Minimizing Information-Centric Convergence Cost in Multi-Agent Agreement Problems

Kiran Lakkaraju
Department of Computer Science
University of Illinois at Urbana-Champaign
klakkara@illinois.edu

Les Gasser
Graduate School of Library/Information Science
& Department of Computer Science
University of Illinois at Urbana-Champaign
gasser@illinois.edu

ABSTRACT

The basic paradigm of learning has shifted significantly, from single agents that learn in single, static environments, to collective learning: multiple, interacting agents with diverse goals learning from each other across different local environments. Instances of collective learning abound in sensor networks, peer-to-peer systems, distributed recommender systems, and in social systems in general. A critical, but unexplored, activity in collective learning is information gathering: the exchange of information about other agents' individual preferences, that will guide collective decision making processes. A set of tradeoffs exists between the amount of information agents gather, the effort of this information gathering, and agents individual and collective performance. Reducing the amount of information gathered may reduce information costs, but reduced information can produce interpretation errors that create suboptimal behavior in the agents and the collective.

In this paper we define and study the impact of these tradeoffs using the well known "Generalized Simple Majority" decision process in the model problem of *norm emergence*, a type of multi-agent agreement process in which agents converge to a common strategy. We present a new metric, "Information-Centric Convergence Cost" (ICCC), that combines information cost with the cost of time, and a new decision process, "Generalized Simple Sampled Majority," and we study these in several agent network topologies. Surprisingly, we find that as the level of information gathering is reduced the amount of error increases non-linearly, giving a non-linear impact on performance. Thus the careful manipulation of information-processing effort can minimize ICCC while still achieving quick norm emergence.

1. INTRODUCTION

The predominant learning paradigm in AI is single agents learning from a single, static environment. However, in many cases learning is a collective process of multiple, interacting agents (Multi-Agent Systems or MAS) with a variety of goals learning from each other across different local environments, and acting together in a shared world. In systems of loosely coupled, autonomous sensing acting machines (e.g., sensor networks, UAV coalitions, peer-to-peer networks) there are numerous scenarios (e.g. collaborative information processing, distributed sensing [2]) in which agents must learn from each other.

The acquisition and evolution of language in humans is one particularly intriguing instance of collective learning. Children initially learn a language from a small set of other people, often their close family. Then they extend their language competence by learning from a variety of possibly contradicting sources such as their peers. Across many instances in a population, variances in the exchange and use of language can lead to evolution in the underlying language itself. Thus "learning" occurs for a single agent through interaction with many others, and it also occurs for the collective as a whole as the shared language itself develops in expressivity, efficiency, variety, etc.

In this paper we investigate collective learning using the case of *norm emergence*. Norms – collective behavioral conventions that have the effect of constraining, structuring, and making predictable the behaviors of individual agents — are an important organizing principle in MAS. Examples of norms are behaviors such as driving on the correct side of the road, tipping, and many elements of language. For instance, using the term "sick" to mean "good" or "cool" is a type of linguistic norm in a certain youth culture – a collective constraint on linguistic behavior for that set of speakers. While norms may be imposed on a population through the design of mechanisms such as "social laws," the more interesting case is where in many instances norms emerge through repeated interaction and modification of behaviors between agents. Finally, norm emergence is a *model problem* through which we can study general principles of learning in collectives.

We study norm emergence through the lens of the Distributed Optimal Agreement Framework (DOA) [5], which identifies the core processes underlying multi-agent agreement in general, and hence in norm emergence. Two main processes of DOA are especially critical and are studied here: (1) *information gathering* – in which agents exchange information about their current individual preferences and develop collective views; and (2) *information use* – in which each agent uses this collective information to modify its own preferences. We study the interactions of speed and cost in norm emergence by focusing on these two processes.

