



AFRL-RI-RS-TR-2022-157

HBCU-MI CENTER IN CYBER SECURITY

NORFOLK STATE UNIVERSITY

NOVEMBER 2022

FINAL TECHNICAL REPORT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

STINFO COPY

**AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE**

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report is the result of contracted fundamental research deemed exempt from public affairs security and policy review in accordance with SAF/AQR memorandum dated 10 Dec 08 and AFRL/CA policy clarification memorandum dated 16 Jan 09. This report is available to the general public, including foreign nations. Copies may be obtained from the Defense Technical Information Center (DTIC) (<http://www.dtic.mil>).

AFRL-RI-RS-TR-2022-157 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE CHIEF ENGINEER:

/ S /

LAURENT Y. NJILLA
Work Unit Manager

/ S /

JAMES S. PERRETTA
Deputy Chief,
Information Warfare Division
Information Directorate

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

1. REPORT DATE		2. REPORT TYPE		3. DATES COVERED	
NOVEMBER 2022		FINAL TECHNICAL REPORT		START DATE APRIL 2015	END DATE MAY 2022
4. TITLE AND SUBTITLE HBCU-MI CENTER IN CYBER SECURITY					
5a. CONTRACT NUMBER FA8750-15-2-0120		5b. GRANT NUMBER N/A		5c. PROGRAM ELEMENT NUMBER OTHER AF	
5d. PROJECT NUMBER		5e. TASK NUMBER		5f. WORK UNIT NUMBER R1NP	
6. AUTHOR(S) Chung-Chu 'George' Hsieh					
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) NORFOLK STATE UNIVERSITY 700 Park Ave Ste 310 Norfolk VA 23504-8090				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Research Laboratory/RIGA 525 Brooks Road Rome NY 13441-4505			10. SPONSOR/MONITOR'S ACRONYM(S) AFRL/RI		11. SPONSOR/MONITOR'S REPORT NUMBER(S) AFRL-RI-RS-TR-2022-157
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited. This report is the result of contracted fundamental research deemed exempt from public affairs security and policy review in accordance with SAF/AQR memorandum dated 10 Dec 08 and AFRL/CA policy clarification memorandum dated 16 Jan 09.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT Sustainable institution to conduct and attract research to which it otherwise would not have access. Increased level of student involvement, papers produced, and impact on the educational program. Demonstrates ability to work within DoD business processes. Desire by Service S&T organizations to initiate their own HBCU-MI programs based on this program's success.					
15. SUBJECT TERMS CoE Cyber, Cyber Analysis, Simulation and Experimentation Environment (CASE-V), Open Source Intelligence (OSINT)					
16. SECURITY CLASSIFICATION OF:				17. LIMITATION OF ABSTRACT	
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U	SAR		18. NUMBER OF PAGES 126
19a. NAME OF RESPONSIBLE PERSON LAURENT Y. NJILLA				19b. PHONE NUMBER (Include area code) N/A	

Table of Contents

List of Figures	ii
List of Tables	iii
1.0 SUMMARY	1
2.0 INTRODUCTION	5
3.0 METHODS, ASSUMPTIONS, AND PROCEDURES	14
4.0 RESULTS AND DISCUSSION	60
5.0 CONCLUSIONS.....	68
6.0 REFERENCES	71
APPENDIX A – Publications and Presentations	94
LIST OF SYMBOLS, ABBREVIATIONS, AND ACRONYMS.....	118

LIST OF FIGURES

Figure 1a: An example of adversarial attack in wireless communications. An adversary adds perturbation to a test sample to create an adversarial example. Then, the pre-trained ML model wrongly predicts the test sample.....	7
Figure 2. CASE-V architecture.....	10
Figure 3: CASE-V testbed connectivity among major components.....	13
Figure 4: CASE-V automation and UI architecture.....	20
Figure 5: Register and create a new user.	21
Figure 6: Twitter Streaming API [2].....	22
Figure 7: NLP Pipeline [6].....	24
Figure 8: Key components of the Elastic Stack.	27
Figure 9: Overview of the arrangement of the materials in this review.	29
Figure 10: Communication system diagram.	32
Figure 11: System model for the applications of DL in wireless communications.....	33
Figure 12: Categorization of attacks based on adversarial machine learning.....	39
Figure 13: Attacks in adversarial machine learning.	42
Figure 14: Significant terms distribution (pie chart).....	62
Figure 15: Significant terms aggregation (tag cloud).	63
Figure 16: Total count of tweets with relevant information.	63
Figure 17: Tweets with top 20 significant terms.	64
Figure 18: All relevant tweets.....	65

