Game Theory in Distributed Systems Security: Foundations, Challenges, and Future Directions

Mustafa Abdallah | Indiana University-Purdue University Indianapolis
Saurabh Bagchi | Purdue University
Shaunak D. Bopardikar | Michigan State University
Kevin Chan | Army Research Lab
Xing Gao | University of Delaware
Murat Kantarcioglu | University of Texas at Dallas
Congmiao Li | University of California at Irvine
Peng Liu | Pennsylvania State University
Quanyan Zhu | New York University

Combining insights from distributed system security and game theory could effectively address security challenges. We present foundational concepts from both fields that can be integrated to better secure distributed systems and outline several research challenges for the community to tackle.

Many of our critical infrastructure systems and personal computing systems, which have a distributed structure, face increasing levels of attacks. There has been vast research on using both game theory and distributed system security to face these increasing attacks. Therefore, we feel it is time to bring in the rigorous reasoning from game theory advanced models to better secure such distributed systems. The distributed system security and the game theory technical communities can come together to effectively address this challenge of securing distributed systems. In this article, we lay out the foundations from each domain that we can build upon to achieve a successful integration of game theory and distributed system security for better securing large-scale distributed systems. We then describe a set of research challenges for the community, organized into three categories—analytical, systems, and integration challenges, each with "short-term" time horizon (two to three years) and "long-term" (five to 10 years) items.

Digital Object Identifier 10.1109/MSEC.2024.3407593

Introduction

Today's distributed systems face sophisticated attacks from external adversaries where the attacker aims to breach specific critical assets within these systems. Such attacks pose a serious danger to large-scale critical

Copublished by the IEEE Computer and Reliability Societies 1540-7993/24©2024IEEE

infrastructure, such as the massive supply chain attack on SolarWinds in 2020 and the Colonial Pipeline ransomware attack in 2021. Such attacks have motivated several attempts to improve the cybersecurity of these systems.[7]

In response to such attacks, there has been significant work in understanding vulnerabilities in large-scale distributed systems and putting together technological patches to address specific classes of vulnerabilities. However, the works often lack an understanding of the impact of cascading attacks or of mitigation techniques on the security of the overall system. Due to the large legacy nature of many distributed infrastructures and budgetary constraints, a complete rearchitecting and strengthening of the system is often impossible. Rather, rational decisions must be made to strengthen parts of the system, taking into account the risks and the interdependencies among the assets. In this context, significant research has investigated how to better secure these systems, with game-theoretical models receiving increasing attention. Such models have shown the power to capture the interactions of different players (strategic attackers and defenders) in different settings (see the survey[10])

While researchers have studied static game theory extensively for several decades, large-scale distributed systems present critical challenges that preclude the direct application of existing theory. Specifically, there is a need for new techniques to account for both the interdependencies and the dynamical nature of the subsystems. Furthermore, some of these dynamical subsystems may be complex in their own right (for example, a perception system that employs multimodal sensors) and may have the limitation of being represented only by simulation models. Thus, advanced game theory models can be proposed to better model attacker/defender realistic scenarios, where such modeling should be connected more to distributed system security to find new insights into securing distributed systems.

This problem context leads to four overarching questions that form a starting point for enhancing the usage of game theory for distributed system security.

1. Can the security community extend traditional game theory to develop tractable analysis and design techniques that can be applied to securing *large-scale* and *interdependent* distributed systems?
2. What are the main foundations in the game theory and distributed system security literature that can help us achieve such a goal of securing distributed systems?
3. What are the advantages and disadvantages of different game-theoretic models when applied to distributed systems security?
4. What are the main challenges and related research directions for integrating game theory for securing distributed systems?

In this article, we present a proposed vision for answering these questions. In particular, we organize our article as follows. We first lay out the foundations that the research community can build on when applying game theory concepts to enhance distributed system security. We then present the main challenges for such a synthesis, which we categorize into 1) analytical directions, 2) systems directions, and 3) integration directions. Figure 1 provides the main flow of this article.

Foundations: Build on Them

We have significant foundations on the topics of distributed systems security and game-theoretic security that we should build upon. Here, we survey the notable foundations categorizing them into two—game-theoretic security and distributed systems security.

Game-Theoretic Security

There have been notable successes in developing and applying game theory for the security of distributed systems.[1] This has been used in the context of proactive or reactive and fixed or adaptive schemes. The most commonplace game-theoretic model for security is that of two-player games, where a single attacker attempts to compromise a system controlled by a single defender. Game-theoretic models have been further used to study the interaction between (multiple) defenders and (multiple) attackers [for example, analyzing distributed denial of service attacks (DDoSs) and the security of cyberphysical systems (CPSs)]. The literature on game-theoretic models (and their unique differences) for different security scenarios can be categorized as follows.

Static and Complete Information Games. The static and complete information two-player games are benchmark security models that capture the incentives or objectives of the players as well as their constraints. The game assumes that the players have a common knowledge of the game and that it does not change over time. The Nash equilibrium of the game can be interpreted as the outcome of repeated plays between the two players or the consequence of homogeneous pairwise interactions of a large population. The analysis of this class of games provides a quantitative approach to assess risks and to design mitigating mechanisms. FlipIt games and Blotto games[11] are two notable games that have been widely used in understanding the competitive scenarios of resource takeover and subjugation in CPSs and military applications. The Nash equilibrium is the traditional concept of capturing efficient solution(s) of complete information games. For instance,

the Nash equilibrium has helped researchers understand botnet defenses.[4] In particular, this line of work has provided a comprehensive game-theoretical framework that models the interaction between the botnet herder and the defender group (network/computer users).

