Detection Games Under Fully Active Adversaries

Benedetta Tondi, *Member, IEEE*, Neri Merhav, *Fellow, IEEE*, Mauro Barni *Fellow, IEEE*

Abstract

We study a binary hypothesis testing problem in which a defender must decide whether or not a test sequence has been drawn from a given memoryless source P_0 whereas, an attacker strives to impede the correct detection. With respect to previous works, the adversarial setup addressed in this paper considers an attacker who is active under both hypotheses, namely, a fully active attacker, as opposed to a partially active attacker who is active under one hypothesis only. In the fully active setup, the attacker distorts sequences drawn both from P_0 and from an alternative memoryless source P_1, up to a certain distortion level, which is possibly different under the two hypotheses, in order to maximize the confusion in distinguishing between the two sources, i.e., to induce both false positive and false negative errors at the detector, also referred to as the defender. We model the defender-attacker interaction as a game and study two versions of this game, the Neyman-Pearson game and the Bayesian game. Our main result is in the characterization of an attack strategy that is asymptotically both dominant (i.e., optimal no matter what the defender's strategy is) and universal, i.e., independent of P_0 and P_1. From the analysis of the equilibrium payoff, we also derive the best achievable performance of the defender, by relaxing the requirement on the exponential decay rate of the false positive error probability in the Neyman–Pearson setup and the tradeoff between the error exponents in the Bayesian setup. Such analysis permits to characterize the conditions for the distinguishability of the two sources given the distortion levels.

Index Terms

Adversarial signal processing, binary hypothesis testing, statistical detection theory, game theory, the method of types.

M. Barni and B. Tondi are with the Department of Information Engineering and Mathematical Sciences, University of Siena, Siena, ITALY, e-mail: {benedettatondi@gmail.com, barni@dii.unisi.it}; N. Merhav is with the the Andrew and Erna Viterbi Faculty of Electrical Engineering - Israel Institute of Technology Technion City, Haifa, ISRAEL, email: {merhav@ee.technion.ac.il}.