Social Learning Applications in Resource Constrained Networks

Ali Saidi
Systems Technology Research Lab
Motorola ARTC
Schaumburg, IL USA
ali.saidi@motorola.com

Mahesh V. Tripunitara
ECE Department
University of Waterloo
Waterloo, ON, Canada
tripunit@uwaterloo.ca

Mojdeh Mohtashemi
MITRE Corp., Bedford, MA, USA
MIT CSAIL, Cambridge, MA, USA
mojdeh@mit.edu

Abstract

Efficient design of social networking applications must take account of two guiding principles: the adaptive processes by which humans learn and spread new information, and the communication and technological constraints that in turn define the boundaries of human social behavior in virtual communities. In this paper, we introduce the concept of social learning in decentralized, resource-constrained networks. We present a mathematical model for spread of information and derive the optimum strategy that minimizes the total cost of learning in cooperative social networks. We then extend our model to allow individuals to limit their cooperative behavior in spreading their knowledge. Our results demonstrate that increased cooperation reduces the overall cost and accelerates the rate of learning new information.

1. Introduction

Although much remains unknown about the evolution of intelligence and learning, today a growing body of research suggests that natural selection must have favored particular sets of rules, or strategies, for animals to learn new information adaptively and efficiently [1]. Implicit in this statement is the hypothesis that the proecesses by which individual members of a population learn and react to their environment are optimal and evolutionarily stable.

Learning, or acquisition of new knowledge, can be asocial or social. Asocial learners, or innovating members of a population, generate new knowledge at some cost, while social learners acquire the new information by observing and adopting the behavior of others at no cost [2, 3]. Thus, social learning is inherently dependent on free utilization of the content of social, or public, information embedded in one's social network. Because social learners adopt new information for free, and are instrumental in the spread of the new knowledge, they are deemed to thrive due to a greater and immediate payoff (the cost-benefit differential). However, a population of pure social learners will gain no new knowledge in the absence of innovators. Today, many scholars agree that a winning strategy[1] is likely a combination of social and asocial behaviours, and that social learners must copy others discriminantly by selectively adopting the behavior of successful individuals [4, 5]. However, it remains an open problem as to what set of rules makes up a winning strategy for social learning in animals [6].

Though humans likely exercise similar strategies, what sets them apart from the rest of animals capable of social learning is the ability to generate and diffuse complex knowledge through sophisticated means of communication [7]. In the virtual world, where making social contact relies on the use of modern communication and information technology, the extent and nature of social behavior are influenced by resource constraints (e.g., energy, bandwidth, processing power, and memory). Conversely, the design of social learning applications must take account of social learning concepts that appear to govern human social behavior and the manner by which they spread information. Thus, the delicate fusion of these dichotomous domain constraints, namely the social and technological, underscores the guiding principles for efficient design of social learning systems.

Contrary to traditional social networks, in online settings information is generated and exchanged across virtual communities. Hence, regardless of whether a node generates new information (asocial learning) or scrounges it (social learning), there is a cost, albeit different, associated with both types of learning. In the virtual context, it is not as easy for social learners to acquire new knowledge or technology at the expense of innovators. Members of a social network must often

---

[1] A winning strategy, akin to game theory, is a strategy whose final payoff (sum of cost-benefit differentials over some time) is the greatest relative to other competing strategies.

establish trust and cooperate to learn new information. How to best design a social learning application that facilitates cooperation among members of a social network while keeping the communication or processing cost at bay, is the key question we aim to address here.

In this paper we introduce a computational framework for modeling the spread of information in decentralized, resource-constrained networks. We use this framework to design social learning systems that minimize the overall cost to the group.