Current studies of norm emergence limit performance met-

ALA Workshop at 2009 AAMAS Conference

rics to the *time until convergence on a norm*; e.g., [4, 3, 12] all focus on the time it takes for a norm to emerge. Placing time-to-convergence into economic terms, the *cost* of convergence is the sum of the costs each agent pays for being collectively unconverged (e.g., lost-opportunity costs). Beyond the cost of unconverged time, the key activities of information gathering and use may also incur costs. But these information cost aspects of convergence haven't yet been studied carefully. This is a gap we need to rectify, since information costs can be critical and dominant design issues. For instance, in energy-constrained wireless sensor networks the energy cost of communication is orders of magnitude greater than the energy cost of local computation [10]; principled approaches to balancing information processing costs and benefits is critical.

To address this issue, we propose a new measure of the cost of norm emergence, called the *Information-Centric Convergence Cost* (ICCC), that is based on both the time, T, and the "effort" (the expenditure of resources for information exchange/gathering and processing by an agent/system) expended for a norm to emerge. We study ICCCs that are a linear combination of these two factors, as follows:

$$\mathrm{C}\left(\mathrm{T},\mathrm{E}\right) = c_t T + c_e E. \tag{1}$$

c_t is the cost incurred for every time step in which no norm has emerged. We can view this as the system "paying" c_t dollars for every time step there is no norm. c_e is the cost of every unit of "effort" an agent must expend in order for a norm to emerge.

Clearly there should be a tradeoff between expenditures of information processing effort, time to convergence, and the ICCC. Intuitively, the more effort agents expend, the better their individual estimates of collective preferences, the better their decision making, and thus the lower the convergence time. However, if information is expensive, the overall ICCC increase (due to information cost) might overcome the value of the respective convergence time decrease.

On the other hand, and again intuitively, restricting the information gathering/processing effort may create suboptimal local decisions, and lead to longer convergence times.

In this paper we study this critical tradeoff through careful empirical simulation and analysis of a new model of norm emergence, the *Generalized Simple Sampled Majority* (GSSM) process. GSSM is an extension of the *Generalized Simple Majority* process in [3] that employs *sampling* to vary the level of information gathering effort. Surprisingly, we find that under significant reduction of information processing effort the system can still converge to a norm quickly; we analyze the impact of limited information and show that the tradeoff between effort, time and Information-Centric Convergence Cost can be effectively managed. A key insight is the surprising robustness of the GSSM rule to a lack of information under certain conditions.

2. RELATED WORK

Norms and their impact have been under a lot of study under numerous domains recently. [9] provides an overview of the salient characteristics of norms.

There are numerous models of norm emergence, however none to our knowledge fully considers the role of information cost. Several models exist for the emergence of norms in situations where there are no constraints on interaction [15, 12]. [4] and [3] consider the time-mediating effect of various complete and incomplete social networks in norm emergence. Graph incompleteness does constrain information propagation, so degree distribution of a network has information cost implications (we demonstrate some effects below). However, none of these studies considers information costs in analyses of norm convergence.

Recent work has focused on norm emergence in a variety of situations – with heterogeneous populations of learners [11]; and with populations that are physically constrained [8, 11, 14].

Models of the emergence of norms can model the creation of *linguistic norms* [6].

Communication costs are sometimes considered when evaluating distributed algorithms, using the concept *message complexity* [7]. In energy-constrained wireless sensor networks the cost of communication (for sharing views) dominates the cost of local computation; thus minimizing communication is an important goal [16, 10]

3. ICCC IN THE GENERALIZED SIMPLE MAJORITY PROCESS

[3] developed the *Generalized Simple Majority (GSM)* process as a generalization of [15]'s *simple majority rule* for norm emergence[1]. In GSM each agent in a population holds one of two possible states or opinions, denoted by 0 and 1 (i.e., a *binary agreement space*). Agents are placed on a social network and can interact only with their neighbors. At each time step the system evolves by following three steps:

Agent Choice Pick a random agent a_x. Let s denote the state of a_x.

Information Gathering a_x determines the distribution of states among its neighbors. Let k_s be the number of neighbors with state s and $k_{\overline{s}}$ be the number of neighbors with the opposite state, $\overline{s}$, where $k_x = k_s + k_{\overline{s}}$ is the number of neighbors of a_x.

