

Research Statement

Hui Lin

My research spans multiple disciplines in computer science/engineering and electrical engineering, including cyber security, intrusion detection systems (IDS), cyber-physical systems (CPS), the Internet of things (IoT), software-defined networking (SDN), operating systems, virtual machines, power system analysis, and control theory. My primary interests are in *cyber security, especially in the context of CPSes and IoTs*.

In the early stage of my research, I gained expertise in the security of general computing environments, with the design of computer architecture, virtual machines, and operating systems [2][3]. Then I developed a keen interest in cyber-physical systems as attackers can introduce physical damage with a short latency in those environments. This latency can vary from hundreds of milliseconds in power systems to a few milliseconds in robotic surgical systems [1]. The diverse computing environments of CPSes make general security approaches insufficient. Consequently, I developed unprecedented methods that distinguish my research from others':

- **Cyber-physical interdisciplinary design.** I combined the knowledge of both cyber and physical domains to detect attacks that introduce little anomalies in either domain of a CPS, such as a power system. I have applied this concept to solve different problems: (i) detect attacks that introduce physical damage by combining network IDS with a newly designed power flow analysis algorithm; (ii) reconnect isolated power grid devices (due to attacks) to maintain the availability of grid operations by a self-healing network; and (iii) preemptively disrupt and mislead attackers by obfuscating network configurations and physical measurements.
- **Real data to drive research.** I analyzed real data from CPSes, including logs and alerts from campus computing and network infrastructures and operational measurements from real power systems at UIUC. The data, which were used to construct experiment benchmarks, help to reshape my research into practical approaches that can be applied in real CPS environments.
- **Evaluations by both theoretical analysis and real-world experiments.** For evaluations in physical domains, I used theoretical analysis, e.g., power system analysis and control theory. For evaluations in cyber domains, I constructed large-scale wide area networks consisting of real hardware switches at different geographical sites in Geni network experiment platform; I integrate power grid simulations with the networks to build a CPS testbed.

My research is conducted under the TCIPG (Trustworthy Cyber Infrastructure for the Power Grid) center, a major multi-university collaborative effort between academia and industry. In this project, I am actively working in collaborations with researchers in diverse areas, including Dr. Robin Sommer (International Computer Science Institute, UC Berkeley), Dr. Jianhui Wang (Argonne National Lab), Prof. Kevin Jin (Illinois Institute of Technology), Prof. Pete Sauer (UIUC), and Prof. Rui Tan (Nanyang Technological University, Singapore). This collaboration helps me to build a solid interdisciplinary background; I will continue the collaborations to extend my research to broad CPSes and IoT environments in future work.

CURRENT RESEARCH

My dissertation focuses on *improving the security of the cyber-physical infrastructure of power grids against remote insider attacks*. Power systems include a control structure that is common to other CPSes, as shown in Figure 1 [1]. In this structure, system administrators use a control network to collect measurements related to physical processes and to issue commands to perform control operations. In remote insider attacks, external attackers penetrate control networks, through which they can use the collected measurements to study operational state and prepare for attacks. To execute attacks, attackers issue or modify commands that are crafted in legitimate formats. I proposed original designs to counter the attack at each stage of its timeline:

1. **Detection: Detect attacks' execution.** I developed the first IDS that fully supports network communications in industrial control systems used by power grids and extended the IDS with a newly designed power flow analysis algorithm.
2. **Response: Remedy attacks' consequences.** I designed a self-healing network infrastructure that, under the constraints of both cyber and physical infrastructures, simultaneously (i) reduces the overhead to reconnect compromised devices to networks and (ii) increases the redundancy of power system measurements.
3. **Preemption: Disrupt and mislead attacks' preparations.** In this ongoing work, I am designing a moving target defense (MTD) mechanism that (i) collects measurements from randomly selected devices instead of from all devices, and (ii) obfuscates measurements from physical operations to provide misleading information, based on which attackers will design malicious operations.



Figure 1. My Research Approach against Remote Insider Attacks in CPS Control Structure.

1 Detection: *Semantic Analysis of Malicious Commands Crafted in Legitimate Formats*

Challenge. After penetrating control networks, attackers can issue malicious commands crafted in legitimate formats. Because those malicious commands do not introduce any anomalies in operating systems and communication networks, existing IDSEs will be ineffective in detecting them.