The remainder of the paper is organized as follows. In Section 2, we provide two examples of potential applications for the design of social learning systems in resource-constrained networks. In Section 3, we present a class of strategies that decouples asocial and social learning into separate phases. In Section 4 we present the information diffusion and cost models for these strategies and derive the optimum strategy that minimizes the total cost of learning. In Section 5, we extend our basic model to allow individuals to limit their cooperative behavior in spreading their knowledge. Finally, in Section 6 we draw conclusions and point to some further directions for research.

2. Motivation

To make our discussion more concrete, we consider the learning behaviors and how they translate to specific computational/communication task with respect to two social networking applications. In the context of these applications, a social network comprises members that trust one another to cooperate in sharing application specific information when necessary despite the cost.

2.1. Security in social networks

The first application is the problem of spreading information about the identities of adversaries (e.g., spammers in an IP-based telephony system) in decentralized, resource-constrained settings. We consider the following question: how do the members of a social network cooperate to identify an adversary that has begun to victimize them? In this application members trust one another to not be an adversary [8, 9]. A context in which this problem arises is Spam over IP Telephony (SPIT) [10]. SPIT is the act of sending unsolicited messages over an IP-based telephony system. We consider settings such as mobile networks, in which there are resource-constraints; that is, communication is associated with a cost.

In this setting, innovation, or asocial learning, is analogous to the act of accepting calls from unknown

sources at the risk of getting spammed for which there is an associated cost. This cost may consist of measurable quantities, such as the cost of cell phone minutes and battery life. It can also relate to non-quantifiable effects, such as the psychological stress caused by being spammed. Observation, or social learning, relates to calling other members of the group in the hope of learning the identity of the spammer. In this case, given a relatively high degree of cooperation, it is fair to assume that the cost of innovation is much higher than the cost observation.

2.3. Targeted advertisement

In targeted advertisement, the objective is to have a subset of group members learn about and adopt a new product and subsequently recommend the product to other members of their network [11, 12]. The ultimate goal is to spread the knowledge about the product through word-of-mouth advertising which should lead to further product adoption by other group members.

In this application, the advertising source incurs all of the cost by offering large incentives to initially targeted group members to try a product, and smaller incentives to spread the word about the product to other members of the group. Hence, the marketing strategy of the advertiser is to minimize the total cost of incentives given to the members while effectively targeting groups and spreading the information about the product. Here, the initially targeted members who adopt the product are viewed as innovators (asocial learners) and those who receive information about the products are considered observers (social learners).

3. Strategy

A strategy is defined as a set of rules that specify which behavior should be adopted at each point in time to achieve a desired objective. Our objective is that all members of a social network, represented by the set F, learn a particular set of information in a decentralized setting. An optimal strategy is a strategy by which the objective is achieved while minimizing some cost parameter such as latency or resource-consumption. As a starting point for our work, we adopt a particular strategy that we call σ. Our strategy comprises three phases. In the first phase, members of F innovate, in the second, they observe, and in the last, they exploit. To simplify the analysis we assume that there is no cost associated with exploitation. This simple model represents a wide range of strategies which begin with innovation.

In the innovation phase of σ, some members of F are targeted to receive a particular set of application-

dependent information at a relatively high cost. We associate a per-member cost, c_s, with innovation. In the observation phase of σ, members of F stop innovating. For example, in the security application, they accept calls only from members of F, and in the advertising application we stop targeted ads and let the information propagate in the social network.

Inherent to the innovation and observation behaviors is the notion of time; we need a notion of time in our formulation for the cost. We model time discretely. Each unit of time is called a round; we have M rounds of innovation and N rounds of observation. We assume that in each round of the observation phase, a member who does not have the information has one instance of communication with another member of F. The former acquires the information if the latter has it.

