Bayesian learning: engineered versus social networks

Kimon Drakopoulos, Asuman Ozdaglar, John N. Tsitsiklis
LIDS, MIT
77 Massachusetts Avenue
Cambridge, MA 02139
{kimondr,asuman,jnt}@mit.edu
http://web.mit.edu/jnt/www/home.html

We consider a classical model of distributed decision making, originally developed in engineering contexts [2,4,5], but which has also attracted recent attention in the social sciences [1]. Suppose that there are two hypotheses on the state of the world and that each of several nodes (or agents, decision makers, sensors, etc.) receives a conditionally independent measurement, with a different distribution under each hypothesis. The nodes act in sequence, and their interactions are modeled by a given directed acyclic graph. Each node computes a binary message, on the basis of its own observation and the messages it receives from its predecessors, and sends it to its successors. If we interpret the binary message as a decision in favor of one or the other hypothesis, a central question is whether the network will "learn," that is, whether the decision by the nth node is correct, with probability that approaches zero as $n \to \infty$.

In the engineering version of the problem, the decision making rules of each node are chosen by a global designer, whose objective is to miminize the probability of an incorrect decision by the last node. In contrast, in the social science literature each node is assumed to act myopically, that is, to choose a message that minimizes its own probability of error. We overview positive and negative results under the two alternative formulations, for two simple network structures: a tandem, whereby a node only hears the message of its immediate predecessor, and an augmented tandem, where a node hears the message of its K immediate predecessors (with $K > 1$). While for $K = 1$, the assumptions needed for learning are more or less the same under the two formulations, this is no longer the case for $K > 1$ [3].

We also discuss a more complex model of selfish (but somewhat altruistic) nodes where the cost to a node is the discounted sum of the error probabilities of itself and all of its successors. We focus on Nash equilibrium strategies and indicate that the partial altruism introduced in this model is not enough to reverse some of the negative learning results due to selfish behavior.

References

1. D. Acemoglu, M. Dahleh, I. Lobel, and A. Ozdaglar, "Bayesian Learning in Social Networks," NBER Working Paper No. 14040, May 2008.

2. T. M. Cover, "Hypothesis Testing with Finite Statistics," *Ann. Math. Statist.,* 40, 1969.

3. K. Drakopoulos, "Observational Learning with Finite Memory," M.S. thesis, Dept. of EECS, MIT, June 2011.

4. J. Koplowitz, "Necessary and Sufficient Memory size for m-Hypothesis Testing," *IEEE Transactions on Infromation Theory*, 21, 1975.

5. J. N. Tsitsiklis. "Decentralized detection," In *Advances in Statistical Signal Processing*, pages 297344. JAI Press, 1993.

Topic: learning theory
Preference: oral