Deep packet inspection. I developed the first IDS that fully supports the communication protocols used in common CPSes [4]. My IDS is now included in Bro, an open-source network analyzer developed at the University of California at Berkeley. The IDS can fully extract from network packets the semantics related to control operations of power systems. Its capability for deep packet inspection can detect anomalies in networks, such as the inconsistency of operational measurements (caused by attacks or accidents) observed at different locations [5].

Extending network IDS with adaptive power flow analysis. To detect malicious commands crafted in legitimate formats, I extended the IDS with a power flow analysis algorithm to estimate the consequences of executing suspicious commands [7]. A critical challenge in using the existing power flow analysis algorithm is that the algorithm is based on iterations of computation, so can take a long time to finish. To shorten detection latency while preserving accuracy, I proposed a new algorithm that dynamically adapts the parameters of power flow analysis, e.g., the number of iterations of computations to perform, based on information from network communications [8]. This algorithm reduces the computation time by fifty percent relative to the classic AC power flow analysis, and increases the accuracy by two orders of magnitude relative to DC power flow analysis.

Impact. Combining knowledge of cyber and physical infrastructures shows great potential to detect attacks in many CPSes. This research work has drawn interest from many companies, e.g. SEL (Schweitzer Engineering Lab), HoneyWell, NERC (North American Electric Reliability Corp), NRECA (The National Rural Electric Cooperative Association), ABB Inc, Ameren, and Kaspersky Lab. I have deployed the IDS in the control networks of power systems at UIUC (Ameren and Abbot Lab) and collected real operational measurements. The measurement data have enabled me to gain greater domain-specific knowledge of CPSes, especially interactions between the cyber and physical domains. The data were also used in evaluation benchmarks in my subsequent research work. This work is also selected multiple times to be presented for researchers from Department of Energy (DOE) in TCIPG Industry workshops [5][6]. Furthermore, this concept of combining knowledge of cyber and physical infrastructures became the foundation of two significant research projects, one supported by a 2012 National Science Foundation (NSF) grant and another by a 2016 DOE grant. My former colleague Dr. Homa Alemzadeh (now an assistant professor at the University of Virginia) has also successfully applied this concept to detect attacks in a very different CPS: robotic surgical systems [9].

2 Response: *Self-Healing Network for Phasor Measurement Units*

Challenge. In power systems, phasor measurement units (PMUs) are used to collect measurements of voltage phasor approximately 60 to 120 times per second. To improve data collection efficiency, a phasor data concentrator (PDC) combines measurements from multiple PMUs and delivers them to a control center. When a PDC is compromised by attacks and quarantined from communication networks, all connected PMUs and their measurements are lost.

Self-healing network for CPSes. To restore the lost measurements from PMUs, I proposed a self-healing network infrastructure [10]. This network infrastructure uses SDN (software-defined networking) to reconnect PMUs to uncompromised PDCs at runtime [13]. To determine the optimal reconnection scheme, I used an integer linear

programming (ILP) model that simultaneously (i) reduces the performance overhead of reconfiguring communication networks and (ii) increases the redundancy of measurements. The ILP model considers resource constraints in both communication networks and physical infrastructures. ILP solvers can suffer from the dimensionality of the problem; I proposed a greedy heuristic that reduces the computation time by one order of magnitude while maintaining near-optimal results.

Impact. This work has become the basis of several presentations, including the one at INFORMS 2016 Annual meeting [11] and another one at DOE peer reviewing meeting [12]. Furthermore, the methodology of this work became one of the core concepts in a proposal submitted to *NSF Cybersecurity Innovation for Cyberinfrastructure program* (still pending at the time of my application for this position).

3 Preemption: *Randomize Network Communications in CPSes (ongoing work)*

Challenge. Disrupting attackers' preparations can prevent damage from happening. However, doing so is challenging, as attackers' reconnaissance activities can appear legitimate. In many CPSes, such as power grids, a control center periodically retrieves measurements from field devices in plain text. After establishing a foothold in control networks, attackers can passively monitor the transmitted measurements and determine malicious commands.

