Research | July 21, 2016

Game Theory and Uncertainty Quantification for Cyber Defense Applications

By Samrat Chatterjee (https://sinews.siam.org/AbouttheAuthor/TabId/918/ArtMID/2225/ArticleID/761/Samrat-Chatterjee.aspx), Mahantesh Halappanavar (https://sinews.siam.org/AbouttheAuthor/TabId/918/ArtMID/2225/ArticleID/762/Mahantesh-Halappanavar.aspx), Ramakrishna Tipireddy (https://sinews.siam.org/AbouttheAuthor/TabId/918/ArtMID/2225/ArticleID/763/Ramakrishna-Tipireddy.aspx), and Matthew Oster (https://sinews.siam.org/AbouttheAuthor/TabId/918/ArtMID/2225/ArticleID/764/Matthew-Oster.aspx)

Cyber system defenders face the challenging task of continually protecting critical assets and information from a variety of malicious attackers. Defenders typically function within resource constraints, while attackers operate at relatively low costs. As a result, design and development of resilient cyber systems that support mission goals under attack, while accounting for the dynamics between attackers and defenders, is an important research problem. The goal of this article is to increase awareness among practitioners and researchers about uncertainty quantification within cybersecurity games, and encourage further advancements in this area.

In order to address cybersecurity challenges, researchers are increasingly adopting game theory-based mathematical modeling approaches that involve strategic decision makers within non-cooperative settings [5-6, 10]. Various taxonomies for classifying game-based modeling approaches exist (see Figure 1). These game formulations contain assumptions about rounds of game plays, past player actions, types of players, number of cyber system states, number of player actions in a given system state, and payoff (reward or penalty) functions associated with player actions.

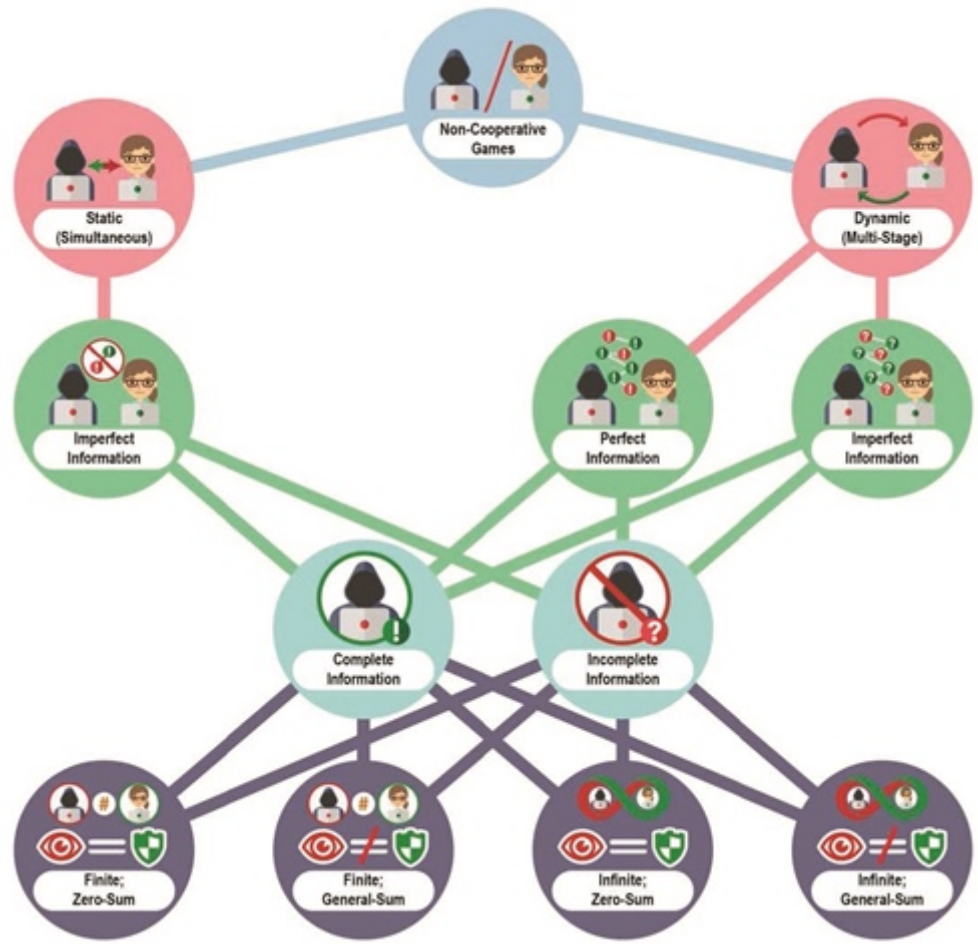

Figure 1. Types of non-cooperative game models for cybersecurity. Figure created by authors.

