclusion will inevitably be over-confident in their conclusions. We refer to this effect as *double counting*. Without communicating actual sensor data or having detailed knowledge of the entire network structure and message sequence, it is impossible to completely remove double counting.

An agent a_i will communicate *true* if $P_{a_i}(b \rightarrow true) > \sigma$ and *false* if $P_{a_i}(b \rightarrow true) < 1-\sigma$. Unless otherwise specified $\sigma = 0.8$. If the communication causes a neighbor's belief to cross a threshold, it too will communicate with all its neighbors. We refer to this as an *cascade*. In the simulation of the agent team, when an agent receives a sensor reading, we allow the resulting cascade to stop propagating before introducing any subsequent sensor readings to the system. The probability $P(k)$ that k agents change their belief during an cascade is a key measure of the dynamics of the system used throughout the remainder of this paper.

The most important metric for each agent and for the team overall is reliability, r_a for a single agent and R for the whole team defined as r_a = total correct conclusions/total incorrect conclusions and $R = \sum_i r_i/|A|$. If agents connected to sensors did not communicate until a very large number of sensor readings arrived, they could be very confident their conclusions was correct and the team would be very reliable. However, the team would also be very slow to make decisions and would not leverage the presence of multiple sensors. We use a second metric, convergence time, C_n, as the time it takes for n agents to reach the same conclusion. Below, we use $n = 0.8|A|$.

Finally we define $n_o(cp)$ as the number of sequential observations, having the same truth value, that would be required to change the conclusion of an agent starting with a belief in opposition to those observations. By this definition an undecided agent would need to receive $n_o(cp)/2$ consistent readings to communicate.

3 Model Dynamics

In this section we discuss the dynamics of the model presented in Section 2. In [1] Glinton et al. developed equations to relate the system parameters in this model to $P(k)$, the probability that a cascade will encompass k agents as a result of a single sensor observation. The analysis was predicated on the assumption that the network G has a random topology with $|A| \rightarrow \infty$.

Glinton et al. showed that the system exhibited three distinct distributions $P(k)$ depending on the values of the system parameters cp and $< d >$. Furthermore, they showed that the performance of the system as measured by the metric R was extremely sensitive to these parameters and consequently to which of the cascade distributions governed system dynamics.

They showed that when parameters cp and d are chosen such that $\alpha =< d > /(n_o(cp) + 2) = 1$ the cascade distribution for the system is given by Equation 2.

$$P(k) \propto k^{-3/2} \tag{2}$$

Furthermore, when $\alpha =< d > (1/(n_o(cp) + 2)) < 1$, $P(k)$ is distributed according to Equation 3.

$$P(k) \propto k^{-3/2} exp\left(-\frac{k}{(1- < d > /(n_o(cp)+2))^{-2}}\right) \tag{3}$$

Finally, when $\alpha =< d > (1/(n_o(cp) + 2)) > 1$, $P(k)$ is distributed according to Equation 4.

$$P(k) \propto k^{-3/2} exp\left(\frac{k}{(1- < d > /(n_o(cp)+2))^{-2}}\right) \tag{4}$$

Equation 4 is identical to 3, except the sign on the exponential term is positive. Consequently, in this parameter range large cascades are enhanced and are much more probable than in the other parameter ranges.

Glinton et al. showed empirically that the performance of the team measured by the metric R was up to 1000 times better than in other parameter ranges, when the parameter cp was chosen relative to $< d >$ such that $\alpha = 1$ and the cascade distribution was governed by the scale-invariant distribution given by Equation 2. Furthermore, the convergence time C_{800}, the time for 80% of the team to come to the correct conclusion, was simultaneously very low. This is shown in Figure 1. The figure also shows that the team's performance is clearly optimized for $cp = 0.63$ which corresponds with scale invariant dynamics. We also see that the performance of the team, measured via both reliability and convergence time, is extremely sensitive to cp. The figure shows how R peaks dramatically while C_{800} is very small. As cp is increased, reliability goes up at first slowly and then very dramatically before falling off even more dramatically. The convergence time drops dramatically at the high end of the cp spike. Notice that for high cp, C_{800} is high because there is not agreement by 80% of the team even after a long period.