Randomize network communications of physical devices. To prevent attackers from preparing attacks, I am using a moving target defense (MTD) mechanism to randomize the cyber and physical infrastructures of a power system. Instead of collecting measurements from all devices, a control center collects measurements from randomly selected devices at remote field sites. I achieve the randomized measurement collections by dynamically manipulating the network flows in control networks. Specifically, I allow network traffic to reach "online devices," which send real measurements. Meanwhile, I disallow traffic to "offline devices," but intelligently spoof the measurements on behalf of the "offline devices." Because of this design, attackers' observations are limited, as they lack the knowledge to distinguish real and spoofed measurements.

Construct decoy measurements. To mislead attackers, I spoof decoy measurements for "offline devices." I modify the state estimation algorithm used for power systems to craft the decoy measurements. The crafted decoy measurements strictly follow the physical model of a power system. Being misled by the decoy measurements, attackers fail to design effective strategies for physical damage, and the attack can become easy to detect. By performing evaluations on power systems with real operational data, I find that the decoy measurements are effective against two major threats to power systems, i.e., false data injection attacks and remote insider attacks (discussed before). The probability of successful attacks is reduced from 70% (achieved by attackers with complete system knowledge targeting critical devices) to 5%, which is even less than the probability observed in random attacks. To evaluate the impact of this approach on communication networks, I build six real wide area networks in Geni nationwide network experiment platform and deploy a network operating system named "Onos" to spoof measurements at runtime. As one of the first few researchers to use Onos in the Geni testbed, I strive to develop device drivers to make Onos communicate with switches of different vendors. My experiments on those real networks show that randomizing network communications introduces less than 5% additional latency in a highly congested network.

Impact. Even though this work is under submission, I was invited to present it at a workshop sponsored by NSA (National Security Agency) [14][15]. In addition to preventing attacks from happening, SDN has great potential to increase the resilience of CPSes against accidental incidents, human operational errors, and even natural disasters. SDN can adjust the network infrastructure to reduce the impact of unexpected events. During normal operations, SDN can also help to dynamically allocate network resources to meet quality of service (QoS) requirements in different CPSes. There are still many related research problems that I can focus on in my future career.

Applying SDN in CPSes attracts attention from researchers in academia, e.g., Prof. Rakesh Bobba (Oregon State University), Prof. Anna Scaglione (Arizona State University), and Prof. Ehab Al-Shaer (University of North Carolina-Charlotte), as well as researchers in industry, e.g. Dennis Gammel (Schweitzer Engineering Laboratories) and Thomas Williams (Dispersive Technologies). Through the communications with those researchers, I initiate the collaborations with them to extend this work in the future.

OTHER RESEARCH

Multi-agent communications to detect false data injection attacks. By deploying a software agent in each substation in a power system, I designed a new method based on inter-agent communications to detect false data injection attacks [16]. False data injection attacks compromise measurements to mislead operators into estimating system states that are different from the real ones. Attackers rely on knowledge of the transmission networks of the whole power system to make the compromised measurements bypass the existing state estimation algorithm. In this design, with the help of the software agent, substations can share measurements with their neighbors that are connected through transmission lines. By using the shared measurements, I built in each substation a small virtual grid, which included only that particular substation and its neighbor substations. Because the topology of the transmission network of each virtual grid is different from the one of the whole grid, state estimation in these virtual grids can detect compromised measurements. Because of this protection, attackers will need to bypass the state estimation of all virtual grids in order to make false data injection attacks successful, and that would be very challenging to achieve.

FUTURE RESEARCH

The digital world has entered the era of cyber-physical systems (CPS) and the Internet of things (IoT). The recent Dyn DDoS (distributed denial of service) attack that exploits Internet-connected cameras shows that the security of CPSes and IoTs can pose a severe threat to our daily lives [17].

In the future, I am interested in using system and network virtualization, cyber-physical interdisciplinary method, and big-data analytic to design a customizable security approach that can be adapted based on the domain-specific knowledge of target applications; I will consider different types of attack models, as well as a broad range of CPSes or IoTs that have different computing environments. In addition to attacks that introduce damage, I will consider attack models that can cause economic loss, the leak of critical information, and other negative social impacts (e.g., using misleading information to impact elections). The application-driven approach makes the security enhancement suitable, effective, and efficient in CPSes and IoTs.

