

A Study into Detecting Anomalous Behaviours within HealthCare Infrastructures

Aaron Boddy, William Hurst, Michael Mackay, Abdennour El Rhalibi

Department of Computer Science
Liverpool John Moores University
James Parsons Building, Byrom Street
Liverpool, UK, L3 3AF

A.Boddy@2011.ljmu.ac.uk; {W.Hurst, M.I.Mackay, A.ElRhalibi}@ljmu.ac.uk

Abstract—The theft of medical data, which is intrinsically valuable, can lead to loss of patient privacy and trust. With increasing requirements for valuable and accurate information, patients need to be confident that their data is being stored safely and securely. However, medical devices are vulnerable to attacks from the digital domain, with many devices transmitting data unencrypted wirelessly to electronic patient record systems. As such, it is now becoming more necessary to visualise data patterns and trends in order identify erratic and anomalous data behaviours. In this paper, a system design for modelling data flow within healthcare infrastructures is presented. The system assists information security officers within healthcare organisations to improve the situational awareness of cyber security risks. In addition, a visualisation of TCP Socket Connections using real-world network data is put forward, in order to demonstrate the framework and present an analysis of potential risks.

Keywords—*Medical Device; Healthcare Infrastructures; Data Visualization; Cyber Security*

I. INTRODUCTION

Hospital infrastructures are classified as mission-critical infrastructures [1]. Damage to network communications and the loss of patient data would have a detrimental impact on the healthcare services they provide. In addition, mobile devices are being increasingly deployed within their networks, in the form of applications ranging from biomonitors to materials handling and transportation [2].

Medical devices are instruments, appliances or software, which are used for the purpose of diagnosis, monitoring and treatment of disease or long term injury. They are essential for modern medicine and allow for automated patient monitoring and management functions [3]. These devices have become increasingly lightweight and ubiquitous in recent years. They are available for continuous use by patients and not restricted to use within clinical settings. Their interconnectedness is increasing both wired and wirelessly to external entities including the Internet, or local infrastructure-less wireless networks. These factors have the potential to make healthcare accessible to everyone and to reduce costs. However, they also provide potential for attackers to gain access to sensitive data for opportunistic purposes, such as for profit [4]. Implanted Medical Devices (IMDs) have become increasingly popular due to the precision control of dosage and the rapid access to data they provide to healthcare professionals [2].

Security is crucial for the long term viability of all types of networked medical devices [5]. These devices have the

potential to be tampered with, reprogrammed by unauthorized users or subject to device-specific hazards [5]. Devices can be targeted through their firmware upgrades or through connections to the network interface when connected through remote attacks, in addition to local attacks. For example, telemetry data of an implantable cardiac defibrillator could be reprogrammed remotely by researchers [6]. There are obvious physical hazards and privacy implications for these attacks. Security concerns for IMDs are particularly challenging due to the potential for patient injury or death due to adversarial tampering [2]. Additionally, the malfunctioning of high-profile medical devices results in potential loss of life due to network compromise or medical device tampering, in addition to the cost to the hospital in terms of brand damage due to loss of patient information and enforced regulations.

This risk is further exacerbated by the Bring Your Own Devices (BYOD) revolution. This is a term referring to the technologies allowing employees to access and utilise internal corporate IT resources, with their personal devices [7]. BYOD policies have numerous benefits including reduced costs and improved productivity, convenience and efficiency of work. However, BYOD also carries numerous risks including data loss/leakage or theft, application security, network availability, legal liability and regulatory compliance and loss of brand identity [7]. Additionally, wireless connections on smart devices can be attacked more easily than on a desktop computer. The increasing number of phone operating systems and carrier combinations are invariably changing with technical advancements and becoming outdated quickly [7]. This poses various challenges for IT departments to support and secure.

BYOD poses a particular risk to the healthcare network in that potentially insecure devices are granted access to hospital infrastructure and confidential data. An attacker can use this to their advantage by hacking a BYOD in order to gain back-door access onto a hospital network. For that reason, this paper presents a visualisation of real world hospital data, showing TCP socket connections to a server offering an Active Directory Domain Controller. The research involves analysing the number of differing devices connected to the network and points of vulnerability or entry for potential attackers.