While game-based attack-defense models consider complex scenarios and effectively represent dynamic interactions, an increased focus on uncertainties in attacker payoff functions could enhance them. In a realistic setting, a defender cannot assume that all necessary information—both about the attackers and their own system—will be available. Since a cyber attacker's payoff generation mechanism is largely unknown, appropriate representation and uncertainty propagation is a critical task. One must also account for the lack or absence of perfect cyber system state information; such uncertainties may arise due to inherent randomness or incomplete knowledge of the behavior of or events affecting the system. For example, partial observability may make a cyber system's state uncertain over time. Moreover, multiple types of attackers could potentially target a system at a given point in time.

Advances in state-space modeling of cyber systems and reinforcement learning approaches for Markov decision processes have inspired the development of partially observable stochastic games (POSGs) and their potential applications for cybersecurity [1, 4, 6-9, 11]. A POSG is comprised of multiple players. Each player independently chooses actions, makes observations, and receives payoffs while the system state transitions based on player-action combinations. A POSG is defined as a tuple (N, A, S, O, P, R, s^0) where:

- N is the set of players
- $A := \Pi_{i \in N} A_i$ is the set of action tuples (pairs when $|N| = 2$), where A_i is the i^{th} player's action set
- S is the set of system states
- $O := \Pi_{i \in N} O_i$ is the set of observations, where O_i consists of the i^{th} player's observations
- P is the probability transition function, where $P(s'|s, a)$ denotes the probability of reaching state s' given a starting state of s and an action tuple a chosen by the players

- R is the reward function, where R_i denotes the individual reward function of the i^{th} player
- s^0 is the initial system state.

POSGs are very general formulations, and thus become intractable. Identifying joint policies (that map from observation history and system states to actions) of players forming a *Nash* equilibrium is the decision-making goal. Under equilibrium conditions, no player gains by unilaterally changing his/her policy. Typically, these problems may be categorized into the following two categories: (1) *Planning* – where complete specification of the cyber-system environment is known and optimal joint policies are desired; and (2) *Learning* – where players need to interact with the cyber-system environment to learn about the system and each other, while updating their policies based on these interactions. Solving such problems involves iteratively finding policies that achieve high rewards, on average, over the long run. A POSG's typical objective is to maximize the expected cumulative value (i.e. a function of payoffs) for each player [8]:

$$V_{p_1}(\pi) = \mathbf{E}\left[\sum_t R_{p_1}(s, \mathbf{a})|\pi, b^0\right],$$

where:

- $V_{p_1}(\pi)$ is the value function for the first player, i.e p_1, associated with a tuple of policies π
- $R_{p_1}(s, \mathbf{a})$ is the reward over time t for the first player in state s for a joint action $\mathbf{a}$
- b^0 is the initial system state distribution.

Researchers have proposed various approaches for solving POSGs, including dynamic programming with iterative elimination of weakly dominated strategies [1] and transformations of POSGs to a series of Bayesian games (with incomplete information about other player payoffs) that have properties similar to the original POSG [7].

In realistic cybersecurity settings, insufficient and uncertain information about system properties and attacker goals may be available to a defender. A recent approach proposed a probabilistic framework for quantifying attacker payoff uncertainty within a stochastic game setup that accounts for dependencies among a cyber system's state, attacker type, player actions, and state transitions [2-4]. This approach adopts conditional probabilistic reasoning to characterize dependencies among these modeling elements. The application of probabilistic theories (such as total probability theorem) and functions (such as marginal and conditional) may then lead to simulation of attacker payoff probability distributions under various system states and operational actions. The framework is flexible and accounts for multiple types of uncertainties—such as *aleatory* (statistical variability) and *epistemic* (insufficient information)—in attacker payoffs within an integrated probabilistic framework (see Figure 2).

Figure 2. Probabilistic attacker payoff framework. Figure created by authors.

Mathematically, as presented in [2-4], the discrete version of the marginal probability of attacker payoff utility (involving notions of time and cost), $Pr(u_{p_1})$, is:

$$Pr(u_{p_1}) = \sum_i \sum_j \sum_k \sum_l Pr(u_{p_1}|s'_l, a_k, \alpha_j, s_i) \cdot Pr(s'_l|a_k, \alpha_j, s_i) \cdot Pr(a_k|\alpha_j, s_i) \cdot Pr(\alpha_j|s_i) \cdot Pr(s_i)$$

where:

- $Pr(s_i)$ is the initial (prior) probability of system states s_i
- $Pr(\alpha_j|s_i)$ is the conditional probability of attacker type α_j for a given system state
- $Pr(a_k|\alpha_j, s_i)$ is the conditional probability of attacker and defender action combinations a_k for a given attacker type and initial system state
- $Pr(s'_l|a_k, \alpha_j, s_i)$ is the conditional probability of system state transition from s_i to s'_l for given action combinations, attacker type, and initial system state
- $Pr(u_{p_1}|s'_l, a_k, \alpha_j, s_i)$ is the conditional probability of attacker payoff utility.