The Nash equilibrium showed the effectiveness of available defense strategies and control/strategy switching thresholds, specified as rates of infection. The two Nash equilibria obtained are either 1) the defender group defends at a maximum level while the botnet herder exerts an intermediate constant intensity attack effort or 2) the defender group applies an intermediate constant intensity defense effort while the botnet herder attacks at full power. This model also showed that integrating game-theoretical analysis with susceptible, infectious, or recovered epidemic models could be useful in understanding system behavior during botnet attacks. Overall, although complete information game-theoretic models for security games enable proactive security planning and predicting worst-case attack scenarios on these distributed systems, actual conflicts are dynamic and involve incomplete information for one or both players, which are discussed next.

Imperfect/Incomplete Information Games. These are games in which at least one player (defender or attacker) does not have complete information. This may be due to lacking complete knowledge of the system or to imprecise sensing. To analyze multistage multihost attacks that may be launched on networks, a defender needs to model long sequences of actions that can circumvent the system defenses.[3] These actions lead to policy spaces that grow exponentially with the number of attack stages, especially under partial/imperfect information. Monte Carlo sampling can confine the search to a decision tree of reduced size by guessing the other player's moves and then using a conventional min-max search to determine the best strategy.[12]

One promising line of work for games with imperfect/partial information is the use of deception (see a recent survey[20]) The key idea is for one or both players to synthesize new actions/policies that leverage limitations induced by the belief of the opponent. Notable classes of problems that fall within this class are *signaling games* that model information corruption, *Bayesian games* that model uncertainty in an opponent's type/cost, and *asymmetric constraints* that enforce stealth and partially observable stochastic games. Akin to general imperfect information games, the complexity of solving deception problems grows exponentially with the number of stages, beliefs, and actions.

Drawing inspiration from robust optimization, the application of randomized sampling methods has proven effective in computing policies that are robust security measures against adversaries employing randomized strategies.[5] These methods utilize randomized sampling techniques to explore the strategy space to choose effective strategies with high confidence.

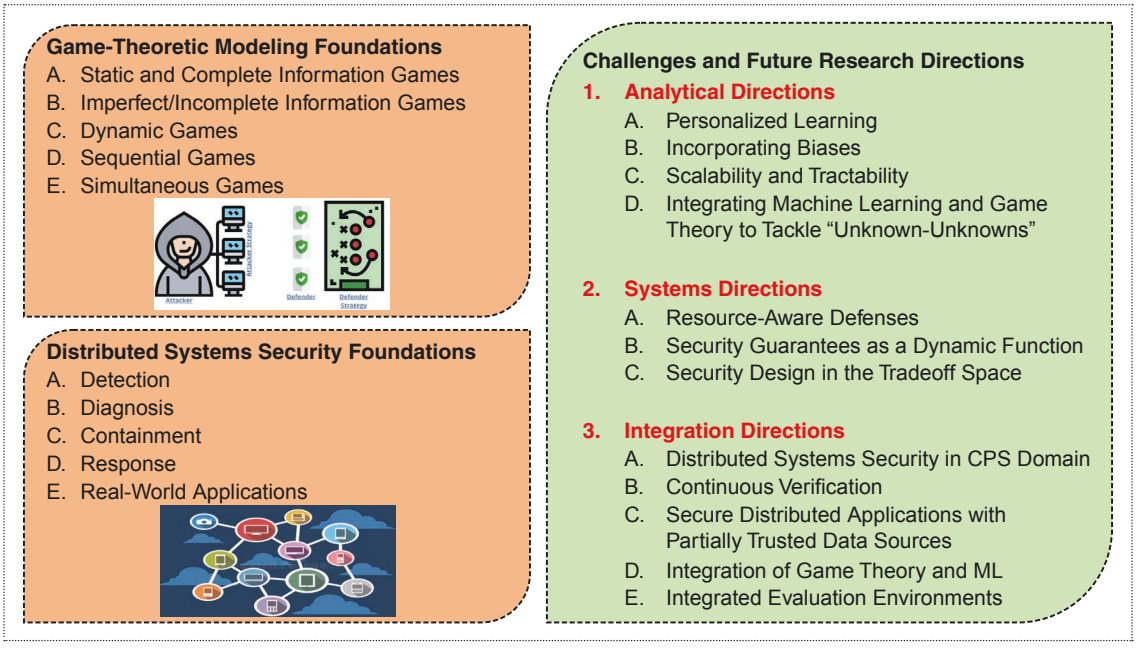

Figure 1. An overview of the flow of this article. We first show the main foundations for game theory modeling and distributed systems security. We then outline the research challenges and future directions that will need the integration of the advancements of the analytical side and systems side for securing distributed systems.

Overall, leveraging incomplete information game-theoretic models for distributed systems security captures the uncertainty about the adversary's actions and payoffs, along with the actions of other stakeholders, which can give a more accurate quantitative estimation of the security level of the distributed system. However, if the player's information evolves over time, then they are more effectively modeled as dynamic games.

Dynamic Games. These are games in which the information, the players' actions, or the payoffs vary over time. One promising line of work has been in leveraging *reinforcement learning (RL)*. Examples of such usage of RL are the malicious falsification of cost signals that are used to mislead agent policy.[6] Another example is applying RL and the infinite-horizon semi-Markov decision process to characterize a stochastic transition and the sojourn time of attackers in a honeynet. Another line of work is to model distributed systems using hybrid input–output automaton. This can help in characterizing the continuous time evolution of the security game.