Information Use a_x changes state with probability $f(k_{\overline{s}})$

Where:

$$f(k_{\overline{s}}) = \frac{1}{1 + e^{2\beta\left(1 - 2\frac{k_{\overline{s}}}{k_x}\right)}} \tag{2}$$

Where β is a parameter of the system. Figure 1 shows how the probability changes with $k_{\overline{s}}$.

[3] refers to $f(\cdot)$ as the *Generalized Simple Majority Rule*; we use the term *Generalized Simple Majority Process* (GSM process) to denote the system dynamics as well as the rule.

It is interesting to note that by setting $f(k_{\overline{s}}) = \frac{k_{\overline{s}}}{k_x}$ we recover the dynamics of the well known *Voter model*; i.e., the probability to switch is equal to the fraction of an agents neighbors that have the opposite state [17, 3]. Importantly, even though the resulting convergence dynamics would be the same as the dynamics of the Voter model, the ICCC would be quite different as the active agent would gather information from all of its neighbors.

Let us develop a precise notion of effort in the GSM process. We will assume that the energy required for internal

[1]We describe GSM in terms of the DOA framework, to create a general basis for comparative analysis.

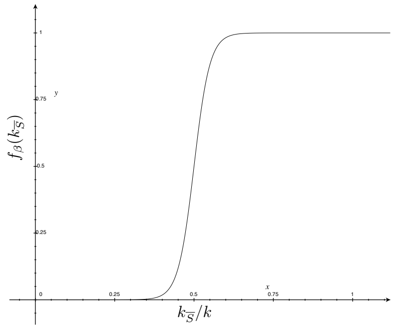

Figure 1: Generalized Simple Majority. The x-axis is the fraction of neighbors with state $\overline{S}$. The y-axis is the probability an agent will switch. $\beta = 10$

computation of information is dominated by the energy required to exchange information. Then the effort undertaken by an agent is proportional to the number of other agents it gathers information from. Let e_x be the effort of agent x at a single time step. Then:

$$e_x = w_x \times k_x.$$

Where w_x is a constant for each agent that indicates the individual energy required per communication interaction, i.e. the "overhead" for every communication interaction. In this work we assume a homogeneous population of agents and set $w_x = 1.0$ for all agents. We also make the simplifying assumption that the energy required to "physically" interact and exchange information is the same for all neighbors.

The effort of the population at time t is:

$$E_t = \sum_x e_x.$$

Then the total effort is:

$$E = \sum_t E_t.$$

What is the Information-Centric Convergence Cost for the GSM process? Consider a population of N agents on a complete social network that converges in T time steps using the GSM process. Delgado conjectured, through extensive simulations, that the time for a norm to emerge in this situation is $T = O(N)$. Based on this conjecture we can estimate the ICCC.

Since at each time step only one agent is active and it samples all other agents, the effort at a time step t is: $E_t = (N-1)$ and thus $E = T * (N-1)$. Substituting $T = O(N)$ into Ewe have $E = O(N^2)$. Substituting into Equation 1 we find that $\mathrm{C}(\mathrm{T}, \mathrm{E}) = O(N^2)$. Even though the time till norm emergence is linear, the ICCC is actually quadratic in the number of agents, which is a tremendous cost.

3.1 Minimizing ICCC

To minimize ICCC we need to minimize the amount of effort which means modifying the amount of information gathered. However, as we reduce the level of information gathering agents will receive increasingly erroneous estimates of the preferences of other agents. This could lead to suboptimal decision making, which in the case of norm emergence might result in a longer time to convergence. The underlying tradeoff is simple, the more information an agent gather, the better its decisions are, however this means more effort, and thus a higher ICCC. The less information gathered, the worse an agents decisions are, which could impact time, and thus could also result in a higher ICCC. Understanding the precise nature of this trade off in the context of norm emergence can shed light onto other collective learning problems.

One widely studied approach is to modify the interaction topology of the population. Instead of a complete graph, a variety of topologies have been studied, including random, scale free and small world graphs. The interaction topology limits the number of neighbors of an agent, thus limiting information gathering. The downside is that information will take longer to propagate to the entire population, and in some cases dramatically increase the time till convergence (e.g. contract nets in [3]).