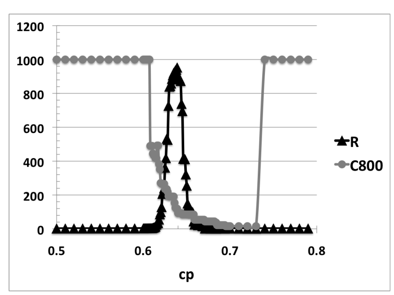

Figure 1: **R** vs **cp**, the peak which is several orders of magnitude larger than the surrounding points occurs at $\alpha = 1$.

4 Analysis of Model Dynamics

In this section we provide an explanation for the optimal performance of scale invariant dynamics as compared to the other two dynamics modes of the system discussed in Section 3. To aid in the understanding of the mechanisms which make scale-invariant dynamics optimal, it is useful to view the agent system from a signal processing perspective. Specifically we consider the network of agents as a central filter that transforms inputs to outputs. Figure 2 illustrates this view. In the figure $s[t]$ gives the input to the system as a function of time, and $S[t]$ gives the output. In signal processing, the action of a linear filter on its inputs is characterized by $h[t]$ the impulse response of the filter. The impulse response of a filter is the response of the filter to a sudden, brief, change in the input. For a linear filter, the output is the sum over the products of the impulse response of the filter with sequential discrete samples of the input over time. The intuition is that if we break up the input signal into individual impulses, for a linear filter, the total response is given by the sum over the responses to individual inputs. This is expressed mathematically by the operation of convolution between the input signal and the impulse response: $S[t] = s[t] * h[t] = \sum_{m=-\infty}^{m=\infty} s[t-m]h[m]$. The result is the output of the system. In this paper we study a nonlinear system but the analogy is still useful to elucidate the important concepts.

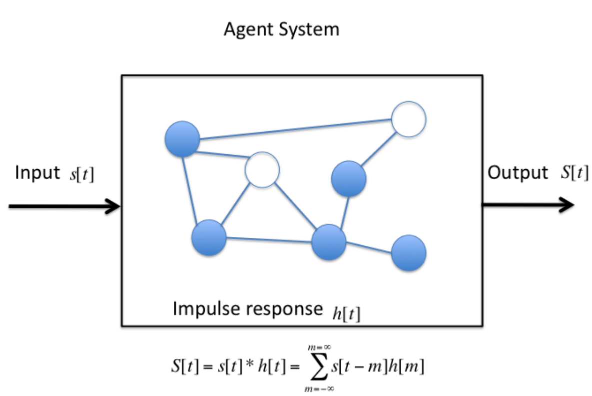

Figure 2: The agent system viewed as a filter that transforms inputs to outputs at a macroscopic level.

From this viewpoint, the input to the system is the pooled inputs from individual sensors. The internal dynamics of agent interactions within the system, as characterized by the distributions of Section 3 act as a filter which transforms inputs to outputs with the distributions of Section 3 acting as the impulse response of the system. The system response is given by the macroscopic response of the system, that is, the response of the system at it's largest scale on the order of the number of agents in the system $O(|A|)$. The output can be thought of as the macroscopic state of the system as expressed by averages of metrics of interest over all agents in the system.

For an information filter, the convolution operation shown in Figure 2 acts to integrate sensor inputs to produce a conclusion as output. The output of an information filter is extremely sensitive to the *memory* of the impulse response $h[t]$ of the filter. The memory of the impulse response is the degree to which past sensor inputs contribute to the current response of the filter. Intuitively, for an information filter, the best impulse response is the one where the memory adapts to the accuracy of the sensor inputs. For example, if the sensor data is extremely inaccurate then a long memory is necessary, with the current response as a function of many previous sensor inputs. However, if the data is very accurate and we desire an answer as quickly as possible, then a long memory is undesirable and we want a response that is only a function of recent inputs. The high performance of the agent system for scale invariant dynamics, is because the scale invariant cascade distribution is analagous to an impulse response with a memory that adapts to the accuracy of the input. Conversely, the exponential distributions perform poorly as an impulse response because they both have a fixed memory length parameterized by cp and $< d >$, the trust agents put in their neighbors information and the average number of neighbors that agents have. In the investigation of Glinton et al. the sensor accuracy was very low with $r_s = 50\%$. Because of this a fixed length memory exhibited particularly poor performance.