Statistical probability distributions typically address aleatory uncertainty, while mathematical intervals address epistemic uncertainty. Depending on these representations, uncertainty propagation methods may include Monte Carlo sampling analysis, interval analysis, and/or probability bounds analysis. Application of uncertainty propagation techniques generates probability distributions, intervals, or intervals of distributions associated with attacker payoffs that serve as critical inputs within stochastic cybersecurity games. These probabilities may be informed and updated based on empirical event and system data, simulation experiments, and/or informed judgments of subject matter experts.

The game-theoretic and uncertainty quantification methods outlined above model the dynamics between cyber attackers and defenders, and have real-world potential to address proactive resource allocation challenges within resilient cyber systems. However, challenges to their implementation exist, including real-time, data-driven system state determination, "realistic" payoff uncertainty representations, and scalability of uncertainty propagation and stochastic game algorithms. Nevertheless, these approaches represent steps toward practical uses of game theory as an effective tool for rigorous cyber defense analysis.

Acknowledgments

This research was supported by the Asymmetric Resilient Cybersecurity (ARC) initiative at the Pacific Northwest National Laboratory (PNNL). PNNL is a multi-program national laboratory operated by Battelle Memorial Institute for the United States Department of Energy under DE-AC06-76RLO 1830.

References

[1] Bernstein, D.S., Hansen, E.A., Zilberstein, S., & Amato, C. (2004). Dynamic programming for partially observable stochastic games. *Proceedings of the 19th National Conference of Association for the Advancement of Artificial Intelligence (AAAI).* San Jose, CA.

[2] Chatterjee, S., Halappanavar, M., Tipireddy, R., Oster, M.R., & Saha, S. (2015). Quantifying mixed uncertainties in cyber attacker payoffs. *Proceedings of the 2015 IEEE International Symposium on Technologies for Homeland Security (IEEE-HST).* Waltham, MA.

[3] Chatterjee, S., Tipireddy, R., Oster, M.R., & Halappanavar, M. (2015). A probabilistic framework for quantifying mixed uncertainties in cyber attacker payoffs. *National Cybersecurity Institute Journal, 2*(3), 13-24.

[4] Chatterjee, S., Tipireddy, R., Oster, M., & Halappanavar, M. (2016). Propagating mixed uncertainties in cyber attacker payoffs: exploration of two-phase Monte Carlo sampling and probability bounds analysis. *Proceedings of the 2016 IEEE International Symposium on Technologies for Homeland Security (IEEE-HST).* Waltham, MA.

[5] Liang, X., & Xiao, Y. Game theory for network security. (2013). *IEEE Communications Surveys and Tutorials, 15*(1), 472-486.

[6] Lye, K., & Wing, J.M. (2005). Game strategies in network security. *International Journal of Information Security, 4*(1-2), 71-86.

[7] MacDermed, L., Isbell, C.L., & Weiss, L. (2011). *Markov games of incomplete information for multi-agent reinforcement learning*. Workshop paper from the 25th Association for the Advancement of Artificial Intelligence Conference (AAAI). San Francisco, CA.

[8] Oliehoek, F.A., Spaan, M.T.J., Robbel, P., & Messias, J.V. (2016). *The MADP toolbox 0.4*. (p. 37).

[9] Ramuhalli, P., Halappanavar, M., Coble, J., & Dixit, M. (2013). Towards a theory of autonomous reconstitution of compromised cyber-systems. *Proceedings of IEEE International Symposium on Technologies for Homeland Security (IEEE-HST)* (pp. 577-583). Waltham, MA.

[10] Roy, S., Ellis, C., Shiva, S., Dasgupta, D., Shandilya, V., & Wu, Q. (2010). A survey of game theory as applied to network security. *Proceedings of the 43rd Hawaii International Conference on System Sciences*. Honolulu, HI: IEEE Computer Society.

[11] Sutton, R.S., & Barto, A.G. (2012). *Reinforcement Learning: An Introduction* (2nd ed.) (p. 334). Cambridge, MA: MIT Press.

Samrat Chatterjee is a research scientist in applied statistics and computational modeling at the Pacific Northwest National Laboratory. Mahantesh Halappanavar is a staff scientist in the physical and computational sciences directorate at the Pacific Northwest National Laboratory. Ramakrishna Tipireddy is a postdoctoral researcher in the physical and computational sciences directorate at the Pacific Northwest National Laboratory. Matthew Oster is an operations research scientist with the national security directorate at the Pacific Northwest National Laboratory.