In contrast to static game-theoretic models, these dynamic games capture the realistic evolution of vulnerabilities and adversary actions, which can lead to the effective usage of learning-based techniques for guiding human (or automated) decision making toward better security policies for securing current distributed systems that have such a dynamic nature. However, if there is a natural order in the conflicts that requires one player to play first or if the actions of both players are not visible to each other until a specified time, then such situations are more effective when modeled as sequential or simultaneous games, respectively, as discussed next.

Sequential Games. Game theory for security has been found to be tractable when considering sequential attacks, through *Stackelberg security games*. In these games, the defender moves first and allocates their resources to the assets under their ownership. Then, the attacker can observe the allocations made by the defender to each asset, after which the attacker targets part (or all) of the assets. Such games may incorporate real-time observations and consideration of nonmyopic players. In reality, many such games may be partially observable as the actions of a player may not be visible to other players (for example, an attacker may conceal their steps).

There have been several applications that have benefited from Stackelberg security games for distributed systems, as diverse as countering man-in-the-middle attacks and screening airport passengers throughout the United States.[15] The sequential order in these games also identifies realistic cases where adversaries attack distributed systems (in firms or government infrastructure) after the security decision-makers invest in securing these systems.

Simultaneous Games. A particular class of simultaneous-move games involving attackers and defenders (where the players have to choose their strategies simultaneously, without first observing what the other player has done) has been studied in various contexts. For example, the Colonel Blotto game is a useful framework to model the allocation of a given number of resources on different potential targets (for example, battlefields) between the attacker and the defender. Specifically, Schwartz et al.[13] proposed a solution for the heterogeneous Colonel Blotto game with asymmetric players (that is, with different resources) and with many battlefields that can have different values. While Colonel Blotto games typically involve deterministic success functions (where the player with the higher investment on a node wins that node), other work has studied cases where the win probability for each player is a probabilistic (and continuous) function of the investments by each player. Overall, simultaneous-move games arise in military-based distributed system security applications. Furthermore, simultaneous-move games may be a better way to model real-world situations in which attackers may choose to act without acquiring costly information about the defense security strategy, particularly if the security measures are difficult to observe (for example, undercover officers, strong privacy measures, and nonavailable insiders).

Advanced Games Examples for the Security of Distributed Systems. Game-theoretic models have also been used to study DDoS attacks, critical infrastructure security, censorship-resilient proxy distribution, wireless network security, and protecting computer networks from cascade attacks (see the survey[10]) Further, Abdallah et al.[2] studied mechanism design to incentivize defenders toward beneficial security investments in distributed systems.

Summary of Game Theory Literature on Distributed Security. Figure 2 provides an overview of the literature on game-theoretic models for distributed systems security. We highlight the advantages and disadvantages of different game-theoretic models when applied to distributed systems security and discuss the potential applications of each model in the various research directions outlined in our vision.

Distributed Systems Security

One way to organize the foundations that have been developed here is through each step of the workflow for distributed systems security, namely, detection, diagnosis, and containment and response.

Detection. This is a mature area of work in which there is influential work on collaborative intrusion detection using multiple sensors placed in a distributed system. This line of work has contributed algorithms to determine where to place the sensors and how to integrate outputs from multiple sensors to devise an integrated decision on the detection of an attack. A survey work on this topic is by Vasilomanolakis et al.[16] This area saw some of the early applications of machine learning (ML) to security. In the context of game theory, game-theoretic analysis has helped develop various intrusion detection systems for distributed systems to increase detection accuracy with reduced cost. The game-theoretical approach has also been used for mitigating edge DDoS attacks.[10]

Diagnosis. This has contributed algorithms to identify the root cause of the attack. This was initially substantially rule based, of the form if metric A > threshold τ_1 and B < threshold τ_2, then A is the root cause.[9] Later, foundational work was done on this topic by using ML, such as causal theory.[17] One significant challenge that has been successfully addressed is how to maintain effective diagnosis capabilities in security algorithms when the interactions and connections between elements within a distributed system are constantly changing.

Containment and Response. This concept has had notable success in the topic of moving target defense, which seeks to change some parameters of the defended system, such as using IP addresses to thwart an adversary. This can be done proactively as a preventive measure in response to a detected threat.[14]

The integration of game analysis for critical infrastructure protection has proven highly successful, effectively encompassing containment, response, and recovery measures.

Applications: CPS and Critical Infrastructure

Security games have played a crucial role in addressing the resilience and interdependence of critical infrastructures, including our nation's legacy CPSs, such as power grids,[2] transportation,[3] and manufacturing systems. With the increasing connectivity of these systems, they face a larger attack surface. The application of game-theoretic methods is vital in developing strategic mechanisms for detection, diagnosis, containment, and response, ensuring their resilience. Having outlined the foundational aspects of game-theoretic models and distributed systems security, we now turn our attention to the