Consider the memory of two types of impulse response. An exponential impulse response $h[t] = e^{-bt}$ and a scale invariant impulse response $h[t] = t^{-3/2}$. For both of the impulse responses the output $S[t]$ is obtained by convolution of the impulse response with the input $s[t]$ this is expressed by Equation 5.

S[t]=s[t]*h[t]=$\sum_{m=-\infty}^{m=\infty} s[t-m]h[m]$ (5)

Figure 3 shows a plot of $log(h[m])$ vs m for both types of input response. Note that the exponential impulse response has a window over a finite range of m outside of which it becomes arbitrarily close to zero. Notice from Equation 5 that this implies that inputs at time $t-m$ for m outside of this window are effectively ignored in the output. This window gives the memory of the impulse response and hence the total output response for sensor inputs. If accuracy in sensor input is low many sensor inputs will need to be integrated to obtain high accuracy in the output. However if the window of the impulse response is small then the output will always be a function of a small number of inputs and output accuracy will be low. Contrast this with the scale invariant response which decays extremely slowly. In the convolution of Equation 5 this means that the output at any time t will be a function of all inputs. Furthermore, the slow decay means that the impulse response is small and Equation 5 is dominated by the signal $s[t-m]$, this produces the adaptation of the response to the signal which happens with a scale invariant impulse response. Also note that the impulse response is always small until a sufficient number of sensor readings has accumulated.

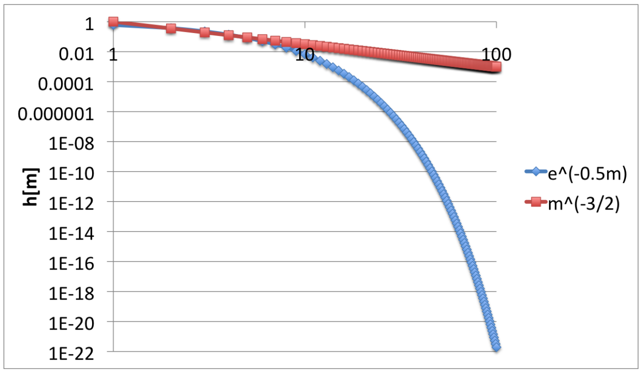

Figure 3: An exponential impulse response has a fixed length memory while a scale invariant response has an infinite memory.

In the remainder of this Section we map the parameters and dynamics of the agent system to the analagous concepts presented thus far. In so doing we show that the efficiency of the scale invariant dynamics are because they act like the scale invariant impulse response shown in Figure 3. That is they provide a response to input sensor readings that is dominated by the accuracy of the signal and is a function of all previous sensor readings.

The remainder of this Section is organized as follows. First we formalize the notion of the agent system response. Next, we develop a probability distribution over the response of the agent system as a function of sensor inputs and the cascade distribution, this plays the role of the convolution operation (Equation 5) of the filter with the cascade distribution acting as impulse response. To close out the section we investigate the memory of this distribution by calculating the probability of a response at time t_n conditioned on previous inputs for $t < t_n$. We show that for any exponential distribution over cascade sizes the probability of a response at time t_n is independent of previous inputs. We further show that for a scale invariant distribution over cascade sizes, the probability of a response at time t_n is both dependent on the accuracy of the sensor (an adaptive memory) and is a function of all previous sensor readings (an infinite memory).

4.1 Definition of System Response

In this Section we formalize the notion of a system level response for the agent system. The system under study responds with cascades of belief changes, we formalize a system level response of the agents as a cascade where the size of the cascade k is on the order of the number of agents in the system. That is when $O(k) = O(|A|)$. Given this definition of a system response, we need to relate the probability of a system response to the probability of the response of a single agent. This is necessary because some cascades will not be large enough to constitute a system level response, however, they will cause individual agents in the system to respond (communicate with neighbors). For this reason we must develop response equations in terms of the response of individual agents. However, giving a link between the two we can then relate the probability of the response of individual agents to the probability of the system level response that we actually care about. The link between the probability