Specifically, in this paper, a system design for modelling data flow within healthcare infrastructures is presented. An illustration of the framework operation, using network TCP

Socket Connections data of a Liverpool-based hospital, is employed as a case study. The system provides a real-time solution to analyse cyber-transactions and traffic. It assists information security officers within healthcare organisations to improve the situational awareness of cyber security risks.

The remainder of this paper is as follows. Section II presents a literature review of the background research on industrial network, medical device security and visualization techniques. Section III outlines our system design. Section IV presents our results and a sample of test data. Section V discusses conclusions within the work and the future work to be done.

II. BACKGROUND RESEARCH

Medical device monitoring systems must be highly automated to reduce user involvement in deployment, operation and management [3]. A lack of security for healthcare devices could mean incorrect data being introduced or legitimate data being modified or suppressed by adversaries [5]. This may lead to both loss of patients' privacy and potential physical harm to the patient. Many medical devices employ the use of wireless communication which technologies are intrinsically vulnerable and attacks can exploit this at patient side [3]. In this section, medical device security is discussed, along with the motivation and research problem being addressed.

A. Medical Device Security

For safety-critical Medical Cyber-Physical Systems (MCPSs), the ability to detect attackers, whilst limiting false alarms, is of critical importance [4]. Attackers who penetrate medical cyber-physical systems have the potential to cause harm to patients through reprogramming devices [8]. Currently typical known targets include the patient, the data, the device and the interaction between the internal network and MCPSs [8].

Real world attacks on MCPS components are increasing with the attackers aiming to cause node compromise, particularly against Insulin Pumps and Cardiac Devices [4]. These attacks can be initiated through over-the-air software updates, stack overflow exploits or logic bombs through third party developers. It is clear that security is a growing concern, particularly for small medical devices attached to a patient [8].

Remote communication channels, such as those used by IMD insulin pumps, can be compromised by an adversary within wireless transmission range [2]. This can be achieved through a strong antenna or subversion of a networked device. In addition, the process can be achieved by being physically close to the device, allowing the adversary to inject a potentially fatal dose into the patient. For example, 100mg of insulin into a patient with normal blood sugar, could induce a diabetic coma [2].

The most common outcome of a cyber-attack on a healthcare-based system, however, is the unavailability of patient care due to computer outages. Real world incidents have happened with dramatic effects. In one case, a virus in a catheterization lab resulted in patients being transferred to another hospital [9]. In another instance, a factory-installed device arrived at a hospital infected with malware [9].

Medical devices also experience unexpected interactions between devices and systems that have not been triggered maliciously [5]. Wireless technologies can suffer from interference caused by these devices and make it easier for malicious persons to access the network. These devices often have no safeguards and are susceptible to buffer overflows when unexpected signals are received. Several devices have been proven to be affected in this way. Such as, Mechanical Ventilators, which have been susceptible to total switch-off and change in ventilation rate; Syringe Pumps which have been completely stopped; External Pacemakers which have malfunctioned and Renal Replacement Devices which have also completely stopped [10].

Most wearable devices record and collect medical data and then transmit to a remote server. This leaves the data vulnerable to 'man in the middle' attacks [11]. It has been considered that security policies should be implemented between the wearable devices and the remote server. However, this is unsuitable, as it does not protect the patient wherever they travel, unless they carry a portable device to perform the security policy and communicate with the server about their persons at all times [11].

Mobile devices have unique security challenges; being custom built, low-power and resource constrained, which lack the processing and security capabilities of a computer [2]. Insulin pumps and pacemakers, for example, communicate wirelessly with a wearable external monitoring and control unit which needs to be accessible to emergency responders and medical personnel. It is a challenge, therefore, to implement effective key-based encryption techniques due to the complexities of key management and revocation, in addition to the issues of limited power and heat-dissipation within the device [2].

There are three key issues with securing updates for embedded devices; namely, Untrusted Infrastructures, Sporadic Network Connectivity, and Limited Local Resources [12]. Embedded devices often do not have user interfaces, meaning a user cannot give consent to a software update. In an example such as an RFID tag, therefore, they must communicate entirely through untrusted readers and infrastructures. With regards to sporadic network connectivity, RFID tags and other implanted devices can only connect to networks when in range. Finally, RFID tags lack the resources for advanced cryptographic protocols, they have limited working memory and so offload computationally exhaustive operations onto RFID readers [12].