However, there are some topologies that can converge quickly. [3] shows that a population with a Scale-Free social network can converge in $O(N)$ time.

While modifying the social network provides some benefit, it is often the case that social network topologies are not changeable. For instance, a distributed sensor network deployed randomly in a geographical location has no means of modifying who it can gather information from.

To explore this, we have developed a new model of norm emergence that allows management of the level of information gathering per agent through a "sampling" mechanism.

3.2 Generalized Simple Sampled Majority Process

The *Generalized Simple Sampled Majority* Process (GSSM Process) introduces the concept of *sampling*. Unlike the standard GSM process where an agent queries all its neighbors in the GSSM process agents sample a fraction, θ, of their neighbors. By varying the sampling fraction we can control the level of information gathering, and thus impact the amount of effort to arrive at a norm.

The steps of the GSSM are as follows.

Let θ be the fraction of an agents neighbors that are randomly chosen to be evaluated. Then at each timestep these three processes are carried out:

- Pick a random agent a_x. Let s denote the state of a_x.
- Let Π be a uniformly randomly chosen subset of the neighbors of a_x with size $\max\{1, \lfloor\theta * k_{x]}\rfloor\}$. Let $\pi_{\overline{s}}$ be the number of agents in Π with state $\overline{s}$.
- a_x changes state with probability $f(\frac{\pi_{\overline{s}}}{|\Pi\ |})$

θ provides a way to control the effort of the system by limiting the number of interactions per agent. Following our example from above, for N agents on a complete graph that converge in T time, $E_t = \max\{1, \lfloor\theta*(N-1)\rfloor\}$ and thus $E = T*\max\{1, \lfloor\theta*(N-1)\rfloor\}$. While this does not change the fact that $\mathrm{C}(\mathrm{T}, \mathrm{E}) = O(N^2)$ there is a substantial decrease in effort. By setting θ to low values we can significantly decrease the amount of effort and thus the ICCC. However, as mentioned above, the possibility exists that agents will behave incorrectly due to erroneous estimates. To investigate this, we first determine the probabilities of making an error in Section 4. Surprisingly, we find that even with a low sampling fraction the probability of error is low. We then empirically evaluate the GSSM system to explore this tradeoff between error and effort.

4. ERRORS DUE TO LIMITED INFORMATION

The probability for an agent in state s to switch to $\overline{s}$ is based on the proportion of its neighbors with that state (Figure 1). Let k_x be the number of neighbors of agent a_x. In the following we drop the subscripts for clarity and we assume the agent in question has state s. Let k_s be the number of neighbors of a_x that are in state s, $k_{\overline{s}}$ is the number of neighbors in the opposite state. Let $f_s = \frac{k_s}{k}$ be the fraction of neighbors in state s. Let π_s be the number of agents in Π that are in state s and then $\widehat{f_s} = \frac{\pi_s}{|\Pi|}$. The state that the majority of agents are in is called the *majority state*; the opposite is called the *minority state*.

We can differentiate between two types of errors occurring during the GSSM process:

Mistaken Majority The majority opinion of the agents in Π differs from the actual majority opinion of the neighbors; i.e.:

1. $f_{\overline{s}} \geq 0.5$ but $\widehat{f_{\overline{s}}} \leq 0.5$; or
2. $f_{\overline{s}} \leq 0.5$ but $\widehat{f_{\overline{s}}} \geq 0.5$.

Difference in Strength The majority is preserved, but $\widehat{f_{\overline{s}}}$ differs significantly from $f_{\overline{s}}$

Even if an agent using a reduced sample does correctly detect its neighbors' majority opinion it still may misjudge the *strength* of that majority. Referring to the decision rule in Figure 1, strength misjudgments can significantly alter the probability of state change (by positioning an agent on the correct side of 50% but incorrectly far along the X axis). The shape of the curve in Figure 1 is determined by β; the higher β is the smaller the effect of this positioning error will be. We leave a complete study of the impact of difference-in-strength errors for future work, and focus here on the more salient mistaken majority errors.