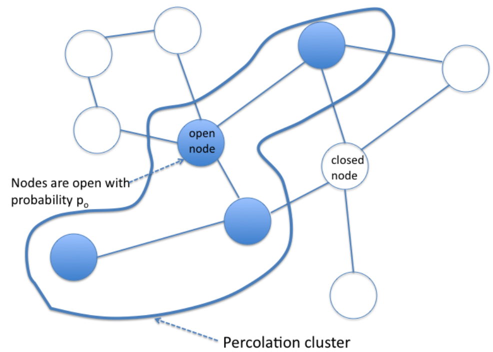

Figure 4: Percolation theory relates the probability p_o of a node being open to a flow to the probability that the flow will encompass the network.

of a response by an individual agent and a system response is provided by Percolation Theory [**?**]. The variables of interest in percolation theory are shown in Figure 4. Given a static probability p_o that a randomly selected node in the network is open to a flow, Percolation theory provides the tools for calculating the probability that the size of a connected cluster of open nodes is on the order of the size of the graph. In the case of a lattice, this is often interpreted as the probability that a flow starting on one edge of the lattice will propagate to the other side. The chief result of percolation theory is that for any finite network topology, there is a corresponding critical value of p_o, called p_c such that when $p_o = p_c$, the largest connected cluster of open nodes spans the network. Note that p_c is unique to a given network topology and is independent of the specifics of the fusion mechanism employed by agents. For a homogeneous random graph with link density $< d >$, $p_c = 1/(< d > -1)$ [**?**]. For the model under study we know that an agent in the unknown state will communicate after receiving $n_0(cp)/2$ readings. Consequently for the system under study an agent is *open* to the passage of a cascade if it has received $(n_0(cp)/2) - 1$ readings. Therefore, for this model we can define p_o as $p_o = P(\text{An average agent has received } no_(cp) - 1 \text{ readings})$. We are now free to develop equations that relate the distributions of Section 3, the impulse response, to p_o the probability that an individual agent will respond. We then automatically know the probability of a system response which is given by $P(p_o > p_c)$ which can be obtained by simple integration of the distribution p_o. We conducted an experiment to test the prediction of percolation theory that a system response will occur deterministically when $p_o > p_c = 1/(< d > -1)$. In the experiment we averaged 5000 simulation runs of the system and calculated p_o at each simulation time step.

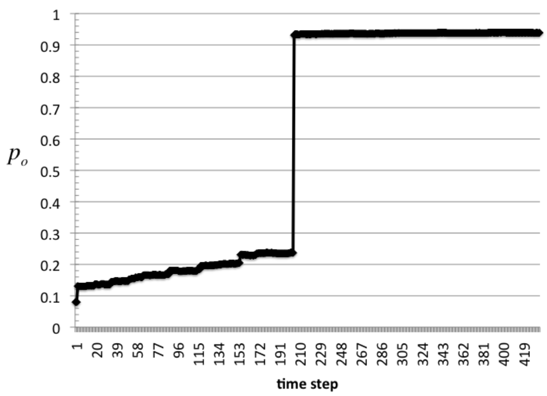

Figure 5: One instance of a macroscopic system response with cascade size $O(|A|) = O(1000)$ when $p_0 > (p_c = 0.33)$

The graphs shown here were produced using $|A| = 1000$, $r_s = 0.55$, $< d >= 4$, $cp = 0.74$, but the results are typical. Figure 5 is a plot of p_o vs time step for a single run. Note the sharp transition around $p_o = 0.24$ to high values of p_o meaning that a system level cascade has occurred. For network density $< d >= 4$, $p_o = 0.33$. The discrepancy is because in our simulation cascades happen in their entirety in a single simulation time step. Once a cascade has begun, the cascade itself increases p_o meaning that a cascade that will eventually propagate system wide might start below p_c. Figure 6 shows the same plot averaged over all 5000 runs.

