Multi-faceted approach for cyber-resilience of extreme scale systems

Mahantesh Halappanavar, Sutanay Choudhury, Samrat Chatterjee, John Johnson
Pacific Northwest National Laboratory, Richland, WA
{FirstName.LastName}@pnnl.gov

Focus area: Extreme Scale Data, Knowledge, and Analytics
Narrative:

The success of DOE research efforts is driven by the collaborative efforts of its research community that is critically dependent on availability of extreme scale computing resources, data from experimental resources and simulations, and the ability to share these resources. As a consequence, cyber security plays an important enabling role in the process of scientific discovery. However, there are tremendous challenges, especially arising from the advent of extreme scale resources resulting in unprecedented amounts of data and compute resources. We therefore believe that a multi-faceted approach is needed to address the cyber security challenges of DOE. In the following discussion, we outline a framework for real-time decision-making built upon two important areas, *game theory* and *real-time anomaly detection*. We will briefly describe several facets that enable this framework.

Modeling system vulnerabilities: DOE supports complex machinery for scientific discovery that comprises of extreme scale leadership class machines, experimental resources, data repositories, high speed networks interconnecting different resources, and numerous users accessing these resources from various systems. Vulnerabilities exist at different levels and are interconnected, with potential for cascading effects on the system. The ability to model these interdependencies in a scalable manner is therefore an important research topic. Current methods for modeling such as attack graphs [1] and Bayesian attack graphs are fundamentally limited in their ability to scale to the levels of interest to DOE. Multi-level methods can be a potential approach for scaling attack graphs by coarsening the graph at several levels and accumulating vulnerabilities and exploits at those levels. However, tradeoffs between accuracy and scalability need further study.

Optimal responses: Scalable models for the cyber system can be used to compute optimal responses from system administrators. Several optimization techniques have been proposed in the literature. However, many of these methods are based on complete information and apply to static snapshots of the system [2]. Game theoretic approaches therefore hold the promise to overcome some of the limitations [3]. Through dynamic interaction of attackers and defenders of a cyber-system, game-theoretic methods can be used to compute the most effective strategy that a defender can deploy against a particular attack for a given state of the system. In addition to computational costs that arise from the need to solve linear and non-linear programs modeling the games, the current game-theoretic methods suffer from a fundamental limitation in their lack of uncertainty quantification from different sources -- a facet that we discuss next.

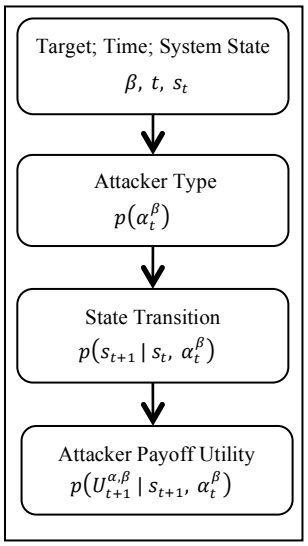

Figure 1. Uncertainty quantification framework.

Uncertainty quantification techniques: In the context of a game-theoretic approach, uncertainties arise from several sources. For example as illustrated in Figure 1, assessing the state of a system (expressed as a state space model), the type of attacker, objectives and payoffs of an attacker, probabilities associated with the state transition, etc. We modeled some of these uncertainties in [4].

Albeit its importance, research on uncertainty quantification and propagation for cyber security games are rudimentary at best. Therefore, this area will not only benefit DOE in developing resilient cyber systems, but will also benefit from the expertise of DOE community.

The facets we discussed so far help build situational awareness at a system level. In order to enable the application of these techniques, several other facets of the problem need to be built. Information for graph and game-theoretic methods for decision-making is built from systematic approaches for information collection and processing. The ability for anomaly detection, classification, traffic monitoring, and monitoring changes in system state are important. We present them next.

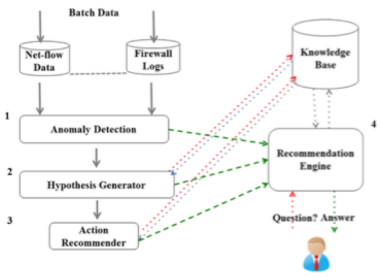

Figure2. Real-time anomaly detection

Real-time Anomaly Detection: The objective of an anomaly detection system is to generate hypotheses from observed anomalies; threat signals raised by third party software, or alerts from the attack graph system and provide high quality action recommendations to system administrators. An illustration is provided in Figure 2. We have shown that graph-based pattern matching and discovery can scale to near real-time on high-speed data streams, such as network traffic [5]. By combining these capabilities we can detect matches to more complex known patterns, learn new patterns, and detect anomalies to these patterns, all while operating on the high-speed data streams inherent to cyber-security tasks.

In conclusion, cyber security challenges posed by DOE's distributed resources are enormous and need a multi-faceted approach for effective solutions. Several aspects of the solution need large-scale data analytics tools to process massive amounts of data, especially data arriving at high speed and volume, data that is heterogeneous in nature, and data that needs complex analyses. At the extreme scale of computing, each of these facets is exacerbated. Extreme scale computing systems can be leveraged concurrently with advanced simulation and analysis to provide the analytics necessary to ensure resilient cyber systems and trustworthy results. We believe the framework we presented holds promise towards this end.

References:

[1] Xinming Ou, Wayne F. Boyer, and Miles A. McQueen. A scalable approach to attack graph generation. In *Proceedings of the 13th ACM conference on Computer and communications security* (CCS '06). ACM, New York, NY, USA, 336-345.
[2] Rinku Dewri, Nayot Poolsappasit, Indrajit Ray, and Darrell Whitley. 2007. Optimal security hardening using multi-objective optimization on attack tree models of networks. In Proceedings of the 14th ACM conference on Computer and communications security (CCS '07). ACM, New York, NY, USA, 204-213.
[3] Kong-wei Lye, Jeannette Wing. Game strategies in network security. *International Journal of Information Security*, V 4, N 1-2. P 71-86. Springer-Verlag. 2005.
[4] Chatterjee S, M Halappanavar, R Tipireddy, MR Oster, and S Saha. Quantifying Mixed Uncertainties in Cyber Attacker Payoffs. In *Proceedings of the IEEE Symposium on Technologies for Homeland Security*. 14-16 April, 2015. Waltham, MA.
[5] Choudhury S, L Holder, G Chin, K Agarwal, JT Feo. 2015. "A Selectivity based approach to Continuous Pattern Detection in Streaming Graphs", In Proceedings of the 18th International Conference on Extending Database Technology.