Game Theory for Distributed Security Literature	Game Model	Pros (+)	Cons (-)	Directions
(a) Interdependent Security Games [Ref. no. 37 in Laszka et. al. 15] (b) Insurance-based Games [Ref. no. 37 in Abdallah et. al. 2022] (c) Information Security Games [Ref. no. 25 in Laszka et. al. 2015] (d) Blotto games for CPS [Robinson 2013, Laszka et. al. 2015] (e) Botnet Defenses using Game Theory [Bensoussan et. al. 2013] (f) Mechanism-based Security Games [Abdallah et. al. 2022]	Static and Complete Information Games	- Equilibrium rigor analysis of costs of security risks - Mechanism design for optimal defense method - Exploration of effective defense/control strategy	- Lack of dynamic interactions - Limited learning by players and lack of adaptation - Non-unique equilibrium - Ignoring private information - Limited real-world scenarios	- Personalized learning - Evaluation environments - Security of distributed CPS
(a) Incomplete Information Security Games [Alpcan & Basar 2009] (b) Large Incomplete-information Games [Sandholom 2015] (c) Deception in Incomplete Inform. Games [Zhu et. al. 2021] (d) Random Sampling in Zero-sum Games [Shaunak et. al. 2013] (e) Probabilistic Security Games [Chapter 5 in Alpcan and Basar 2010]	Imperfect/Incomplete Information Games	- Bayesian equilibrium analysis of all defenses - Real-world scenarios for different attack strategies - Diverse sampling method for exploring strategies	- Higher solution complexity - Lack of proper coordination among different defenders - No guarantee of equilibrium - High uncertainty levels - Limited learning scenarios	- Incorporating attack biases - Scalability and tractability - Integrating ML and game-theory
(a) Dynamic Security Games [Chapter 3 in Alpcan and Basar 2010] (b) Selfish investments in dynamic network security Games [Ref. no. 11 in Laszka et. al. 2015] (c) Reinforcement learning in Security Games [Yunhan et. al. 2019] (d) Moving Target Defense [Ref. no. 61 in Sengupta et. al. 2020] (e) Learning Attacks in Distributed Systems [Reference no. 25 in Abdallah et. al. 2022]	Dynamic Games	- Varying strategies and players payoffs over time - High prospect of guiding security decision-makers - Usage of diverse set of learning methods(RL, ML) - Risk analysis simulations	- Computational complexity - Changing equilibrium(s) - Accurate modeling of the evolution of attack actions - Quantifying learning levels - Limited-realistic application - Real security losses of trials	- Personalized learning - Integrating ML and game-theory - RL defenses - Partially trusted data sources
(a) Stackelberg Security Games [Sinha et. al. 2018] (b) Sequential Games for Cyber-physical Systems Security [Ref. no. 170 in Humayed et. al. 2017] (c) Sequential one-defender-n-attacker Games [Ref. no. 66 in Abdallah et. al. 2022] (d) Sequential Security Games [Chapter 2 in Alpcan & Basar 2010]	Sequential Games	- Capturing observations - Non-myopic players - An accurate modeling of maximum attack gains - Realistic sequential order in security applications	- Concealing actions problem - Information assumptions - Lack of assumed strategic commitment by the follower - Stackelberg Equilibria issues - Complex analytical solutions	- Security of distributed CPS - Scalability and tractability
(a) Heterogeneous Colonel Blotto Game for Network Security [Schwartz et. al. 2014] (b) System Reliability and Free Riding [Last Ref. in Laszka et. al. 15] (c) Behavioral Interdep. Security Games [Abdallah et. al. 2020]	Simultaneous Games	- Military-based security - Asymmetric resources - No need for observing the actions (lower costs)	- Mapping of all possible equilibria to real scenarios - Lack of defender's control - Possible colluding attackers	- Incorporating defense biases - Resource aware defenses

Figure 2. A summary of the relevant literature on game-theoretic models for distributed systems security. We show the pros and cons of various game-theoretic models as applied to distributed systems security and the prospective usage of each model in the different research directions outlined in our vision.

key challenges faced by the research community and provide prospective research directions.

Challenges Ahead

Here, we summarize the technical challenges to improve the security landscape of distributed systems. We structure our discussion into challenges that are on the *analytical directions* and *systems directions* and those that involve a combination of the two, called *integration directions*. The orthogonal dimension on which we structure these challenges is the time horizon to solve them, with the short term indicating two to three years and the long term indicating five to 10 years. Figure 3 summarizes the main challenges and future research directions for such integration.

Analytical Directions

Personalized Learning (Short Time Horizon). Different actors (let's say different defenders and adversaries) learn differently (as in stochastic learning) and at different rates. The learning happens for human actors as well as for machines (in an ML context). This learning can build on the literature on dynamic games, discussed in

Figure 3. A timeline overview of research challenges and future research directions for both the analytical side and systems side (upper part) along with possible research directions for integrating both sides (lower part). XAI: Explainable AI.

the "Dynamic Games" section. Another form of heterogeneity that comes in is asymmetric capabilities among the various players; for example, some defenders have access to assets with a trusted hardware environment, like ARM's TrustZone or Intel's SGX. The personalization of the learning must also be able to accommodate partial cooperation (among defenders) or partial collusion (among adversaries) (building on literature on incomplete information games, discussed in the "Imperfect/Incomplete Information Games" section.) in addition to the formulation of complete cooperation or collusion (building on the "Static and Complete Information Games" section).