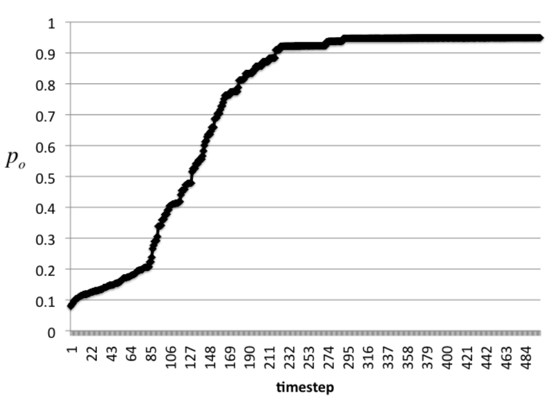

Figure 6: Average over 5000 instances of a macroscopic system response with cascade size $O(|A|) = O(1000)$ when $p_0 > p_c$)

Not surprisingly the transition is not as sharp, as the actual value of p_o before the start of the cascade that eventually causes $p_o > p_c$ varies.

4.2 Probability of a System Response

In this section we develop an equation for the probability of a system level response, $P(p_o^t > p_c)$, that plays the role of the convolution given in Equation 5. We calculate this probability as a function of the probability over possible sensor inputs to the system and the cascade distribution $P(k)$ which is analagous to the impulse response $h[t]$ in Equation 5. We calculate this distribution for use in later sections to investigate the *memory* of this distribution given particular choices of $P(k)$ as discussed in Section 4.

To calculate $P(p_o^t > p_c)$ we start by defining the random variables which describe the operation of the system. We define $\mathbf{C^t} = [\sigma, \psi]$ as the vector valued random variable that returns the number of times that a randomly selected agent was touched by *true* cascades (σ) and the number of times that agent was touched by false cascades (ψ) up to time step t. Next we define the vector $\hat{k_t} = [k_0, k_1, \ldots, k_t]$ where $k_i \in \hat{k_t}$ gives the size of the *false* avalanche that occurred at time $t = i$ (Note that if a true avalanche happened at $t = i$ then $k_i = 0$). Similarly we define $\hat{g_t} = [g_0, g_1, \ldots, g_t]$ where $g_i \in \hat{g_t}$ gives the size of the *true* avalanche that occurred at time $t = i$. We further define $\mathbf{K^t}$ as the random variable that returns the sum of the sizes of false avalanches that occurred up to time t. That is, $\mathbf{K^t}$ takes on values from $\sum_{i=0}^{t} k_i$. We define $\mathbf{G^t}$ as a random variable that plays the same role for true avalanches. (Note for the model presented in this paper, only a single cascade per time step is possible). Finally we define the vector valued random variable $\mathbf{S^t}$ which takes on values from the possible vectors of sensor inputs $\hat{s_t} = [s_0, s_1, \ldots, s_t]$ input to the system up until time t.

With the appropriate variables defined, the joint distribution over these variables is given by

$$P(\mathbf{C^t}, \mathbf{K^t}, \mathbf{G^t}, \mathbf{S^t})$$

. Given the joint we can calculate the probability of any system event by integrating over the appropriate variables. Using successive applications of the chain rule the joint distribution reduces to

$$P(\mathbf{C^t} \mid \mathbf{K^t}, \mathbf{G^t})\mathbf{P}(\mathbf{K^t} \mid \mathbf{S^t})\mathbf{P}(\mathbf{G^t} \mid \mathbf{S^t})\mathbf{P}(\mathbf{S^t})$$

where we have used the fact that $\mathbf{C^t}$ is independent of $\mathbf{S^t}$ given $\mathbf{K^t}$ and $\mathbf{G^t}$. That is if we know how many cascades of *true* and *false* have occurred then we automatically know the type of sensor reading, *true* or *false* that caused them.

Next we develop the individual distributions within the factored joint distribution. The individual sensor readings are binomially distributed therefore $P(\mathbf{S^t} = \hat{\mathbf{s_t}}) = \beta(|\hat{\mathbf{s_t}}^{\mathbf{T}}|; \mathbf{t}, \mathbf{r_s})$ where $\hat{s_t}^T \subset \hat{s_t}$ is the set of *true* readings in $\hat{s_t}$. For notational convenience we simplify this to $\beta(\hat{s_t})$. $P(\mathbf{C^t} \mid \mathbf{K^t}, \mathbf{G^t})$ is the product of two binomial distributions with $P(\mathbf{C^t} \mid \mathbf{K^t}, \mathbf{G^t}) = \beta(|\mathbf{k_i} \in \hat{\mathbf{k_t}} : \mathbf{k_i} \neq \mathbf{0}|; \mathbf{t}, \mathbf{1}/|\mathbf{A}|)\beta(|\mathbf{g_i} \in \hat{\mathbf{g_t}} : \mathbf{g_i} \neq \mathbf{0}|; \mathbf{t}, \mathbf{1}/|\mathbf{A}|)$ where, again for notational simplicity we simplify this to $\beta(\hat{k_t})\beta(\hat{g_t})$.