B. Medical Infrastructures

Figure 1 presents an overview of a typical network infrastructure for enabling remote access within the hospital. The layout enables staff the ability to work and provide on-call services remotely. The blueprint demonstrates where the hospitals LANs and VLANs are situated on a hospital network; in addition to where the firewalls are placed in relation to the Internet. Figure 1 also displays the relationship between the hospitals 'Community of Interest Network' (CoIN)/WAN and the N3 (a WAN used to connect many sites across the NHS).

which medical device data input is visualised in order to be analysed and manipulated by the system operator. In this way, the data can be assessed for potential points of vulnerability within the network. Patterns of behaviour and anomalous data can be studied further and addressed.

The system receives input data directly from the health-care network, which it processes and stores until it is called upon by the system operator. When the data is requested it is processed within the Visualisation Generation Engine and presented to the operator through a Graphical User Interface.

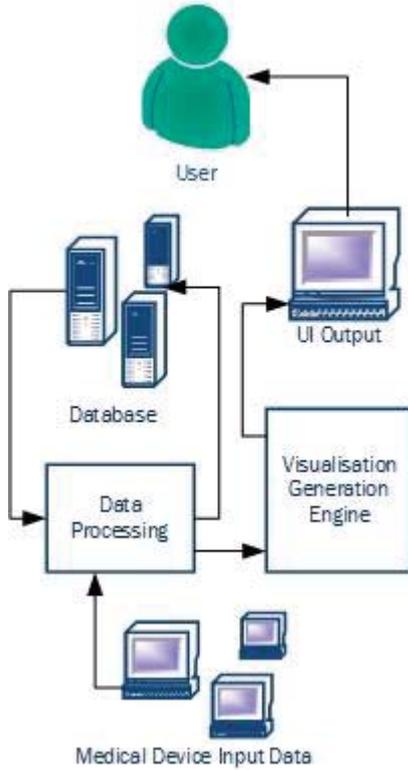


Figure 2. System Design Framework

The components in Figure 2 are explained as follows:

- **Medical Device Input Data:** This data is collected and transmitted by all or selected medical devices used by patients.
- **Data Processing:** This component of the framework involves processing the medical device input data in order to remove irrelevant data from the dataset. The process ensures the data is transmitted and stored safely and securely. This stage ensures that only pertinent and useful data is analysed.
- **Database:** This component stores the data when not in use by the other components. In addition to this function, the database stores known attack behaviours of datasets to compare the input data against. This allows the system to identify anomalous and erroneous data within the system in order to further investigate potentially malicious activity.
- **Visualisation Generation Engine:** This component generates the visualisation for the user. The component uses the system operators input and calls upon the data stored in the

database component, which is then processed and visualise within the generation engine and passed onto the UI Output.

- **UI Output:** This is the output with which the operator interacts. The operator can manipulate the visualisation in this way and set their own data parameters in order to increase their situational awareness of the data flow within the healthcare infrastructure.

IV. EXPERIMENTS

Building on the system design, the following section presents an initial dataset captured from a UK-based hospital network. The data is visualised as a demonstration of the research approach.

The data visualised is real-world network data from a Liverpool-based hospital network employing over 4,000 staff. The data collected is a snapshot of the network infrastructure using the network statistics (netstat) command-line in order to capture incoming and outgoing Transmission Control Protocol (TCP) Data. A sample data and visualisations of netstat snapshot data conducted on two servers, is presented in this section.

A. Patient Administration System (PAS) netstat

Firstly, the network statistics on a server which hosts the hospital's Patient Administration System (PAS), as seen in Table 1 and Figure 3, is presented. Secondly, a network statistics capture on one of the Active Directory Domain Controllers at the hospital, as seen in Table 2 and Figure 6, is put forward.

In Table 1 a sample of the netstat data analysed is shown displaying firstly, the connection type, secondly the IP source connecting to the PAS, thirdly the target of the IP address (the PAS server), and fourthly the state of the connection. In Figure 3 the full dataset is visualised.

Table 1. Patient Administration System – TCP Socket Connections Sample Data (Anonymised)

Conn	Source	Target	State
TCP	0.0.0.0:49876	***SIGAPP05:0	LISTENING
TCP	10.52.***.224:139	***SIGAPP05:0	LISTENING
TCP	10.52.***.224:819	4***sophosman01:**881	ESTABLISHED
TCP	10.52.***.224:819	4***sophosman01:**217	ESTABLISHED
TCP	10.52.***.224:819	4***sophosman01:**772	ESTABLISHED

Figure 3 is a visualisation of data connections for the PAS system at the Liverpool hospital. The different nodes, depicted by blue circles, represent devices accessing the PAS. The clusters of nodes represent different servers using the PAS.