Incorporating Biases and Incomplete Information (Short Time Horizon). For the learning, one has to incorporate incomplete information sharing among the actors (building on prior literature on incomplete information games, discussed in the "Imperfect/Incomplete Information Games" section.). For human learning, one must also incorporate cognitive biases among the human players. Behavioral economics has shown that humans consistently deviate from these classical models of decision making; for example, humans perceive gains, losses, and probabilities in a skewed and nonlinear manner.[8] We have done some nascent work applying behavioral game theory to the security of distributed interdependent systems.[2] The main point here is to explore the effects of such behavioral biases on the security policies of human decision-makers and their effect on securing distributed systems. One example in the related literature is that behavioral human security decision-makers may allocate part of their limited security resources to noncritical parts of the distributed system. The key insight is that one can provide the appropriate incentives to reduce the biases and encourage cooperation among even biased defenders. Such bias incorporation would build on prior simultaneous security games for distributed systems security discussed in the "Distributed Systems Security" section. Such cooperation is in general a more secure strategy than independent decision making. A related theme here is trust building among human agents in multiagent learning. This is to counter the natural tendency for each player to explore and exploit their own strategy spaces (for example, see (a) in "Dynamic Games" in Figure 2).

Scalability and Tractability (Long Time Horizon). A well-known challenge with applying game-theoretic formulation to the security of distributed systems is the scalability and the tractability of the solution (particularly for imperfect information games and sequential games discussed in the "Imperfect/Incomplete Information Games" section.). Scalability implies scaling to large numbers of actors, large amounts of data, or large volumes of interaction among the actors. Tractability implies being able to handle realistic attack models or realistic workload incidents on the protected system. To ease this challenge, we must develop sound approximations of the game-theoretic formulation, for example, leveraging sampling techniques discussed in the "Imperfect/Incomplete Information Games" section. This should allow one to produce bounds for best-case/worst-case outcomes. As an example, one can use scalable techniques from epidemic theory to analyze the effect of cascading attacks while accommodating the case of large numbers of players.

Integrating ML and Game Theory to Tackle "Unknown-Unknowns" (Long Time Horizon). The game-theoretic formulations are often rigidly deterministic in nature; for example, a specific deterministic action is coded in for a particular state. The open question is: Can ML be integrated with game theory and thus incorporate stochastic behavior? The best candidate game for that direction is dynamic games, discussed in the "Dynamic Games" section. This is important as failures and attacks are inherently stochastic in nature. The core challenge here is that ML methods can achieve accurate predictions only if they have been trained with the right set of examples. Security problems such as zero-day attacks remain extremely challenging because of the lack of the appropriate types and numbers of examples until recently. The key question to ask here is: How can a defense scheme be resilient to unanticipated attacks (also known as *black-swan events*)?

Systems Directions

Resource-Aware Defenses (Short Time Horizon). Different nodes have different capabilities and available resources, and the system should be able to calibrate the defense mechanism using node-specific attributes. Some of these node-specific attributes will be static and unchanging, such as the intrinsic hardware capability of the node (which can be captured efficiently using static and complete information games discussed in the "Static and Complete Information Games" section.). In contrast, some attributes will be dynamic, such as the current battery level on the node (which is better captured using dynamic games, discussed in the "Dynamic Games" section.). The cost of an attack may also vary; for example, the cost to corrupt data may be higher if there is some security protection overlaid on the data.

Compared with traditional defense mechanisms, which could be slow due to the lack of awareness of the available resources and capabilities for different nodes, game-theoretic approaches can better allocate the

limited resources of each node to balance the defense tasks and take timely action. This is particularly important for critical infrastructure security.

Security Guarantees as a Dynamic Function (Short Time Horizon). The security guarantees may, under certain situations, be a function of the current system state. As an example, under this regime, the guarantees could be a function of the number and the capability of attackers and defenders rather than an absolute. Thus, the security guarantees that the system can provide are a dynamic property varying with the system state. For example, hardware degrades, and the software ecosystem changes over time. The guarantees could also be a function of the level of collusion among attackers, for example, non-Byzantine or Byzantine attackers (which can be modeled by imperfect information games discussed in the "Imperfect/Incomplete Information Games" section.).

Designing for Security in the Tradeoff Space (Long Time Horizon). A radical design principle would be to design for security in the tradeoff space between security on the one hand and performance and resource usage on the other. For example, the security design may use hardware-level virtualization, when available, rather than (software) containers, the former providing stronger isolation and greater protection against side-channel attacks. This direction can build on the literature on containment and response, discussed in the "Containment and Response" section. Suppose one can design specialized functions, for example, specialized to the resource available at a node. In that case, this has the added benefit of reducing the attack surface, making debugging easier and reducing consumed resources.

The security guarantees must be clearly delineated as a function of the performance and the resource usage so that the end user can understand the guarantees they are getting. Alternately, in the case of automatic composition with other software packages, it becomes clear what security guarantees are in effect in the composed system.

Integration Directions

Security of Distributed Systems in CPS Domain (Short Time Horizon). To secure CPSs, there are several unique aspects that we need to consider. These are prototypical interdependent distributed systems, often with multiple stakeholders as owners. The nodes are embedded in the physical environment and are subject to environmental effects, which contribute to the difficulty of securing them. For example, it is often difficult to tell apart a node malfunction due to the environmental effects of a node compromise. Further, some parts of the system are opaque to defenders as they are developed by an external party. Consequently, security mechanisms that rely on the (fine-grained) observability of the events happening in the software stack are out of bounds. In this context, sequential games and complete information games have a strong foundation to build on for these challenges (see related works in Figure 2).

Continuous Verification (Short Time Horizon). This topic needs to answer the question: Are our models and practical software instantiations generating valid results even under attacks and perturbations? This should happen continually rather than in batch mode as is typically done today, when verification is used at all. The continuous verification should happen as the system processes inputs and generates outputs during its operation. This could use *sparse* human feedback online, that is, without putting undue cognitive burden on the user. Existing methods for incremental verification/testing would be useful for this challenge,[19] as would be recent progress on the verification of highly nonlinear ML models.[19] This continuous verification would be based on dynamic games that can efficiently model this progressive verification setup.