We can calculate $P(p_o^t > p_c) = P(\mathbf{C^t} : \sigma - \psi > \mathbf{n_o}(\mathbf{cp})/\mathbf{2})$ by appropriately integrating the joint distribution.

$$P(\mathbf{C^t} : \sigma - \psi > \mathbf{n_o}(\mathbf{cp})/\mathbf{2})$$

$$= \sum_{\hat{k_t}} \sum_{\hat{s_t}} \beta(\hat{k_t})\beta(\hat{g_t})P(\hat{k_t} \mid \hat{s_t})P(\hat{g_t} \mid \hat{s_t})\beta(\hat{s_t}) \qquad (6)$$

4.3 Memory of the System Response

In this section we investigate the memory of the system response distribution given by Equation 6. Recall that this equation expresses $P(p_o^t > p_c)$, the probability that the system will respond with a system wide cascade. We formally define the memory of the distribution as:

$$P(p_o^t > p_c \mid \text{A specific sequence of cascades up until t-1})$$

That is, the probability that the system will respond at time t conditioned on the history of the dynamics and sensor readings. This will act as a measure of how much the response at the current time step, depends on the past sequence of sensor readings. This is equivalent to the *memory* of the system impulse response discussed in Section 4. Once we have calculated this distribution, we will use it to show that the system's ability to adapt to sensor readings is highest when $P(k)$ the cascade distribution is scale invariant. We can calculate:

$$P(p_o^t > p_c \mid \text{A specific sequence of cascades up until t-1})$$

$$= \frac{P(p_o^t > p_c \wedge \text{A specific sequence of cascades up until t-1})}{P(\text{A specific sequence of cascades up until t-1})}$$

Using Equation 6 this reduces to:

$$\frac{\sum_{k_t}\sum_{s_t}\beta(\hat{k_t})\beta(\hat{g}_t)P(\hat{k_t} \mid \hat{s}_t)P(\hat{g}_t \mid \hat{s}_t)\beta(\hat{s}_t)}{\beta(\hat{k}_{t-1})\beta(\hat{g}_{t-1})P(\hat{k}_{t-1} \mid \hat{s_{t-1}})P(\hat{g}_{t-1} \mid \hat{s}_{t-1})\beta(\hat{s}_{t-1})} \quad (7)$$

We now consider the effect of changing the distributions of $P(\hat{k_t} \mid \hat{s}_t)$ and $P(\hat{g}_t \mid \hat{s}_t)$ on the memory of the response as defined by Equation 7. If $P(k)$, the distribution over cascade sizes is exponentially distributed as is the case when Equations 3 or 4 govern system dynamics. Then $P(\hat{k_t} \mid \hat{s}_t)$ and $P(\hat{g}_t \mid \hat{s}_t)$ are both distributions over sums of exponentially distributed random variables and are Erlang distributed with:

$$P(\hat{k_t} \mid \hat{s}_t) = \frac{\lambda^t (\sum_i^t k_i)^{t-1} e^{-\lambda \sum_i^t k_i}}{(t-1)!}$$

. (8)

Where λ is the exponent from Equation 4 or 3. Since Equation 8 contains an exponential, then the terms

$$\frac{P(\hat{k_t} \mid \hat{s}_t)}{P(\hat{k}_{t-1} \mid \hat{s}_{t-1})} \propto \frac{e^{-\lambda \sum_i^t k_i}}{e^{-\lambda \sum_i^{t-1} k_i}} = e^{-\lambda(\sum_i^t k_i - \sum_i^{t-1} k_i)} = e^{-\lambda k_t}$$

will dominate Equation 7. Notice that the exponentials in these terms are only dependent on k_t the size of the cascade that happened at time t (and consequently the sensor reading at time t). This means that when the cascade distribution $P(k)$ is exponentially distributed, the system *forgets* previous sensor inputs and the response of the system is governed solely by the current sensor input. This causes the poor performance when the sensor reliability r_s is low. In this case many sensor readings are needed to obtain high accuracy.