We say the sample is a *Success* when a mistaken majority error does not occur, otherwise it is a *Failure*. Clearly mistaken majority errors can cause non-convergence. Suppose a mistaken majority error *always* occurs – then agents switch to the minority state with high probability. This increases the fraction of agents with the minority state, eventually turning that state into the majority one and reversing the process. The result will be oscillation around the 50–50 point.

What is the probability of a sample being successful, $P(\text{Success})$? Without loss of generality suppose state 0 is the majority state. Then $P(\text{Success}) = P(\pi_0 > \pi_1)$ where $m = \pi_0 + \pi_1$ is the size of Π . This can be calculated easily by enumerating the number of different ways of choosing a π_0 size subset of k_0 times the possible ways of choosing an π_1 size subset of k_1. This leads to:

$$P(\text{Success}) = P(\pi_0 > \pi_1) = \frac{\sum_{\pi_0 > \pi_1} \binom{k_0}{\pi_0} \times \binom{k_1}{\pi_1}}{\binom{k}{m}} \qquad (3)$$

Figure 2 shows the probability of success for $k = 999$ for $\theta = [0.1, 1.0]$ and $f_0 = (.5, 1.0]$. As θ increases $P(\text{Success})$ increases for values of f_0 close to 0.5. In norm emergence simulations agents' states are randomly initialized, and stochastically one state will have a slight majority. Figure 2 shows that even under those conditions and with low θ the probability of committing a mistaken majority error is slim. As the fraction of majority state agents increase, this probability reduces significantly.

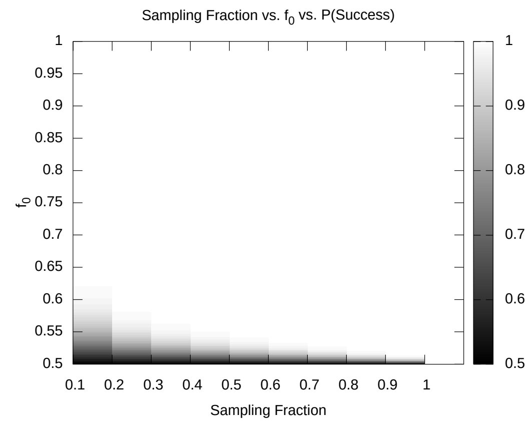

Figure 2: m (shown as the sampling fraction) vs. f_0 vs. $P(\text{Success})$(color) for $k = 999$.

As the number of neighbors decreases $P(\text{Success})$ decreases as well. Figure 3 displays $P(\text{Success})$ for $k = 16$. As can be seen $P(\text{Success})$ does not increase vs. f_0 as it did when $k = 999$.

Figure 3 shows an interesting pattern where every other band has a higher probability. This is due to the fact that we require $\pi_0 > \pi_1$ and $\pi_0 + \pi_1 = m$.The intuition is that when we have an odd number we get a sample "for free" that does not change the minimum value of π_0 as compared to the even number one before. For example, for $m = 6$, $\pi_0 = 4$; however for $m = 7$, it is still true that the minimum value for π_0 is 4. The addition of the extra sample only increased the number of ways to choose a majority sample. Thus, the probability of success for an odd m is greater than the probability of success for $m - 1$.

We can see, however, that the probability of error is quite low in both of these cases, especially as the fraction of agents with the majority opinion grows. Because of this, we believe that setting θ to a low value will not strongly affect convergence time. In the following we develop a computational model to empirically validate this hypothesis.

5. GSSM SIMULATION RESULTS

We empirically investigated norm emergence through the GSSM process for a variety of values of θ. We studied a population of $N = 1000$ agents, and two different social networks, complete and scale-free. Each run of the simulation was executed for $10,000$ time steps. Following [4, 3] we say a norm has emerged when 90% of the population are in the same state. Folliwing Delgado, we set $\beta = 10.0$ for all simulations. All results are averaged over 25 runs.