Secure Distributed Applications With Partially Trusted Data Sources (Long Time Horizon). The overarching question that we need to answer is whether we can build secure distributed applications when the input data are only partially trusted. This is particularly important for the significant class of systems that are stochastic and data dependent in nature. The data may be streaming, rather than at rest, adding to the challenge of verifying the data. The nodes comprising the distributed system are heterogeneous (as argued in the "Simultaneous Games" section) in terms of resources, including secure hardware and access to trustworthy data sources. Finally, trust in data is a dynamic property, increasing, for example, when there has been successful validation of data and decreasing when there is a detected attack. This challenge can benefit from work on causal reasoning in dynamic systems, as outlined in the "Dynamic Games" section.

Integration of Game Theory and ML (Long Time Horizon). The grand open question here is: Can we, in a principled manner, integrate game theory and ML to secure distributed systems? In such integration, we must be cognizant that there can be multiple players (tens to hundreds) in terms of attackers and defenders. The interactions and actions may evolve over time,[20] necessitating learning rather than static spaces for actions and rewards, as is typical today. Dynamic games are the best candidates to tackle such integration challenges.

There may also be partial information sharing among defenders, asymmetric information between attackers and defenders (captured by imperfect information games discussed in the "Imperfect/Incomplete Information Games" section), and cognitive biases among players. A subset of these factors may be relevant in a specific application context, but the framework and the algorithms should be able to encompass them. One direction could be to leverage ML models to learn the approximate utility functions of the participants by analyzing the past behavior data. ML models can also be used to learn approximate best response strategies when solving the game-theoretical model is intractable (for example, Wu et al.[18])

Using Game Theory to Gain a Better Understanding of RL Defenses (Long Time Horizon). RL techniques are being increasingly used to respond to continuous probes (for example, heartbeat requests) to exploit some vulnerability, like the HeartBleed OpenSSL vulnerability and lateral movement attempts. Hence, the following research questions become important: Do the "game plays" between the attacker and the corresponding RL agent converge to a certain notion of equilibrium? And if so, how soon? A clear answer to these questions, which will build on dynamic games from game theory foundations and detection from distributed systems security literature, helps system defenders discover better RL-based defenses against attacks.

Integrated Evaluation Environments (Long Time Horizon). The availability of an integrated evaluation environment (synonymous with a testbed) would be important for the community to evaluate if we are making progress toward the hard security goals. These testbeds across application domains will have two parts; one will provide application-generic functionality, and the other will be specific to the application domain. The desiderata for testbeds will be that they will allow for the injection of various types of attacks, the creation of different kinds of players (with the heterogeneous characteristics mentioned earlier in the "Analytical Directions" section), and the measurement of a diversity of metrics of interest. Such testbeds would be important for educating policymakers in addition to their value in demonstrating research artifacts. The acquired education may help policymakers inform what kinds and degrees of security investments should be made to different parts of an interconnected system to reach a desired level of security. This will make our critical infrastructure more secure.

In summary, our suggested research directions for each side (game theory side and distributed systems security side), along with the integration of both sides, can help in the better development of tractable and practical methods for securing distributed systems, which would represent a consequential advancement for our technical community and our society broadly.

Many of the critical infrastructure systems that we rely on as well as personal computing systems are structured as distributed systems. Take, for example, from the large end of the spectrum—transportation, power grid, and financial infrastructures—to the personal end of the spectrum—cooperating personal computing and Internet of Things devices. As the attack surfaces for such systems become larger and the sophistication and the incentives for attacks against them increase, it is time to usher in rigorous reasoning to secure such systems. At the same time, the rigorous analytical foundations need to be made scalable and tractable to apply to these real-world large-scale applications under realistic resource and timing constraints.

In this context, the technical themes of the security of distributed systems and game theory applied to security have much to contribute to each other. We trust that the two vibrant communities will continue the process of working together for better securing distributed systems. In this article, we have laid out a set of foundations that will serve as useful starting points for our journey plus a set of open research challenges that the community will hopefully take up.

The challenges and research directions outlined in this vision article will help toward the ultimate goal of practical and secure distributed systems. Our broad vision is twofold and applies to distributed systems both at work and in our personal lives. First, we will have a clear understanding of their security properties so that we can decide the level of trust that is warranted in each. Second, such reasoning can be done in a systematic manner, without having to perform one-off reasoning for each system. ■

Acknowledgment

The idea for this article was conceived during the NSF SaTC PI Meeting in Arlington, VA, USA, in June 2022. We thank the attendees of the "Game Theory and Distributed Systems Security" breakout for their valuable feedback. We thank the NSF SaTC meeting organizers, William Enck, Heather Lipford, and Michael Reiter, and the NSF SaTC program managers, Jeremy Epstein, Nina Amla, and Daniela Oliveira, for enabling this session. This work was supported by the Army Research Lab (ARL) under Contract W911NF-2020-221 and the National Science Foundation under Grants CNS-202667, CNS-2134076, and CNS-2038986. Any opinions, findings, and

conclusions expressed in this material are those of the authors and do not necessarily reflect the views of the sponsors. Apart from the first two authors, all other authors are listed alphabetically. The corresponding author is Saurabh Bagchi.