If $P(k)$, the distribution over cascade sizes is scale invariant as is the case when Equation 2 governs system dynamics. Then $P(\hat{k_t} \mid \hat{s}_t)$ and $P(\hat{g}_t \mid \hat{s}_t)$ are both distributions over sums of power law distributed random variables and are power law distributed with the same exponent as $P(k)$. Therefore:

$$\frac{P(\hat{k_t} \mid \hat{s}_t)}{P(\hat{k}_{t-1} \mid \hat{s}_{t-1})} = \left(\frac{\sum_i^t k_i}{\sum_i^{t-1} k_i}\right)^{-\alpha}$$

. (9)

Where α is the exponent from Equation 2. Notice that the terms dependent on the cascade distributions when the cascade distribution is scale invariant, as expressed by Equation 9 depend on the sums of avalanche sizes over all time. This means that the scale invariant distributions give the response distribution an infinite memory. Further notice that the exponent $-\alpha$ causes the change in the response due to the avalanche history to vary slowly. The consequence of this is that the Binomial terms in, Equation 7 which gives the probability of a system level response, have a strong effect on the response. These binomial terms are functions of the history of sensor input. The consequence of this is that the response is strongly dependent on the history of the sensor input. This means that the scale invariant cascade distribution causes the system response to adapt to the accuracy of the sensor input. This explains the excellent performance of the system shown in Figure 1 when the cascade distribution is scale invariant. The system response adapts to the sensor input. When the accuracy of the sensor r_s is low, the system will take a longer time to respond thus raising the accuracy of the information spread by the large cascade when it eventually occurs. Conversely when accuracy r_s is high, the system will respond quickly.

To test the prediction of the analysis, that the scale invariant response causes adaptation to the sensor accuracy, we conducted an experiment where we observed the variation in the time of a system level response as a function of r_s the sensor accuracy. We defined a system level response as one in which the cascade size $k > |A|/2$The result of this experiment is shown in Figure 7. The x-axis of all plots gives the sensor accuracy r_s and the y-axis gives the simulation time step when a system level response occurrs. The graphs were produced using $|A| = 1000$, $r_s = 0.55$, $< d >= 4$. The progression of the graphs from left to right show the result for increasing values of cp. The second plot from the left shows the plot with a value of $cp = 0.63$ when the distribution of $P(k)$ is scale invariant and given by Equation 2. Notice that this is the plot for which there is maximum correlation between the time step when the system response occurrs and the sensor accuracy.

5 Related Work

There have been several studies conducted to investigate models whose dynamics are governed by cascades on complex networks. These include models of fads[5, 4], rumors

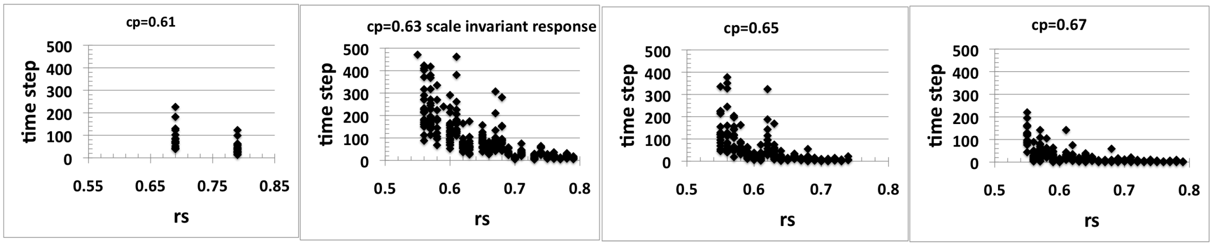

Figure 7: The time step for a cascade of size $O(|A|)$ is maximally correlated with the reliability of the sensors r_s for the scale invariant response.