4. Information Diffusion Model

We now study the innovation and observation phases of σ more closely. We wish to identify the convergence properties and cost characteristics of σ. Let X_n be the proportion of F that has learned the desired information (e.g. knows who the spammer is) at the end of round n; X_0 is the proportion that learned the information at the end of the innovation phase. We have the following progression for X_n.

$$X_{n+1} = X_n(1 - X_n) + X_n \quad for \ \ n \geq 0 \quad (1)$$

The intuition behind (1) is that in round $n+1$, the probability that someone that does not have the information (the proportion $1 - X_n$) calls someone that does, is X_n. An Implicit assumption in (1) is that group members fully cooperate in spreading their knowledge. In Section 5, we model a case of limited cooperation. We can immediately infer some properties of (1). For example, it converges at the limit, that is, for $X_0 \in (0,1)$ and $\varepsilon \in (0, 1 - X_0)$ there exists an N such that $X_N = 1 - \varepsilon$. We say that N is the number of rounds in which X converges.

To understand the behavior of the system depicted by (1), we seek a closed form solution for X_n. We rewrite (1) as follows.

$$X_{n+1} = X_n(2 - X_n) \quad for \ \ n \geq 0 \quad (2)$$

Equation (2) is the familiar logistic equation [13] with growth parameter $r = 2$. If we adopt $Y = X/2$, we can rewrite (2) in its canonical form as follows.

$$Y_{n+1} = 2Y_n(1 - Y_n) \quad for \ \ n \geq 0 \quad (3)$$

The dynamic properties of (3) have been well studied [14]. It has been shown that for the growth parameter $r > 3$, the equation exhibits chaotic behavior. Here $r = 2$, and thus the equation is stable and has the steady state value of $Y^* = 1/2$ (or $X^* = 1$). For $Y_0 < 1/2$, Y_n is monotonically increasing for $n \geq 0$, until convergence is attained. While closed form solutions exist for logistic differential equations for generic values of r, such solutions exist for only a handful of specific values of r for logistic difference equation. For $r = 2$ (i.e. equation (3)), the closed form solution is the following for $n \geq 0$.

$$Y_n = \frac{1}{2}\left[1 - e^{2^n \ln(1 - 2Y_0)}\right]$$
$$\equiv \ X_n = \left[1 - e^{2^n \ln(1 - X_0)}\right] = \left[1 - (1 - X_0)^{2^n}\right] \quad (4)$$

Logistic equations have been previously considered in the literature in different applications including security (in the context of worms) [15-17]. To our knowledge, our proposed model is the first to apply the logistic equation to the problem of information sharing in resource-constrained networks in a cost-minimization context.

We seek a meaningful characterization for the cost in terms of particular strategies we evaluate. Here, we study a set of strategies that decouples innovation and observation behaviors in separate phases. The following is our formulation for the total system cost.

$$C = \left[c_i \sum_{m=0}^{M-1} I_m\right] + \left[c_0 \sum_{n=M}^{M+N-1} O_n\right] \quad (5)$$

In the above formulation, C_i is the cost of innovation incurred by each member of the social network that chooses to innovate, C_0 is the cost of observation, I_k is the number of innovators in round k and O_l is the number of observers in round l. This is a rather abstract characterization of the cost. Our intent is to customize it based on particular strategies.

We adapt the cost function as expressed by (5) for σ; Equation (6) expresses this. The intuition is that in each round, each member in the $1 - X_n$ proportion incurs the cost of 1. The proportion X_0 incurs a cost of C_s in the innovation phase.

$$C(c_s, X_0) = c_s X_0 + \sum_{n=0}^{N-1}(1 - X_n)$$
$$(by \ Equation \ (4)) = c_s X_0 + \sum_{n=0}^{N-1}(1 - X_0)^{2^n} \quad (6)$$

We do not split the innovation behavior into rounds; we assume that it occurs in a single round. Also, we assume that C_s is normalized by the cost of

258

observation. Thus, each round in the observation phase consists of several instances of what we call legitimate communication, each of which incurs the normalized cost of 1. In the exploitation phase of σ, members use what they have learned in previous rounds.

We now make some observations about our system as expressed by (1) and (6). In Figure 1 we show the relationship between $C(c_s, X_0)$ and X_0 for a particular value of C_s. We have one real minimum for the values of X_0 that are relevant, that is, $X_0 \in (0, 1)$. It turns out that this is true for all meaningful values of X_0 and C_s. This is an interesting result because it demonstrates that there is a unique initial condition, X_0^*, that results in minimum total cost for any given choice of C_s. The following theorem establishes this general result.

Theorem 1 *For $X_0 \in (0, 1)$ and $C_s \geq 1$ a constant, $C(X_0)$ has exactly one real minimum.*

Proof. If $C_s = 1$ then the value of X_0 for which $C(X_0)$ is minimized is $X_0^* = 1$ (i.e. $N=0$) because then $C(X_0) = 1$. Otherwise, $C(X_0) > 1$. For $C_s > 1$, we set $y = 1 - X_0$ and take the derivative of $C(y)$ with respect to y.

$$C'(y) = \frac{dC(y)}{d(y)} = \sum_{n=1}^{N-1} 2^n y^{2^n - 1} + (1 - C_s)$$

We consider the roots of $C'(y)$ and apply Descartes' rule of signs [18]. $C'(y)$ has exactly one positive real root - there is exactly one change of sign across consecutive coefficients as $1 - C_s < 0$. There are no negative real roots because $C'(-y)$ has no sign change across consecutive coefficients. To show that the root represents a minimum in (0, 1), we first observe that $C'(y = 0) < 0$. Also, $C'(y = 1) = 2^N - (1 + c_S) > 0$ because $y \rightarrow 1 \Rightarrow N \rightarrow \infty$. That is, $C'(y)$ changes from negative to positive in (0, 1). This means that: (1) $C'(y)$ has a root in (0, 1), and, (2) this root corresponds to a minimum for $C(y)$. ∎

Now that we have established that there exists a real minimum for $X_0 \in (0, 1)$, we can find X_0^* using a root-finding algorithm [19]. Figure 2 shows the relationship between X_0 and N for $\varepsilon \approx 10^{-38}$ (10^{-38} is approximately the smallest single-precision floating point number in MATLAB.) We see an exponential decrease in N as X_0 increases. Figure 3 shows the
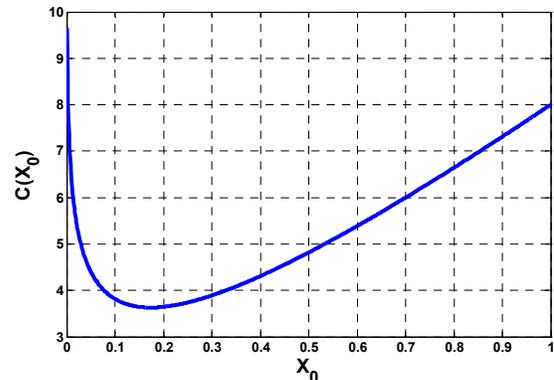
Figure 1. The relationship between X_0 and $C(c_s, X_0)$ for $C_s = 8$.
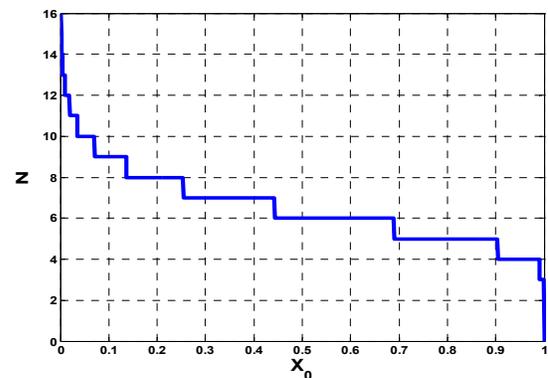
Figure 2. The relationship between X_0 and N for $\varepsilon \approx 10^{-38}$.
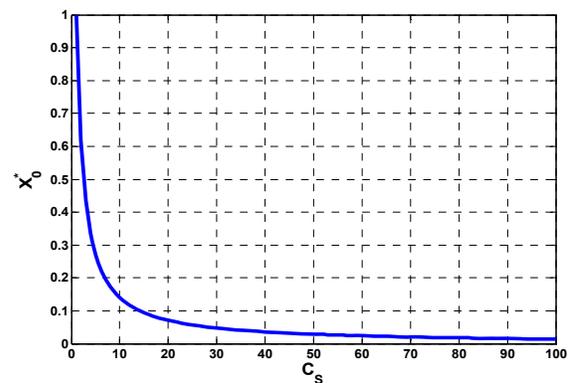
Figure 3. The relationship between C_s and X_0 for which the total cost is minimized.

relationship between the normalized cost of innovation, and the X_0^* for which the total cost, is minimized. We consider only $C_s \geq 1$; the cost of innovation is at least as much as engaging in observation. The graph shows that the decrease in X_0^*

259

slows significantly with increase in C_s beyond a certain threshold.

Figure 4 shows the relationship between C_s and N for which cost is minimized, N^*, across all possible values of X_0. We observe that N^* increases but starts to converge. This is consistent with Figure 3; N^* is a function of X_0^* and X_0^* exhibits little variation for large values of C_s. Finally, Figure 5 shows the relationship between C_s and the minimum total cost across all possible values for X_0. This is also consistent with Figure 3. That is, the minimum total cost grows linearly with C_s for relatively constant X_0. The convergence of the minimum total cost is interesting, because this suggests that even for high values of C_s, we can hope to establish an upper bound on the total cost provided we choose the correct value for X_0, that is, we innovate for as long as is needed to minimize the cost.

5. Modeling Limited Cooperation

We now extend our basic model to allow individuals to limit their cooperative behavior. The extent of cooperation could be based on an individual's decision to control cost during the observation phase and is governed by resource constraints. It has been shown that spread of new information among members of networks has likely played a key role in the evolution and sustenance of cooperation in human populations [21, 22].

Here, we examine the effect of degree of cooperation among members of a network on the overall cost and the rate of spread of new knowledge. We introduce a new parameter α, to represent the degree of cooperation by those who have the information in the observation phase. Then, the information diffusion model in the observation phase can be described by the following recursion.

$$X_{n+1} = \alpha X_n (1 - X_n) + X_n \quad where \quad \alpha \in (0,1] \quad (7)$$

Although the growth of X_n is now a function of α, the total cost as a function of X_0 is still given by (6):

$$C(c_s, X_0) = c_s X_0 + \sum_{n=0}^{N-1} (1 - X_n) \quad (8)$$

The following theorem establishes that the total cost has only one real minimum, and thus is an extension of Theorem 1 when members of a network may no longer be fully cooperative.
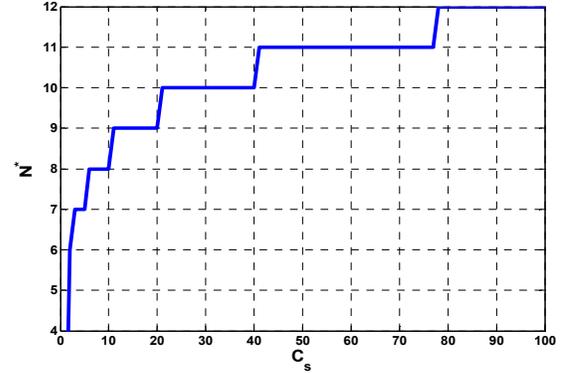

Figure 4: The relationship between C_s **and** *N* **for which the total cost is minimized.**
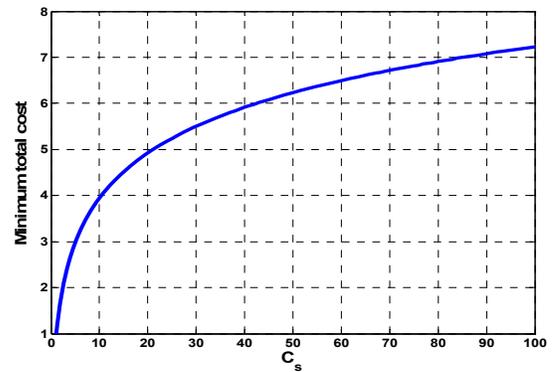

Figure 5. The relationship between c_s **and the minimum total cost across all values of** X_0 **.**

Theorem 2 *For* $X_0 \in [0,1]$, $\alpha \in (0,1]$, *and constant* $c_s \geq 1$, $C(X_0)$ *has exactly one real minimum,* $C_{\min}(X_0) \leq C_S$.

Proof (by contradiction): First note that $C(X_0 = 0) = \infty$ and $C(X_0 = 1) = C_S$ which then implies that there exists at least one real minimum, and the global minimum in [0,1] is at most C_s. We now establish that there is exactly one real minimum. Re-write (7) as

$$1 - X_{n+1} = (1 - \alpha X_n)(1 - X_n) \quad where \quad \alpha \in (0,1] \quad (9)$$

It is clear from (9) that $1 - X_n$ declines at a monotonically decreasing rate of $1 - \alpha X_n$. As X_0 increases from 0 to 1, the first term on the right hand side of (8) increases at a constant rate while the second term declines at a monotonically decreasing rate. Now suppose that (8) has more than one minimum. Then it follows that for some X_0 greater than the value for which the first minimum occurs, the rate of change of the second term on the right hand side of (8) would

260

have to increase to attain a second minimum. This contradicts the monotonicity premise implied by (9). ∎

Note that, unlike the full cooperation model, the minimum total cost may occur at $C(X_0 = 1) = C_S$. That is, when the cost of observation is relatively high, asocial learning is the most cost effective learning strategy. This happens when the degree of cooperation is low (i.e., α is small).

It is possible to numerically find the value of X_0 for which the cost is minimized. However, it can be computationally expensive. We begin by transforming (7) to the more familiar logistic equation with growth parameter $r = 1 + \alpha$. Let $Y = \frac{\alpha}{1 + \alpha} X$ and rewrite (7) in its canonical form as follows:

$$Y_{n+1} = (1 + \alpha)Y_n(1 - Y_n) \quad for \ n \geq 0 \qquad (10)$$

An analytical solution to (9) as a function of Y_0 in matrix form is given in [20]:

$$Y_n = YT^n e_1 \qquad for \ n \geq 0 \qquad (11)$$

where T is defined by:

$$T_{j,k} = (-1)^{j-k} \binom{k}{j-k}(1 + \alpha)^k , \qquad (12)$$

$Y = [Y_0, Y_0^2, Y_0^3, \cdots]$ and $e_1' = [1, 0, 0, \cdots]$ (where e_1' is the transpose of e_1). To solve for Y_n, we need to compute the nth power of T. Although T is an infinite matrix, only the finite $2^n \times 2^n$ upper-left corners of the matrix are needed. Substituting (10) into (8) leads to a cost function which is a polynomial in X_0 from which we can obtain the initial condition that minimizes total cost.

Figure 6 illustrates the relationship between $C(C_S, X_0)$ and X_0 for $C_S = 8$ and $\alpha = 0.5$. As expected, the total cost has only one minimum for $X_0 \in [0, 1]$. It is interesting to note that the minimum total cost in Figure 6 is higher than that of Figure 1. In other words, when members of a network are not fully cooperative, they collectively incur a higher cost. In Figure 7, the minimum total cost is plotted as a function of α for $C_S = 15$. As can be seen, the minimum total cost is a monotonically decreasing function of the degree of cooperation: a less cooperative group of individuals is doomed to pay a higher price for social learning and spread of knowledge in the long run.