References

1. M. Abdallah, P. Naghizadeh, A. R. Hota, T. Cason, S. Bagchi, and S. Sundaram, "Behavioral and game-theoretic security investments in interdependent systems modeled by attack graphs," *IEEE Trans. Control Netw. Syst.*, vol. 7, no. 4, pp. 1585–1596, Dec. 2020, doi: 10.1109/TCNS.2020.2988007.
2. M. Abdallah et al., "TASHAROK: Using mechanism design for enhancing security resource allocation in interdependent systems," in *Proc. IEEE Symp. Secur. Privacy (SP)*, Piscataway, NJ, USA: IEEE Press, 2022, pp. 249–266, doi: 10.1109/SP46214.2022.9833591.
3. T. Alpcan and T. Başar, *Network Security: A Decision and Game-Theoretic Approach*. Cambridge, U.K.: Cambridge Univ. Press, 2010.
4. A. Bensoussan, M. Kantarcioglu, and C. Hoe, "A game-theoretical approach for finding optimal strategies in a botnet defense model," in *Decision and Game Theory for Security*, T. Alpcan, L. Buttyán, and J. S. Baras, Eds., Berlin, Germany: Springer-Verlag, 2010, pp. 135–148.
5. S. D. Bopardikar, A. Borri, J. P. Hespanha, M. Prandini, and M. D. Di Benedetto, "Randomized sampling for large zero-sum games," *Automatica*, vol. 49, no. 5, pp. 1184–1194, 2013, doi: 10.1016/j.automatica.2013.01.062.
6. Y. Huang and Q. Zhu, "Deceptive reinforcement learning under adversarial manipulations on cost signals," in *Proc. Int. Conf. Decis. Game Theory Secur.*, Cham, Switzerland: Springer-Verlag, 2019, pp. 217–237.
7. A. Humayed, J. Lin, F. Li, and B. Luo, "Cyber-physical systems security—A survey," *IEEE Internet Things J.*, vol. 4, no. 6, pp. 1802–1831, Dec. 2017, doi: 10.1109/JIOT.2017.2703172.
8. D. Kahneman and A. Tversky, "Prospect theory: An analysis of decision under risk," *Econometrica, J. Econometric Soc.*, vol. 47, no. 2, pp. 263–291, 1979, doi: 10.2307/1914185.
9. G. Khanna, M. Y. Cheng, P. Varadharajan, S. Bagchi, M. P. Correia, and P. J. Veríssimo, "Automated rule-based diagnosis through a distributed monitor system," *IEEE Trans. Dependable Secure Comput.*, vol. 4, no. 4, pp. 266–279, Oct./Dec. 2007, doi: 10.1109/TDSC.2007.70211.
10. A. Laszka, M. Felegyhazi, and L. Buttyan, "A survey of interdependent information security games," *ACM Comput. Surv.*, vol. 47, no. 2, 2015, Art. no. 23, doi: 10.1145/2635673.
11. B. Roberson, "The Colonel Blotto game," *Econ. Theory*, vol. 29, no. 1, pp. 1–24, 2006, doi: 10.1007/s00199-005-0071-5.
12. T. Sandholm, "Abstraction for solving large incomplete-information games," in *Proc. 29th AAAI Conf. Artif. Intell.*, 2015, pp. 4127–4131 doi: 10.1609/aaai.v29i1.9757.
13. G. Schwartz, P. Loiseau, and S. S. Sastry, "The heterogeneous Colonel Blotto game," in *Proc. 7th Int. Conf. NETwork Games, COntrol OPtim. (NetGCoop)*, Trento, Italy, Oct. 2014, pp. 232–238.
14. S. Sengupta, A. Chowdhary, A. Sabur, A. Alshamrani, D. Huang, and S. Kambhampati, "A survey of moving target defenses for network security," *IEEE Commun. Surv. Tut.*, vol. 22, no. 3, pp. 1909–1941, Thirdquarter 2020, doi: 10.1109/COMST.2020.2982955.
15. A. Sinha, F. Fang, B. An, C. Kiekintveld, and M. Tambe, "Stackelberg security games: Looking beyond a decade of success," in *Proc. 27th Int. Joint Conf. Artif. Intell. (IJCAI)*, 2018, pp. 5494–5501.
16. E. Vasilomanolakis, S. Karuppayah, M. Mühlhäuser, and M. Fischer, "Taxonomy and survey of collaborative intrusion detection," *ACM Comput. Surv. (CSUR)*, vol. 47, no. 4, pp. 1–33, 2015, doi: 10.1145/2716260.
17. H. Wang, G. Yang, P. Chinprutthiwong, L. Xu, Y. Zhang, and G. Gu, "Towards fine-grained network security forensics and diagnosis in the SDN era," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2018, pp. 3–16, doi: 10.1145/3243734.3243749.
18. J. Wu, C. A. Kamhoua, M. Kantarcioglu, and Y. Vorobeychik, "Learning generative deception strategies in combinatorial masking games," in *Proc. Int. Conf. Decis. Game Theory Secur.*, Cham, Switzerland: Springer-Verlag, 2021, pp. 98–117.
19. J. Yao, R. Tao, R. Gu, J. Nieh, S. Jana, and G. Ryan, "{DistAI}:{Data-Driven} automated invariant learning for distributed protocols," in *Proc. 15th USENIX Symp. Operating Syst. Design Implementation (OSDI)*, 2021, pp. 405–421.
20. M. Zhu, A. H. Anwar, Z. Wan, J.-H. Cho, C. A. Kamhoua, and M. P. Singh, "A survey of defensive deception: Approaches using game theory and machine learning," *IEEE Commun. Surv. Tut.*, vol. 23, no. 4, pp. 2460–2493, Fourthquarter 2021, doi: 10.1109/COMST.2021.3102874.