[6], gossip[7], forest fires [8], and diseases[9, 10]. Common to all of these models is that the dynamics are governed by the spreading of a single influence. In contrast, our model investigates competing influences which significantly alters the dynamics of a system. Many of the systems under study in these studies fall into the category of self-organized critical systems, which exhibit scale invariant dynamics over large parameter ranges[11]. In contrast, we found that very small changes in the conditional probability assigned to information from neighbors in our model could cause the team to exhibit other types of dynamics.

Recently there has been significant interest in social networks [12], [13] and the impact of those networks on performance of a group. For example, Xu looked at the impact of networks on routing information to a specific agent [14]. Kleinberg, looked at the impact of the network on the performance of decentralized search algorithms [15], when a single agent has information valuable to the system. We build on both of these contributions by investigating the case when a large percentage of the agents in the team are both sources and sinks for information, which fundamentally changes the dynamics of information exchange in the system. Boyd has looked at the impact of networks on decentralized gossip-based information dissemination [7], our analysis method could be utilized to understand the dynamics of information exchange in Boyd's model.

6 Conclusions and Future Work

In this paper we provided an explanation for the high performance of scale invariant dynamics for information propagation and fusion in a large scale team. Leveraging concepts from signal processing we showed that scale invariant dynamics cause the agent system to act as an information filter that adapts the response of the system to the accuracy of the sensor input. In the future we plan to investigate using the results of this analysis to design large scale information fusion systems which may utilize different fusion mechanisms. In addition we plan to investigate if the performance of more sophisticated models which do not exhibit scale invariant dynamics can be improved by artificially encouraging scale invariant dynamics.

References

[1] R. Glinton, P. Scerri, and K. Sycara, "Exploiting scale invariant dynamics for efficient information propagation in large teams," in *Proc. of AAMAS'10*, 2010.

[2] B. Bellur, M. Lewis, and F. Templin, "An ad-hoc network for teams of autonomous vehicles," in *Proceedings of the First Annual Symposium on Autonomous Intelligence Networks and Systems*, 2002.

[3] J. Casper and R. Murphy, "Human-robot interaction during the robot-assisted urban search and rescue effort at the world trade center," vol. 33, pp. 367–385, 2003.

[4] W. DJ, "A simple model of global cascades on random networks," *Proceedings of the National Academy of Science*, vol. 99, pp. 5766–5771, 2002.

[5] S. Bikhchandani, D. Hirshleifer, and I. Welch, "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of Political Economy*, vol. 100, no. 5, p. 992, 1992. [Online]. Available: http://www.journals.uchicago.edu/doi/abs/10.1086/261849

[6] M. Nekovee, Y. Moreno, G. Bianconi, and M. Marsili, "Theory of rumour spreading in complex social networks," *Physica A: Statistical Mechanics and its Applications*, vol. 374, no. 1, pp. 457–470, 2007.

[7] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE/ACM Trans. Netw.*, vol. 14, no. SI, pp. 2508–2530, 2006.

[8] S. Clar, B. Drossel, and F. Schwabl, "Scaling laws and simulation results for the self-organized critical forest-fire model," *Phys Rev E*, vol. 50, p. 10091018, 1994.

[9] R. Pastor-Satorras and A. Vespignani, "Epidemic spreading in scale-free networks," *Phys. Rev. Lett.*, vol. 86, no. 14, pp. 3200–3203, Apr 2001.

[10] V. M. Eguíluz and K. Klemm, "Epidemic threshold in structured scale-free networks," *Phys. Rev. Lett.*, vol. 89, no. 10, p. 108701, Aug 2002.

[11] P. Bak, C. Tang, and K. Wiesenfeld, "Self-organized criticality: an explanation of 1/f noise," *Physical Review Letters*, pp. 381–384, 1983.

[12] D. Watts and S. Strogatz, "Collective dynamics of small world networks," *Nature*, vol. 393, pp. 440–442, 1998.

[13] A.-L. Barabasi and E. Bonabeau, "Scale free networks," *Scientific American*, pp. 60–69, May 2003.

[14] Y. Xu, P. Scerri, B. Yu, S. Okamoto, M. Lewis, and K. Sycara, "An integrated token-based algorithm for scalable coordination," in *AAMAS'05*, 2005.

[15] J. Kleinberg, "Complex networks and decentralized search algorithms," in *Proceedings of the International Congress of Mathematicians (ICM)*, 2006.