Scale-free networks have a degree distribution of the form $P(k) \propto k^{-\gamma}$. We use the Albert-Barabàsi extended model from [1] to generate a scale-free network. Following [3] we use the parameters $m_0 = 4, m = 2, p = q = 0.4$ (Section 8.1

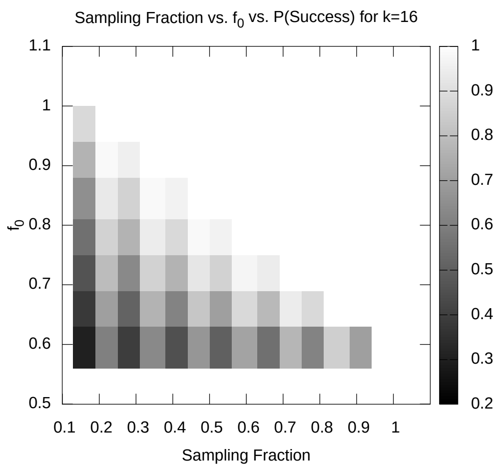

Figure 3: m (shown as the sampling fraction) vs. f_0 vs. $P(\textbf{Success})$(color) for $k = 16$.

provides more details about this process).

In the complete graph case each agent has $k = 999$ neighbors, thus Figure 2 describes the probability of not making a mistaken majority error.

The heterogeneous degree distribution of scale-free graphs complicates the picture, but most scale-free graph nodes will have low degree. For a single graph generated with the parameters specified above we found that 81% of the nodes had degree 16 or less. Thus Figure 3 can be used as an estimate.

Figure 4 shows the time for a norm to emerge for θ between 0.1 and 1.0, in increments of 0.1. For $\theta = 0.0$ none of the simulations converged within $10,000$ iterations, so those results are omitted.

Based on Section 4 we expect that norm emergence on complete graphs will not be affected by θ. This can be seen in Figure 4 as there is very little change in time to convergence.

For scale-free graphs we expect a much wider range of convergence time because of the larger probability of error; clearly θ introduced many errors. However, note the non-linear benefit of more information. As θ is increased the time to convergence decreases, substantially at first but then it tapers off after $\theta = 0.5$.

Clearly there is a difference in time to convergence, but what about effort? Figure 5 shows the effort to converge on a Log-linear plot. As θ is increased we see a increase in the amount of effort; however there are substantial differences between the increase in effort in scale free and complete graphs. For scale-free graphs the effort at $\theta = 1.0$ is approximately 2.5 times the effort at $\theta = 0.1$. In contrast, for complete graphs the change is an order of magnitude, from approximately 2×10^4 to 2×10^5.

5.1 Discussion

From the analyses and simulations above we see that the relationship between the level of information gathering and the impact on time to convergence can be leveraged to manage the effort of the system and thus minimize the Information-

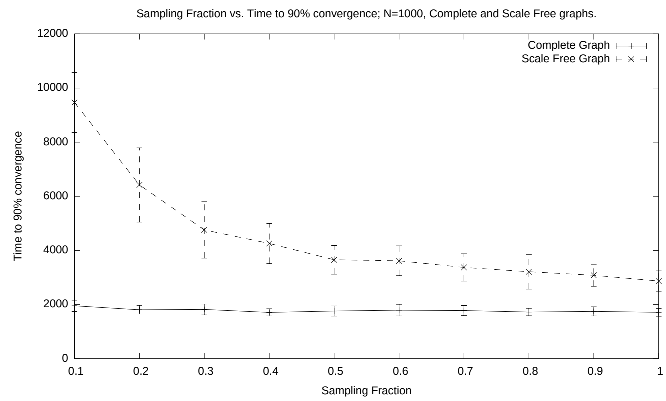

Figure 4: θ vs. Time to converge to 90%. Complete and Scale Free graphs. Mean of 25 runs, error bars indicate one standard deviation.

