SaTC: Core: Small: Statistical Inference in Adversarial Environments: Fundamental Limits and Algorithms

Overview

Advances in data acquisition and high-dimensional information processing are rapidly transforming various technological, social, and economic domains, such the Internet, telecommunication, electricity grids, water management systems, social networks, and medical records. Empowered by these advances, such domains are evolving into complex networked platforms in which high-dimensional and complex data is being routinely generated, exchanged, and processed for monitoring, inferential, and resource management purposes. Due to the inherent scale of the data and complexity of the processes involved on the one hand, and strong coupling between the physical operations of the network and their respective cyber aspects on the other hand, the flow of information in complex networks becomes increasingly more vulnerable to adversarial intrusions, which are expected to grow well into the future. Hence, benefiting from the full extent of such enabling technologies is feasible only when appropriate security measures are implemented so that first, the adversaries are identified very quickly, and secondly, the data processing mechanisms are designed to be robust against such adversaries. Understanding and addressing the security concerns, hence, has a pivotal role in designing and operating the existing and emerging complex networks.

The underlying vision for this framework is that security against active adversaries can never be absolute. For example, an adversary with omniscient knowledge about the network can simply compromise all legitimate operations. Thus, ensuring security always involves a tradeoff: devoting some resources to protect against a larger class of attacks incurs a performance loss. This research goal of this proposed program is to characterize the fundamental limits of statistical inference in the potential presence of active adversaries. Specifically, we propose to produce a coherent statistical decision-theoretic framework that will characterize the fundamental limits of secure statistical inference, and provide design principles for secure strategies that reach these limits and are amenable to real-time and scalable implementation. The existing approaches are often designed without regard for the inherent security, resulting in a pattern in which security vulnerabilities are addressed as they are discovered. This proposal aims to break free of this cycle, developing inferential algorithms that are secure at their core against large classes of adversaries.

Keywords: sequential statistics; compound inference, secure inference, quickest detection.

Intellectual Merit

The proposed research advances the theory and practice of the statistical decision theory and lies at the crossroad of estimation and detection theories, sequential statistics, and stochastic optimization. The key drive underpinning the proposed research is that the existing art lacks a theory that is versatile enough to address the growing challenges pertinent to secure and real-time inference from complex and dynamic data, in which inference cannot be reduced to hitherto known inference operations. The developed theory provides new insights and tools for performing real-time compound estimation and detection operations in a jointly optimal fashion, while recognizing the complexities and structures of data in adversarial environments. It also establishes novel decision-making rules under stochastic uncertainties and solves some long-lasting problems, which have been open since 1940s.

Broader Impacts

Societal, economical, and environmental interests increasingly emphasize the significance of enhancing security in distributed data networks. The proposed research stands to increase the resiliency of the emerging complex networks for enabling more secure operations. This is in line with the vision of various regulatory and research agencies for major infrastructures such as electricity grids, transportation systems, and telecommunication networks. The PI's educational plan is composed of several parts aimed at students ranging from the high school to graduate level and from several demographics. In collaboration with the Engineering Ambassadors program at Rensselaer, the PI is designing interactive presentations for high-school students about modern notions of cyber security, which involve undergraduate students as assistants in delivering these presentations. This complements the PI's plans for engaging them in research.