There are plenty of opportunities to obtain funding from NSF, NSA, and DOE for my future research. In recent years, NSF has been demonstrating a strong interest in multidisciplinary approaches in the areas of cyber security and cyber-physical systems. Multiple divisions of NSF, including those on *Advanced Cyberinfrastructure (ACI)*, *Computer & Network Systems (CNS)*, and *Information & Intelligent Systems (IIS)*, include programs, such as *Secure and Trustworthy Cyberspace (SaTC)*, *Cybersecurity Innovation for Cyberinfrastructure (CICI)*, and *Cyber-Physical Systems (CPS)*, that focus in part on security and resilience of CPSes [18]. In addition, IT companies, such as Intel, are exploring collaborations with NSF in order to conduct research on designing secure and trustworthy infrastructure for CPSes [19]. In the following, I provide some details of my future research directions.

1 Trusted Computing Base by Network Virtualization in CPS/IoT

As software-defined networking technologies are being deployed in today's network infrastructures, physical network resources can be dynamically assigned to different virtual domains. The virtualization of network resources can become similar to the existing virtualization implemented on hardware and operating systems.

In this direction, I will do research on adapting security approaches based on the system virtualization and make them suitable in the "virtual" network environment. A specific research direction will be to find a method and implementation to select network resources, such as end hosts and communication links, to construct a trusted computing base (TCB) for applications running in CPSes or IoTs. The TCB will serve as a root of trust. By verifying communications with the TCB, we can extend the range of the trusted domain when applications require access to more resources at runtime. Based on this approach, I hope to design a customizable "security-as-you-go" mechanism for different applications, as the "one-size-fits-all" approach can have unacceptably high overhead. To achieve this goal, I plan to follow two procedures:

- *Make SDN fit into networks with different QoS (Quality of Service) requirements in CPSes/IoTs.* Many CPSes and IoTs have very different computing environments, such as wireless networks that consist of energy-constrained devices running customized embedded systems. As communication media change, control applications and security monitoring that are implemented on top of those media should change accordingly. A specific research topic will be the determination of how these applications can provide trusted services despite

unreliable network environments. I plan to integrate research from the area of approximate computing with the existing security analytics. I can identify the “approximable” security logs based on the monitoring of CPS/IoT applications. Then I can add redundancies for “inapproximable” logs to ensure the trustworthiness of security monitoring.

- *Use knowledge of physical control operations to determine the range of the TCB.* The design of the TCB will involve a trade-off between the security and usefulness of applications used in CPSes/IoTs. A small TCB would be easy to maintain, but would require more runtime overhead to ensure its secure interactions with untrusted resources. I plan to focus on a few specific applications in CPSes, namely ones related to distributed power generation in smart grids. I plan to build a “dependency graph” among the cyber and physical components used in these applications. Then I can use formal methods and graph theory to extract the components that are sufficient to observe security properties. In addition, I will implement the theoretic findings in real systems to evaluate their effectiveness and efficiency.

2 Economic Impacts of Cyber-Attacks on CPS/IoT

The majority of the existing research in the fields of cyber security and cyber-physical systems focuses on the physical damage that can be caused by cyber-attacks in CPSes or IoT systems. In addition to physical damage, obtaining economic benefits can be another motive for attackers. Many CPS operators, such as power system operators and oil producers, decide on productions based on auctions that receive bids from different suppliers. As auctions are commonly performed over Internet-based platforms, their integrity, confidentiality, and availability during the transaction have attracted security researchers’ attention. However, there has still been a lack of work focusing specifically on auctions performed in CPSes or IoT systems. The auctions in these contexts have unique characteristics because physical constraints and models (e.g., the balance of generation and consumption in power systems) must be observed continuously.

In this direction, I plan to do research on two threat models. In the first, attackers (e.g., a potential supplier) are working to compromise parameters related to the auction itself; I will study whether such auction-related attacks can result in any perturbations of the control operations or even damage to competitive suppliers. In the second threat model, attackers are targeting the physical parameters and measurements of CPSes; I will study how intelligently compromised information can corrupt the bids from competitive suppliers and allow attackers to obtain economic benefits. What makes this research challenging and also interesting is that suppliers can disguise themselves as users as well. In each threat model, I will consider the trustworthiness of interactions among all three entities, i.e., system administrators, suppliers, and users. If two of these entities are malicious, a “coordinated attack” can be performed, e.g., malicious suppliers deliver misleading and malicious information on behalf of some users. Based on the study of these threat models, I can further design a verifiable and trusted auction platform that can protect the target CPSes.

3 Event-Driven Data Analytics in CPS/IoT

Today’s CPSes or IoTs can generate large amounts of heterogeneous data, which provide valuable information for detecting anomalies. However, data can be polluted by oversampling, making the extraction of valuable information challenging.

Just as environmental benefits can be achieved by reducing production of pollutants (instead of merely cleaning up pollution after the fact), I will use my interdisciplinary expertise on system and network designs to reduce the amount of “polluted” data while maintaining the accuracy with which anomalies are identified. The choice of the range of data to be collected can be modeled as a function of the control application that we are trying to protect and the events observed from the application. Based on the characteristics of the application, I can decide on an initial set of monitors; that decision will be a trade-off between the detection accuracy and the volume of data to be collected. As the control application continues its execution, I can use the physical model to estimate the possible outcome of the application and use this outcome to adjust the monitors to produce more focused data on the events that are critical in detecting anomalies. Following this approach, I can further design new analytics for the “refined” data.

4 Cyber-Physical HoneyNet

In recent years, government agencies, including NSF, DOE, and NSA, have emphasized transfer of research achievement into practical applications. In my future work, I am interested in building a cyber-physical HoneyNet to include and demonstrate the usability of my current and future research. Current HoneyNet projects for CPSes

mimic just their cyber-infrastructures. In this work, I plan to extend the previous work by adding a mimic of physical infrastructures as well. There will be two major challenges in achieving that goal. The first is the need to craft measurements that can prevent leaking of real measurements and follow the physical models in CPSes. My experience in crafting decoy measurements that mislead attacks and follow the physical model in power systems can help me achieve this goal in other CPSes, such as those in water treatment facilities. The second challenge will be to simulate the interactions of cyber and physical infrastructures at runtime. Simulations of communication networks are often done in the form of discrete events, while measurements and states in physical infrastructure change continuously. I plan to study ways to combine the simulations in these two domains without affecting the accuracy of the simulations at runtime.

To demonstrate its usability, I will search for an opportunity to deploy the developed HoneyNet in the public Internet environment after careful evaluations in laboratory environments. Currently, we have limited access to security incidents that have occurred in CPSes or IoT systems, because of the involved companies' need for privacy. I hope to use this cyber-physical HoneyNet to catch attackers who are specifically targeting CPSes and IoT systems, and to analyze their activities in detail. The HoneyNet can be used to design and evaluate original research and can also provide firsthand materials for teaching security courses.

