On-line monitoring of high-dimensional data streams

Zhaojun Wang and Changliang Zou

Abstract

There has been increasing interest in how to monitor high-dimensional data streams for real-time detection of abnormal activities in a data-rich circumstances. We are interested in detecting an occurring event as soon as possible, but we do not know which subset of data streams will be affected by the event. By illustrating the connection with the problems of detecting heterogeneous mixtures, a new control scheme is developed based on a powerful goodness-of-fit test of the local cumulative sum statistics from each data stream. Both asymptotically theoretical analysis and numerical results show that the proposed approach is able to balance protection against various the fractions of affected streams, and generally outperforms existing methods.