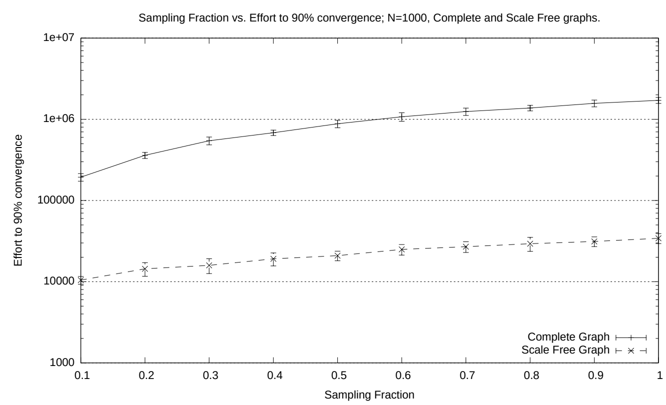

Figure 5: Log-linear plot of θ vs. Effort to converge to 90%. Complete and Scale Free graphs. Mean of 25 runs, error bars indicate one standard deviation.

Centric Convergence Cost. The key insight is that increasing information provides a non-linear benefit in terms of time to convergence on scale-free graphs. For complete graphs it is clear that the level of information has a very small effect on time to convergence, which in itself is also somewhat surprising.

Now, given a weighting of the cost of time versus the cost of effort, can we find the optimal level of information gathering?

Figure 6 shows how the optimal sampling frequency (i.e., resulting in the lowest ICCC) changes as we weigh time and effort differently. When c_t is low the cost of not converging is minimal, so the sampling fraction is low to minimize the effort. In contrast when c_t is high we can expend more effort to converge faster.

The dynamics for the complete graph are similar, however since there is no significant difference in convergence time over a variety of sampling fractions, the optimal strategy is to use the minimal converging sampling fraction, i.e., 0.1.

In Figure 6 we set $c_e = 1 - c_t$. Thus when $c_t = 1.0$ the optimal value of θ corresponds to the minimum point in Figure 4. For complete graphs the minimum can be found at $\theta = 0.4$, although we believe this is merely a statistically insignificant fluctuation.

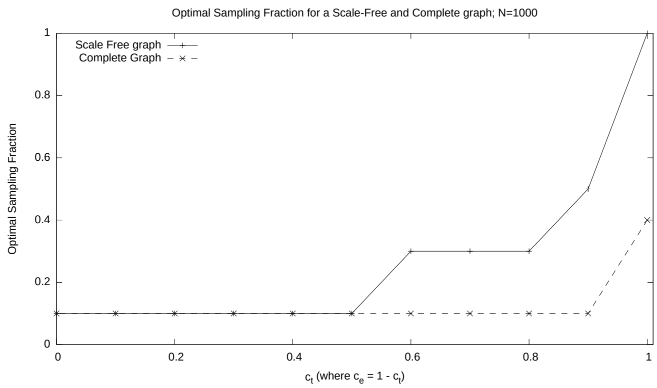

Figure 6: Optimal θ for value of c_t where $c_e = 1 - c_t$.

6. CONCLUSIONS AND FUTURE WORK

As the paradigm of collective learning becomes more important it becomes critical to explore its basic principles. In this paper we studied the tradeoff between effort and convergence time using a new model of norm emergence, *Generalized Simple Sampled Majority*. We noted some interesting results indicating how Information-Centric Convergence Cost can be minimized robustly, through sampling. One major avenue for further work is in larger (i.e. greater than 2) agreement spaces. This is an important direction for the field, as much norm emergence work focuses on the simple case of binary agreement spaces. Further, we have studied just two types of networks, complete and scale-free. We have seen significant differences in the ICCC between the two topologies; others have noted significant differences in convergence time between random, small-world and other types of graphs. A thorough exploration of these network topologies is needed. Finally, our analysis of success probabilities is limited to understanding the occurrence of mistaken majority errors. In future work we will consider the occurrence and impact of difference-in-strength errors as well. Further analysis of convergence time and Information-Centric Convergence Cost may be pursued through looking at the analysis of voter models (i.e. [13]).

7. ACKNOWLEDGMENTS

The authors would like to thank Shivani Agarwal for useful discussion and formulation of Equation (3); and Samarth Swarup for useful comments.

8. APPENDIX

8.1 Albert-Barabàsi Extended Model

The Albert-Barabàsi extended model depends upon four parameters; m_0 is the initial number of nodes, $m(\leq m_0)$ is the number of links that are added or rewired every step of the algorithm, p is the probability of adding links, and q is the probability of rewiring an edge ($p + q = 1$). The algorithm to generate the network is as follows. Start with m_0 isolated nodes, and at each step perform one of these three actions:

1. With probability p add m new links. Choose the start of the link uniformly randomly and the end point with distribution:

$$\Pi_i = \frac{k_i + 1}{\sum_j (k_j + 1)} \tag{4}$$

   where Π_i is the probability of selecting the ith node and k_i is the number of edges of node i. This process is repeated until m new links are added to the graph. If m links cannot be added we add as many as possible.

2. With probability q rewire m edges. Pick uniformly randomly a node i and link l_{ij} between node i and node j. Delete this link and choose another node k according to the probability distribution Π_i with the constraints that $k \neq i, j$ and l_{ik} does not already exist. Add the link l_{ik}.

3. With probability $1-p-q$ add a new node with m links – the new links with connect the new node to m other nodes chosen according to Π_i.

9. REFERENCES

[1] R. Albert and A.-L. Barabàsi. Topology of evolving networks: Local events and universality. *Physical Review Letters*, 85(24):5234–5237, 2000.

[2] C.-Y. Chong and S. Kumar. Sensor networks: Evolution, oppotunities, and challenges. *Proceedings of the IEEE*, 2003.

[3] J. Delgado. Emergence of social conventions in complex networks. *Artificial Intelligence*, 141(1-2):171–185, Oct. 2002.

[4] J. E. Kittock. Emergent conventions and the structure of multi-agent systems. In *Proceedings of the 1993 Santa Fe Institute Complex Systems Summer School*, 1993.

[5] K. Lakkaraju and L. Gasser. A unified framework for multi-agent agreement. In *Proceedings of AAMAS '07*, Honolulu, Hawaii, 2007.

[6] K. Lakkaraju and L. Gasser. Norm emergence in complex ambiguous situations. In *Proceedings of the Workshop on Coordination, Organizations, Institutions and Norms at AAAI*, 2008.

[7] N. A. Lynch. *Distributed Algorithms*. Morgan Kaufmann, 1997.

[8] P. Mukherjee, S. Sen, and S. Airiau. Norm emergence under constrained norm emergence under constrained interactions in diverse societies. In *Proceedings of AAMAS 2008.*, 2008.

[9] K.-D. Opp. How do norms emerge? an outline of a theory. *Mind and Society*, 3:101–128, 2001.

[10] G. Pottie and W. Kaiser. Wireless integrated networked sensors. *Communications of the ACM*, 43(5), 2000.

[11] S. Sen and S. Airiau. Emergence of norms through social learning. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence*, 2007.

[12] Y. Shoham and M. Tennenholtz. On the emergence of social conventions: modeling, analysis, and simulations. *Artificial Intelligence*, 94(1-2):139–166, July 1997.

[13] V. Sood, T. Antal, and S. Redner. Voter models on heterogenous networks. *Physical Review E*, 77, 2008.

[14] D. Villatoro and J. Sabater-Mir. Norm selection through simulation in a resource-gathering society. In *Proceedings of the 21st European Simulation and Modelling Conference (ESM07)*, 2007.

[15] A. Walker and M. J. Wooldridge. Understanding the emergence of conventions in multi-agent systems. In *ICMAS95*, pages 384–389, San Francisco, CA, 1995.

[16] J. L. Williams, J. W. F. III, and A. S. Wilsky. Approximate dynamic programming for communication-constrained sensor network management. *IEEE Transactions on Signal Processing*, 55(8), August 2007.

[17] F. Wu and B. A. Huberman. Social structure and opinion formation. arXiv:cond-mat/0407252v3 [cond-mat.other], 2005.