REFERENCES

- [1] **Hui Lin**, Homa Alemzadeh, Daniel Chen, Zbigniew Kalbarczyk, Ravishankar K. Iyer, "Safety-critical Cyber-physical Attacks: Analysis, Detection, and Mitigation," in *Proc. of the Symposium and Bootcamp on the Science of Security (HotSoS)*, pp. 82-89, 2016.
- [2] Gyungho Lee, Yixin Shi, and **Hui Lin**, "Indirect Branch Validation Unit," *Microprocessors and Microsystems*, vol. 33, no. 7 (2009): 461-468.
- [3] Aashish Sharma, Zbigniew Kalbarczyk, James Barlow and Ravishankar K. Iyer, "Analysis of security data from a large computing organization," in *Proc. of IEEE/IFIP 41st International Conference on Dependable Systems & Networks (DSN)*, pp. 506-517, 2011.
- [4] **Hui Lin**, Adam Slagell, Catello Di Martino, Zbigniew Kalbarczyk, and Ravishankar K. Iyer, "Adapting Bro into SCADA: Building a Specification-based Intrusion Detection System for the DNP3 Protocol," in *Proc. of the 8th Cyber Security & Information Intelligence Research Workshop*, Oak Ridge National Lab, Jan. 2013. (Third Place, Best Paper Award).
- [5] **Hui Lin**, "TCIPG demo: detection of a man-in-the-middle attack in scada network," [online] available at: <https://www.youtube.com/watch?v=unb7b8myNvA>, Nov. 5, 2012.
- [6] **Hui Lin**, "Specification-based IDS for the DNP3 protocol," *2014 TCIPG Industry Workshop*, November 12-13th, 2014 (One of four selected student presentation), [online] available at: <https://www.youtu.be/J9OMzhbgNjA>.
- [7] **Hui Lin**, Adam Slagell, Zbigniew Kalbarczyk, Peter Sauer, and Ravishankar K. Iyer, "Semantic security analysis of scada networks to detect malicious control commands in power grids," in *Proc. of ACM CCS Smart Energy Grid Security Workshop*, Nov. 2013.
- [8] **Hui Lin**, Adam Slagell, Zbigniew Kalbarczyk, Peter W. Sauer, and Ravishankar K. Iyer, "Runtime Semantic Security Analysis to Detect and Mitigate Control-related Attacks in Power Grids," *IEEE Trans. Smart Grid*, March 28th, 2016.
- [9] Homa Alemzadeh, Daniel Chen, Xiao Li, Thenkurussi Kesavadas, Zbigniew Kalbarczyk, and Ravishankar Iyer, "Targeted attacks on teleoperated surgical robots: dynamic model-based detection and mitigation," in *Proc. of the 46th IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, 2016.
- [10] **Hui Lin**, Chen Chen, Jianhui Wang, Junjian Qi, Dong Jin, Zbigniew Kalbarczyk, and Ravishankar K. Iyer, "Self-healing attack-resilient pmu network for power system operation," *IEEE Trans Smart Grid*, July 27th, 2016.
- [11] Chen Chen, **Hui Lin**, Jianhui Wang, Junjian Qi, Dong Jin, Zbigniew Kalbarczyk, and Ravishankar K. Iyer, "Self-healing attack-resilient pmu network for power system operation," *INFORMS 2016 annual meeting*, Nov. 13, 2016 (invited presentations).
- [12] NETL Event Management, "Cybersecurity for energy delivery systems peer review," U.S. Department of Energy, National Energy Technology Laboratory, [online] available at: http://www.netl.doe.gov/File%20Library/Events/2016/ceds/2016-CEDS-Peer-Review_Agenda_11-21-16.pdf.
- [13] Xinshu Dong, **Hui Lin**, Rui Tan, Ravishankar K. Iyer, and Zbigniew T. Kalbarczyk, "Software-defined networking for smart grid resilience: opportunities and challenges," in *Proc. of ACM AsiaCCS Workshop on Cyber-Physical System Security*, pp. 61-68. 2015.
- [14] **Hui Lin**, Ravishankar K. Iyer, Zbigniew Kalbarczyk, "RAINCOAT: randomized network communication in power grid cyberinfrastructure to mislead cyber attackers," *the 47th IEEE/IFIP International Conference on Dependable Systems and Networks (DSN '17)*, in submission.
- [15] **Hui Lin**, Ravishankar K. Iyer, and Zbigniew Kalbarczyk, "RAINCOAT: randomization of network connectivity in industrial control systems to mitigate cyber-attacks," in *Workshop on Science of Security through Software-Defined Networking (SoSSDN)*, June 16-17, 2016 (unpublished invited presentation).
- [16] Esther Amullen, **Hui Lin**, Zbigniew Kalbarczyk, and Lee Keel, "Multi-agent system for detecting false data injection attacks against the power grid," in *Proc. of the Second Annual Industrial Control System Security Workshop (ICSS 2016)* (co-first authors), to appear.
- [17] Michael Kan, "Chinese firm admits its hacked products were behind Friday's massive DDoS attack," [online] available at: <http://www.itnews.com/article/3134037/chinese-firm-admits-its-hacked-products-were-behind-fridays-massive-ddos-attack.html>
- [18] National Science Foundation, "Programs: directorate for computer & information science & engineering (CISE)," [online] available at: <https://www.nsf.gov/funding/programs.jsp?org=CISE>.
- [19] National Science Foundation, "NSF/Intel partnership on cyber-physical systems security and privacy (CPS-Security)," [online] available at: https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505047&org=CISE&from=home.