The visualisation shows data connections between each of the various servers which compose the PAS solution. This includes a server, which hosts the PAS itself, a server which

hosts SQL databases, and a server which hosts the anti-malware solution.

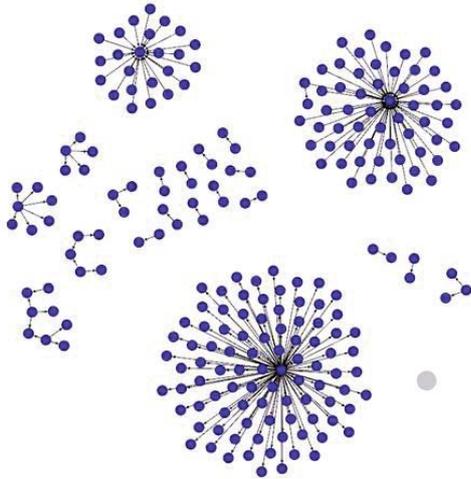


Figure 3. Patient Administration System – TCP Socket Connections Visualisation

B. Active Directory Domain Controller (DC) netstat

Table 2 is a sample of the netstat data analysed is shown. Firstly, the connection type is displayed. In addition, secondly the IP source connecting to the active directory domain controller is presented. Thirdly the target of the IP address, i.e. the device name on the domain controller, and fourthly the state of the connection are both presented. As before, the full dataset is visualised in Figure 6.

Table 2. Domain Controller – TCP Socket Connections Sample Data (Anonymised)

Conn	Source	Target	State
TCP	10.52.***.15:135	***0395:63091	ESTABLISHED
TCP	10.52.***.15:135	***0395:63160	ESTABLISHED
TCP	10.52.***.15:135	***0645:50562	ESTABLISHED
TCP	10.52.***.15:135	***3635:49164	ESTABLISHED
TCP	10.52.***.15:135	***3635:49195	ESTABLISHED

In Figure 4 the most frequent items for Foreign Addresses’ on the Domain Controller is shown. The most frequent value is that of an asterisk as the port has not been established indicating that at this time the Domain Controller had approximately 25% of its ports connected.

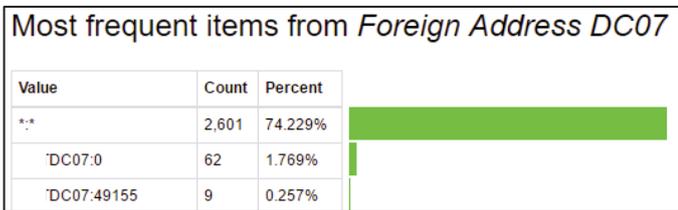


Figure 4. Domain Controller – Frequent Items – Foreign Address

In Figure 5, however, the output is more varied. The largest IP address item counts comprise around 6% of the local IP

address ports but this number quickly decreases and the majority of the ip addresses are unique values.

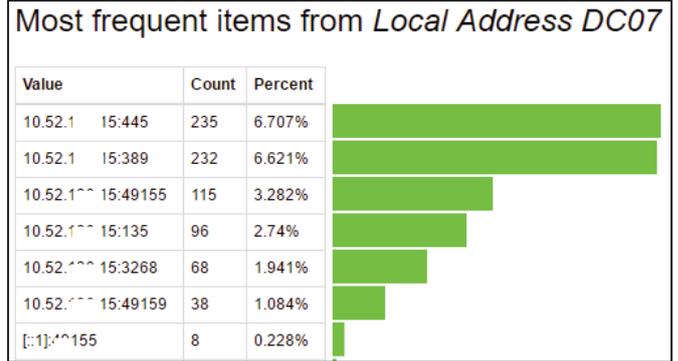


Figure 5. Domain Controller – Frequent Items – Local Address

Visualising complex healthcare network to detect anomalous behaviour is a significant challenge. This is demonstrated in Figure 6, which reveals the complex nature of the networked systems. Each of the nodes displays a different device connected to the domain controller, such as a physical computer or a laptop, or a virtual cloud session. This is further exacerbated by the fact that the data presented is merely a snapshot of data on one of the domain controllers on the hospital network.

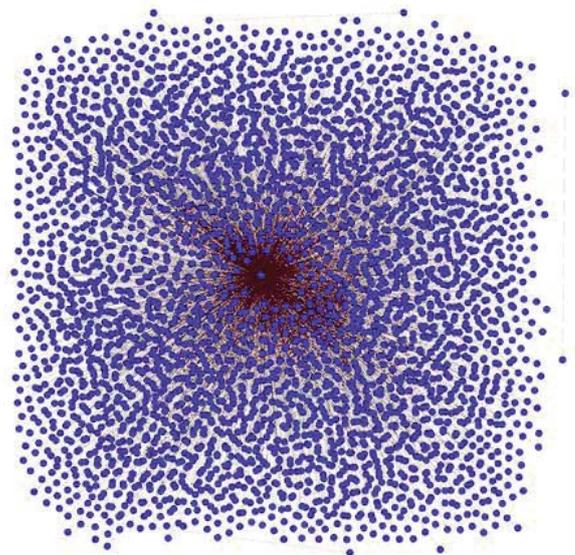


Figure 6. Domain Controller – TCP Socket Connections Visualisation

C. Discussion and Future Work

The visualised data demonstrates how even a small sample of the overall data connections within a section of the hospitals network infrastructure is difficult to analyse for anomalous data behaviours due to sheer quantity of data. Even once the data has been visualised, interactions between the user and the visualisation itself is a challenge. For this reason, captured network data needs to be pre-filtered in order to simplify the visualisation and the visualisation process.

The visualisation will bring together several related data sets and present them in such a way as to identify relationships between them, using the system design presented in this paper. How new data is added to the infrastructure, and when and how it moves within it, will be visualised in order to assist end users in finding the potential cyber vulnerabilities within the health care infrastructure. This will enable end users to be able to identify where further cyber security systems need to be put in place. In addition, identifying where best practices and policies can be implemented to minimize the risk of a cyber-attack on highly confidential personal data. The design stage will involve testing and modifying various data analysis techniques in order to ascertain the most relevant technique for the proposed system. Specifically, the data analysis techniques will involve using machine learning algorithms to interpret dataset patterns and identify potential on-going cyber-attacks.

V. CONCLUSION

With healthcare organisations using electronic records, cyber-based transactions and mobile electronics, the risk of a data breach is an increasing concern. Healthcare data is intrinsically valuable; the repercussions of data compromise within healthcare infrastructures can range from loss of patient privacy and fraud to patient injury or potentially death. Therefore protecting private patient data and preventing data compromise is critically important. Visualisation can be used as a tool for cyber security officers within healthcare organisations to increase their situational awareness of data flow and actively address this issue. Additionally, visualisation tools allow system operators to be proactive about cyber security within healthcare organisations. This is in contrast to the accepted and fundamentally flawed approach of reactivity to cyber security attacks, which do not attempt to address the underlying security flaws within healthcare organisations.

In this paper, a framework for data visualisation within healthcare infrastructures is proposed. Additionally, two real-world datasets were analysed and visualised and are presented as a study to demonstrate the scale of the research challenge. These datasets demonstrate that detecting anomalous data behaviours in healthcare infrastructures is challenging. Even once data has been visualised it is challenging to manage and manipulate in meaningful ways in order to identify irregular and anomalous data which may indicate potentially malicious behaviour and prevent data compromise. Future work will involve the use of date filtering techniques to simplify the visualisation of normal network traffic. This will facilitate understanding patterns of data behaviours and highlight abnormal data behaviours in the network which are the result of an attack taking place.

REFERENCES

- [1] M. Rong, C. Han, and L. Liu, Critical Infrastructure Failure Interdependencies in the 2008 Chinese Winter Storms, in *2010 International Conference on Management and Service Science*, 2010, pp. 1–4.
- [2] R. Skowrya, S. Bahargam, and A. Bestavros, Software-Defined IDS for securing embedded mobile devices, in *2013 IEEE High Performance Extreme Computing Conference (HPEC)*, 2013.
- [3] A. Sawand, S. Djahel, Z. Zhang, and F. Nait-Abdesselam, Multidisciplinary approaches to achieving efficient and trustworthy eHealth monitoring systems, in *2014 IEEE/CIC International Conference on Communications in China (ICCC)*, 2014, pp. 187–192.
- [4] R. Mitchell and I-R. Chen, Behavior Rule Specification-Based Intrusion Detection for Safety Critical Medical Cyber Physical Systems, *IEEE Trans. Dependable Secur. Comput.*, vol. 12, no. 1, pp. 16–30, Jan. 2015.
- [5] D. Arney, K. K. Venkatasubramanian, O. Sokolsky, and I. Lee, Biomedical devices and systems security., *Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.*, vol. 2011, pp. 2376–9, Jan. 2011.
- [6] D. Halperin, T. S. Heydt-Benjamin, B. Ransford, S. S. Clark, B. Defend, W. Morgan, K. Fu, T. Kohno, and W. H. Maisel, “Pacemakers and Implantable Cardiac Defibrillators: Software Radio Attacks and Zero-Power Defenses,” in *2008 IEEE Symposium on Security and Privacy (sp 2008)*, 2008, pp. 129–142.
- [7] R. Ogie, Bring Your Own Device: An overview of risk assessment., *IEEE Consum. Electron. Mag.*, vol. 5, no. 1, pp. 114–119, Jan. 2016.
- [8] Q. Shafi, Cyber Physical Systems Security: A Brief Survey, in *2012 12th International Conference on Computational Science and Its Applications*, 2012, pp. 146–150.
- [9] D. B. Kramer, M. Baker, B. Ransford, A. Molina-Markham, Q. Stewart, K. Fu, and M. R. Reynolds, Security and privacy qualities of medical devices: an analysis of FDA postmarket surveillance., *PLoS One*, vol. 7, no. 7, p. e40200, Jan. 2012.
- [10] R. van der Togt, E. J. van Lieshout, R. Hensbroek, E. Beinat, J. M. Binnekade, and P. J. M. Bakker, Electromagnetic interference from radio frequency identification inducing potentially hazardous incidents in critical care medical equipment., *JAMA*, vol. 299, no. 24, pp. 2884–90, Jun. 2008.
- [11] J. Kim, B. J. Lee, and S. K. Yoo, Design of real-time encryption module for secure data protection of wearable healthcare devices., *Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.*, vol. 2013, pp. 2283–6, Jan. 2013.
- [12] A. Bellissimo, J. Burgess, and K. Fu, Secure software updates: disappointments and new challenges, p. 7, Jul. 2006.
- [13] H. Koike, K. Ohno, and K. Koizumi, Visualizing cyber attacks using IP matrix, in *IEEE Workshop on Visualization for Computer Security, 2005. (VisSEC 05)*, 2005, pp. 91–98.
- [14] M. Bastian, S. Heymann, and M. Jacomy, Gephi: An Open Source Software for Exploring and Manipulating Networks, in *Proceedings of the Third International Conference on Weblogs and Social Media, ICWSM 2009, San Jose, California, USA, May 17-20, 2009*, 2009.
- [15] H. J. Kairam Sanjay, MacLean Diana, Savva Manolis, GraphPrism: Compact Visualization of Network Structure, in *Advanced Visual Interfaces*, 2012.
- [16] K. Ohno, H. Koike, and K. Koizumi, IPMatrix: An effective visualization framework for cyber threat monitoring, in *Proceedings of the International Conference on Information Visualisation*, 2005, vol. 2005, pp. 678–685.
- [17] H. J. Selassie David, Heller Brandon, Divided Edge Bundling for Directional Network Data, *IEEE Trans. Vis. & Comp. Graph. (Proc. InfoVis)*, 2011.
- [18] J. Stoll and R. Z. Bengez, Visual structures for seeing cyber policy strategies, in *2015 7th International Conference on Cyber Conflict: Architectures in Cyberspace*, 2015, pp. 135–152.
- [19] M. Merabti, M. Kennedy, and W. Hurst, Critical infrastructure protection: A 21st century challenge, in *2011 International Conference on Communications and Information Technology (ICCIT)*, 2011, pp. 1–6.
- [20] J. J. Walker, T. Jones, and R. Blount, Visualization, modeling and predictive analysis of cyber security attacks against cyber infrastructure-oriented systems, in *2011 IEEE International Conference on Technologies for Homeland Security (HST)*, 2011, pp. 81–85.
- [21] P. Venkatasubramaniam, J. Yao, and P. Pradhan, Information-Theoretic Security in Stochastic Control Systems, *Proc. IEEE*, vol. 103, no. 10, pp. 1914–1931, Oct. 2015.