LIST OF TABLES

Table 1: CASE-V TESTBED EQUIPMENT	14
Table 2: OpenStack Platform Hardware Components.....	15
Table 3: OpenStack Platform Computing and Storage Capacity.....	16
Table 4: OpenStack Platform Network Traffic Types.....	16
Table 5: Hadoop Hardware Components.....	18
Table 6: Hadoop Hardware Switches	18
Table 7: Hadoop Hardware Switches	19
Table 8: DEEP LEARNING APPLIED TO WIRELESS SYSTEM FUNCTIONALITIES	34
Table 9: Review of Previous Works On AML Attacks In Modulation Classification	48
Table 10: Review of Previous Works On AML Attacks In Modulation Classification(a).....	49
Table 11: Review of Previous Works On AML Attacks In Modulation Classification(b)	50
Table 12: Review of Previous Works On AML Attacks In Modulation Classification (c).....	51
Table 13: Review of Previous Works On AML Attacks In Spectrum Sensing.....	53
Table 14: TWEET SAMPLING RESULTS	67

1.0 SUMMARY

The Center of Excellence in Cyber Security at Norfolk State University (NSU) was established on April 14, 2015 with \$5 million funding plus an additional \$1 million in 2020 from the Office of the Under Secretary of Defense for Research and Engineering (OUSD(R&E)) at the Department of Defense (DoD) through a cooperative agreement managed by the Air Force Research Lab (AFRL)/Rome. The period of performance for the agreement was originally five years from April 14, 2014, to April 14, 2020. The end date was extended to May 31, 2021, with an additional funding of \$1,000,000. A no cost extension to May 31, 2022, was approved in May 2021. So, the overall period of performance for the project is April 14, 2014, to May 31, 2022, with a total funding of \$5,984,352. The CoE-Cyber has two academic collaborators: Old Dominion University (ODU) and its Virginia Modeling, Analysis and Simulation Center (VMASC), and Tennessee State University (TSU). The period of performance for this award is five years which ends on April 14, 2020.

The purpose of this document is to provide a high-level summary of the major activities, accomplishments, and challenges of the CoE-Cyber during its first year of operation from April 14, 2015 to March 31, 2016.

During its first reporting period, the CoE-Cyber has made excellent progress against the planned activities to achieve its three major objectives:

The CoE-Cyber conducts basic research to develop a cloud-enabled, big-data-analytics-capable Cyber Analysis, Simulation and Experimentation Environment (CASE-V) for enhancing situational awareness and decision support for cyber defense and cyber training (especially for countering future Advanced Persistent Threats, or APTs). Perform research-related education and outreach activities. Be a valuable resource for cyber security research, education, outreach and practices for the HBCU/MI community and the nation.

The overall accomplishments of the COE during this reporting period are summarized below: Research: Established four research thrusts. Launched 34 research projects. Engaged a total of 14 faculty and project scientist. Outreach: Led or contributed to 15 major activities including cybersecurity summer internship programs and summer camps.

Capacity Building: Submitted 13 new proposals and assisted the preparation of 2 additional proposals. Six of these new proposals have received approvals for funding. NSU received three new awards for a total funding of \$422,000, submitted two new proposals and worked on three additional continuing projects. ODU (Dr. Shetty) received 6 new awards for a total funding of \$4,000,000 and leveraged 6 continuing projects.

In addition, NSU has continued to include budget and personnel support for the COE in its business operation to help sustain and grow cybersecurity related programs at NSU going forward.

Education: Led or contributed to 8 major educational initiatives and objectives. Provided research assistantship support to a total of 21 students at Ph.D. M.S. and B.S. levels. Four of these research assistants graduated during this reporting period. Mentored and supervised 36 students for their degree-required capstone research projects. **Partnerships:** