Digital Signal Processing 47 (2015) 17–24

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Social learning with heterogeneous agents and sequential decision making

Yunlong Wang, Petar M. Djurić *

Department of Electrical and Computer Engineering, Stony Brook University, Stony Brook, NY 11794, United States

ARTICLE INFO

Article history:
Available online 2 March 2015

Keywords:
Social learning
Bayesian learning
Sequential decision making
Irrational agents
Random/deterministic decision making

ABSTRACT

In this paper, we consider the problem of social learning in a network of agents where the agents make decisions sequentially by choosing one of two hypotheses on the state of nature. Each agent observes a signal generated according to one of the hypotheses and knows the decisions of all the previous agents in the network. The network contains two types of agents: rational and irrational. A rational agent makes a decision by not only using its private observation but also the decisions of each of the agents which already made decisions. To that end, the agent employs Bayesian theory. An irrational agent makes a decision by ignoring the available information and by randomly choosing the hypothesis. We analyze the asymptotic performance of a system with rational and irrational agents where we study rational agents that use either a deterministic or random decision making policies. We propose a specific random decision making policy that is based on the social belief and the private signals of the agents. We prove that under mild conditions the expected social belief in the true state of nature tends to one if the rational agents use the proposed random policy. In a network with rational agents that use deterministic policy, the conditions for convergence are stricter. We provide simulation results on the studied systems and compare their performance.

1. Introduction

An important issue in social learning is how agents make decisions and learn from the decisions of other agents. When agents make decisions based on private and imperfect information, it is natural that they also use information in the decisions of other agents made in a similar situation. In this work, we consider agents that make their decisions on one of two hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ one at a time in a sequential manner. When an agent makes its own decision, it broadcasts it to all the agents that will subsequently be making decisions. The graph that describes our system is a directed graph shown in Fig. 1. We note that the motivation for studying this network is understanding some basics of social learning with random decision making rather than a particular application of signal processing.

Applications of sequential decision making can be found in many fields including distributed detection in wireless sensor networks. There, the need for low-cost and low-power devices forces every sensor to aggregate all its information into a one bit decision [1,2]. In [3], a comprehensive introduction to this problem was provided. Another major application is the understanding of social learning in multi-agent systems. In these systems, the agents make decisions not only from their private observations, but also from the decisions made by others [4]. As indicated in [5], the study of social learning addresses the problems of modeling the interaction of the agents in the network while learning takes place.

In social learning systems, an agent can learn from the decisions of agents that already made them by using either non-Bayesian approaches or the Bayesian methodology. For example, in [6–8], non-Bayesian social learning methods were studied. In [4,9–15], the agents had Bayesian social learning strategies in obtaining information from previous decisions. Although the Bayesian learning methods provide a fast convergence rate to one of the hypotheses, the agents that employ the Bayesian machinery may ignore their private observations and *herd* on the wrong decision. Some work on this issue can be found in [4,9], where the interest was in studying herding behavior and information cascades. In [11], the authors provided the sufficient conditions for asymptotic learning in Bayesian social learning systems. In [16], the author studied this problem from the perspective of quickest time change detection. Recently, random decision making was addressed in [17]. In [18], the authors considered the effect of the decisions of subsequent agents on the utility of a current agent by using a Chinese restaurant game model. They proposed an optimal strat-

* Corresponding author.
E-mail addresses: yunlong.wang@stonybrook.edu (Y. Wang), petar.djuric@stonybrook.edu (P.M. Djurić).

http://dx.doi.org/10.1016/j.dsp.2015.02.016

egy for decision making. This approach was applied to cognitive radio networks [19]. Another extension of this problem was a scenario with noisy links [20], where the decisions were broadcasted through a noisy channel and thereby the decisions could be randomly flipped or erased. An overview of models and techniques for studying social learning can be found in [5,10]. In this paper, our study is within the framework of Bayesian social learning.

In the analysis of social learning systems, one typical question is the following: Does the probability of error in decision making converge to zero as the number of agents goes to infinity? The answer depends on (a) whether the log-likelihood ratio of private observations is bounded and (b) whether the memories of the agents are finite (the size of the memory of an agent is the number of decisions of other agents available to the agent). In [21–23], the problem with finite memory was addressed. In [21], a system where each agent could only observe the decision of its immediate previous agent was investigated. It was shown that if the likelihood ratio test is employed for decision making, the error probability does not converge to zero if the ratio is bounded. In [22], social learning with sampling of past decisions was addressed. There, each agent made inference by the Bayesian method and using very few samples from the decision history of the agents. In the case of bounded log-likelihood ratios, a non-Bayesian decision policy was proposed whose probability of decision error converges to zero [23]. In [11,24], the agents had infinite memory, and the conditions for asymptotic learning were listed.

In the literature, it is common to assume that the agents know the decision making policy of the other agents, or that the policy is identical for every agent (e.g., [4,16,20]). In this work, we address the problem of social learning in a system of heterogeneous agents. The agents can be either rational or irrational and they make the decisions sequentially, one at a time. A rational agent makes a decision according to the posterior of the hypotheses (conditioned on the decisions of the previous agents and its private observation). By contrast, an irrational agent simply ignores all its available information and makes a decision by flipping a fair coin. The agents do not know which of them are rational and irrational but they know the percentage of irrational agents in the network.

More related work on social learning with heterogeneous agents can be found in [25–27]. In [25], the problem of misinformation spread was addressed, where the social system was modeled by a random network and the agents used a gossip style method to update their beliefs from the exchanged information. In [26], with a binary voter model, the authors studied the effect of stubborn agents in social systems. There, the stubborn agents are agents that never change their beliefs. In [27], the authors discussed the relationship between the convergence rate of an average-based learning method and the homophily of the multi-type random networks. In all of these papers, however, there is no randomness in the decisions making.

In this paper, we study two cases. In the first, the rational agents for decision making use a deterministic policy whereas in the second, they employ a random policy. In [17], we proposed a random decision making policy and compared it to the deterministic policy. In that comparison, all the agents were rational. We showed that with the deterministic policy, the probability of decision error may fail to converge to zero due to an information cascade. On the contrary, with the random policy, the probability of decision error is guaranteed to converge to zero. In this work, because of the presence of irrational agents, this probability does not converge to zero. However, if an agent is rational, its probability of a decision error converges to zero. Furthermore, if the social belief is defined by the posterior on the true state of nature conditioned on all the known decisions, then we show that it *always* converges to one in probability if the rational agents adopt the random decision making policy. This is not the case when they

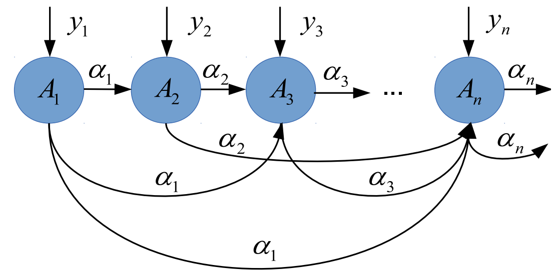

Fig. 1. Social learning in a sequential decision making system. The symbol y_n is the private signal of agent A_n, and α_n is the decision of this agent.

use the deterministic policy. In summary, the main contributions of this paper are (1) a proof that the presence of irrational agents does not affect the asymptotic probability of the rational agents if they adopt random policy and (2) a proof that in a system with random agents herding does not take place.

The paper is organized as follows. In the next section we describe the models of the sequential system and explain the social learning process. In Section 3, we introduce both the deterministic and the random decision making policies. The analyses of the convergence of the social belief and the probability of decision error are provided in Section 4. Simulation results are presented in Section 5, and concluding remarks in Section 6.

2. Problem formulation

The process of social learning is illustrated in Fig. 1. We consider the decision making problem in networks of agents A_n, $n \in \mathbb{N}^+$, where the agents make decisions and broadcast them sequentially one at a time to all the subsequent agents. Mathematically, each agent A_n receives an independent private observation y_n that is generated according to one of the following two hypotheses:

$$\begin{aligned} \mathcal{H}_1: &\quad y_n \sim \phi_1(y_n) \\ \mathcal{H}_0: &\quad y_n \sim \phi_0(y_n), \end{aligned} \tag{1}$$

where $\phi_k(y_n)$ denotes the distribution of observation y_n under $\mathcal{H}_k$, $\forall k \in \{0,1\}$. We assume that these two distributions are known to all the agents and that each agent has noninformative prior probabilities on the hypotheses. Thus, we have $p(\mathcal{H}_0) = p(\mathcal{H}_1) = 1/2$. Let $\log\left(\frac{\phi_1(y_n)}{\phi_0(y_n)}\right)$ be the log-likelihood ratio (LLR) of the hypotheses. We discriminate two types of LLRs, bounded and unbounded. An LLR is bounded if there exist two finite real numbers m and M, such that $\forall\, y_n$, $\log\left(\frac{\phi_1(y_n)}{\phi_0(y_n)}\right) \in [m, M]$. Otherwise, the LLR is unbounded.

In the network, there is a positive probability that an agent A_n behaves irrationally. An irrational agent A_n, makes its decision by drawing from a Bernoulli distribution with parameter $\eta \in (0,1)$, i.e.,

$$\alpha_n \sim Ber(\eta), \tag{2}$$

where α_n denotes the decision made by A_n and $\alpha_n \in \mathcal{A} = \{0,1\}$. Let I_n be an indicator function taking a value of one if the agent A_n is irrational; otherwise, it is zero. Then we set that $p(I_n = 1) = \xi < 1$, $\forall n \in \mathbb{N}^+$. In the rest part of this paper, we assume that both η and ξ are known by all the agents in the network.

Here we remark that the model in (2) is just one possible formulation of irrational behavior. In our work, an irrational agent is an agent that ignores all the available information to it and makes decision randomly according to a certain law. The simplicity of this irrational behavior model notwithstanding, allows us to