Another interesting result is that of Figure 8, where the relationship between the number of rounds (N) and the degree of cooperation (α) is plotted for those values of N where the total cost is minimized. As the
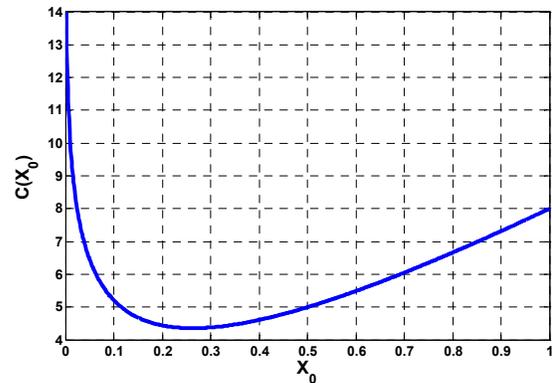

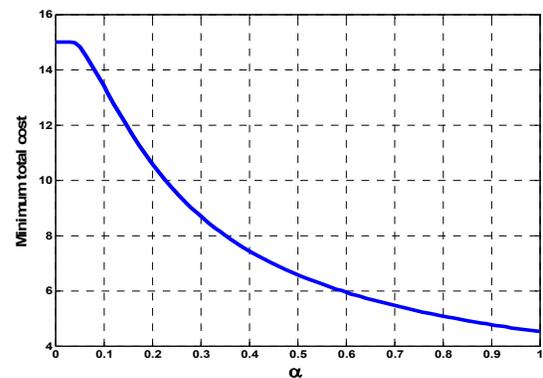**Figure 6. The relationship between X_0 and $C(C_S, X_0)$ for $C_S = 8$ and $\alpha = 0.5$.**

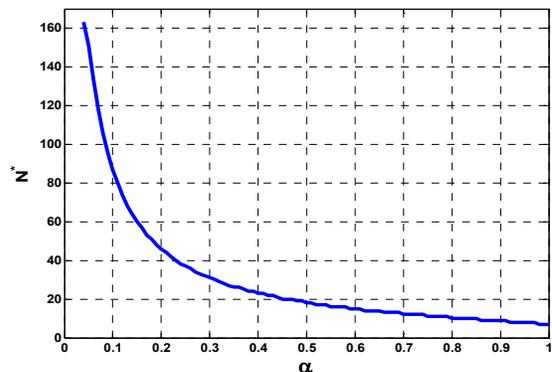**Figure 7. The relationship between α and the minimum total cost for $C_S = 15$.**

Figure 8. The relationship between α and N for which the total cost is minimized across all values of X_0 for $C_S = 15$.

degree of cooperation increases, learning of new knowledge is expedited and new information is propagated at a faster rate. Putting the results of

261

Figures 6-8 together, it becomes clear that the benefit of cooperation is two-fold: it pays off to cooperate because it reduces the cost and increases the rate of learning new information.

6. Concluding Remarks

We introduced a novel computational framework based on social learning concepts and the logistic equation for efficient learning and spread of new information in decentralized, resource constrained networks. We showed that when members of a social network cooperate to learn new information, the overall cost can be minimized. We then developed a model of limited cooperation and demonstrated that both the cost and rate of learning new information are dependent on the extent of cooperation among members of a social network.

Much research in evolutionary biology, anthropology, and economics has been devoted to understanding the conditions under which cooperation has evolved when resources are finite [23, 24]. More recently, it has been shown that learning and transmission of new information has likely been critical in the evolution of cooperation [21, 22]. Here we have shown the converse—that when members of a network cooperate to learn new knowledge, learning and transmission of information can be achieved at a faster rate and a lower cost.

There is considerable scope for future work. One aspect concerns applications of our approach to some of which we allude in Section II. Another extension is to formulate an optimal strategy when it is critical for members of a network to learn and disseminate more than one informational component (e.g., identities of multiple attackers). Finally, we plan to investigate a repertoire of strategies from a more foundational standpoint, where an agent-based simulation framework paired with game theoretic notions is used to examine the optimality of various strategies in a changing environment.