Mustafa Abdallah is an assistant professor in the Purdue School of Engineering and Technology at Indiana University-Purdue University Indianapolis, West Lafayette IN 47907 USA. His research interests include game theory, human decision making, and machine learning with applications including cybersecurity, edge computing, and data science. Abdallah received a Ph.D. from the Elmore Family School of Electrical and Computer Engineering at Purdue University. His research contribution is recognized by receiving the Purdue Bilsland Dissertation

Fellowship and having many publications in top IEEE/ACM journals and conferences. He also was the recipient of an M.Sc. Fellowship from Cairo University in 2013. He is a Member of IEEE. Contact him at mabdall@iu.edu.

Saurabh Bagchi is a professor of electrical and computer engineering and computer science at Purdue University, West Lafayette IN 47907 USA. His research interests include dependable computing and distributed systems. Bagchi received a Ph.D. in computer engineering from the University of Urbana-Champaign. He is the founding director of a university-wide resilience center at Purdue called CRISP (from 2017 to present) and principal investigator of the U.S. Army's Artificial Intelligence Innovation Institute (A2I2) (from 2020 to 2025). He serves as the founder and CTO of a cloud computing startup, KeyByte. He is a Senior Member of IEEE. Contact him at sbagchi@purdue.edu.

Shaunak D. Bopardikar is an assistant professor in the Electrical and Computer Engineering (ECE) Department, Michigan State University (MSU), East Lansing, MI 48825 USA. Bopardikar received a Ph.D. in mechanical engineering from the University of California at Santa Barbara, CA, USA. His research interests include systems and control, game theory, and optimization. He has more than 75 refereed journal and conference publications and holds two U.S. patented inventions. His recent recognitions include a National Science Foundation Career Award and an MSU College of Engineering's Withrow Teaching Excellence Award (ECE). He is a Senior Member of IEEE. Contact him at shaunak@egr.msu.edu.

Kevin Chan is an electronics engineer at the U.S. Army Combat Capabilities Development Command, Army Research Laboratory, Adelphi, MD 20783 USA. His research interests include network science, distributed analytics, and cybersecurity. Chan received a Ph.D. in electrical and computer engineering from the Georgia Institute of Technology. He is the recipient of the 2021 IEEE Communications Society Leonard G. Abraham Prize and multiple best paper awards. He has served as a coeditor of the *IEEE Communications Magazine* Military Communications and Networks Series. He is a Senior Member of IEEE Contact him at kevin.s.chan.civ@army.mil.

Xing Gao is an assistant professor in the Department of Computer and Information Sciences at the University of Delaware, Newark 19716-3106 DE, USA. His research interests include security, cloud computing, and mobile computing. Gao received a Ph.D. degree in computer science from the College of William and Mary, Williamsburg, VA, USA. He is a Member of IEEE. Contact him at xgao@udel.edu.

Murat Kantarcioglu is an Ashbel Smith Professor in the Computer Science Department and the director of the Data Security and Privacy Lab at The University of Texas at Dallas, Dallas 75080 TX USA. He is also a faculty associate at the Harvard Data Privacy Lab and a visiting scholar at the University of California Berkeley RISE Labs. His research interests include the integration of cybersecurity, data science, and blockchains for creating technologies that can efficiently and securely share and analyze data. Kantarcioglu received a Ph.D. in computer science from Purdue University, where he received the Purdue CERIAS Diamond Award for Academic Excellence. He is also a Fellow of IEEE and the American Association for the Advancement of Science (AAAS) and a distinguished member of ACM. Contact him at muratk@utdallas.edu.

Congmiao Li is a postdoc at the University of California, Irvine (UCI), Irvine, CA 92697 USA. Her research interests include security and computer architecture. Li received a Ph.D. degree in computer engineering from UCI. He is a Member of IEEE. Contact her at congmial@uci.edu.

Peng Liu is the Raymond G. Tronzo, MD Professor of Cybersecurity, founding director of the Center for Cyber-Security, Information Privacy, and Trust, and founding director of the Cyber Security Lab at Penn State University, University Park, PA 16802 USA. His research interests include all areas of computer security. Liu received a Ph.D. from George Mason University. He has published more than 350 technical articles, including numerous papers and articles at top conferences and journals. His research has been sponsored by the NSF, ARO, AFOSR, DARPA, DHS, DOE, AFRL, NSA, TTC, CISCO, and HP. He is the co-editor-in-chief of the *Journal of Computer Security* and was an associate editor for *IEEE Transactions on Dependable and Secure Computing*. He is a recipient of the DOE Early Career Principal Investigator Award. He has advised or coadvised more than 40 Ph.D. dissertations to completion. He is a Member of IEEE. Contact him at pxl20@psu.edu.

Quanyan Zhu is an associate professor in the Department of Electrical and Computer Engineering, New York University (NYU), New York, NY 11201 USA. He is an affiliated faculty member of the Center for

Urban Science and Progress and the Center for Cyber Security at NYU. His research interests include cognitive security, artificial intelligence and game theory, cyber and physical system resilience, automation, and robotics. Zhu received a doctorate in computer engineering from the University of Illinois at Urbana-Champaign. He currently serves as the technical committee chair on security and privacy for the IEEE Control Systems Society. He is a Senior Member of IEEE. Contact him at qz494@nyu.edu.