7. References

[1] Albert Bandura, "Social Learning Theory," Prentice-Hall, Englewood Cliffs, N.J. 1977.

[2] C. B. Harley, "Learning the Evolutionary Stable Strategy," J Theoretical Biology **8** (1981): 611-633.

[3] M. Enquist, K. Eriksson, and S. Ghirlanda, "Critical social learning: A solution to rogers's paradox of nonadaptive culture," American Anthropologist, 109(4):727–734, December 2007.

[4] K. N. Laland. "Social learning strategies. Learning & Behavior," 32(1):4–14, February 2004.

[5] K. N. Laland and W. Hoppitt, "Do animals have culture? Evolutionary Anthropolology," volume 12: 150-159, 2003.

[6] Kevin Laland and Luke Rendell, "Cultaptation: Social learning strategies tournament," January 2008. http://www.intercult.su.se/cultaptation.

[7] M. Tomasello, "The human adaptation for culture," Ann. Rev. Anthropol. 28:509-529, 1999.

[8] Vijay A. Balasubramaniyan, Mustaque Ahamad, and Haesun Park, "Callrank: Using call duration, social networks and pagerank to combat SPIT," Proceedings of Fourth Conference on Email and Anti-Spam (CEAS), August 2007.

[9] P. Oscar Boykin and Vwani P. Roychowdhury, "Leveraging social networks to fight spam," IEEE Computer, 38(4), April 2005.

[10] J. Rosenberg and C. Jennings. "The session initiation protocol (SIP) and spam," IETF RFC 5039, January 2008.

[11] J. Leskovec, L. A. Adamic, and B. A. Huberman, "The dynamics of viral marketing," ACM Transactions on the Web, 1(1):5, 2007.

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

[13] Robert M. May, "Simple mathematical models with very complicated dynamics," Nature, 261:459–467, 1976.

[14] S. Wolfram. A new kind of science. Wolfram Media, Inc., Champaign IL, USA, 1 edition, May 2002.

[15] Z. Chen, L. Gao, and K. Kwiat, "Modeling the spread of active worms," Proceedings of the Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM), volume 3, pages 1890 – 1900, April 2003.

[16] Z. Chen and C. Ji, "Spatial-temporal modeling of malware propagation in networks," IEEE Transactions on Neural Networks, 16(5):1291–1303, September 2005.

[17] M. Liljenstam, David M. Nicol, Vincent H. Berk, and Robert S. Gray, "Simulating realistic network worm traffic for worm warning system design and testing," Proceedings of the 2003 ACM Workshop on Rapid Malcode (WORM), pages 24–33, October 2003.

[18] Eric W. Weisstein, "Descartes' Sign Rule," From MathWorld-- http://mathworld.wolfram.com/DescartesSignRule.html

[19] The Mathworks, Inc. MATLAB. http://www.mathworks.com/products/matlab/, 2008.

[20] S. Rabinovich, G. Berkolaiko, S. Buldyrev, A. Shehter, S. Havlin, "`Logistic Map': an Analytical Solution," Physica A 218 (3-4) 457-460, 1995.

[21] M. Mohtashemi and L. Mui, "Evolution of indirect reciprocity by social information: the role of trust and reputation in evolution of altruism," Journal of Theoretical Biology, 223:523–531, 2003.

[22] H. Brandt and K. Sigmund, "Indirect reciprocity, image scoring, and moral hazard," PNAS, 102(7): 2666–2670, 2005.

[23] R. Axelrod, The Evolution of Cooperation, Basic Books, New York, 1984.

[24] J. Henrich, R. Boyd, S. Bowles, H. Gintis, E. Fehr, "In search of homo economicus: experiments in 15 small-scale societies,"Am. Econ. 91,73–78, 2001.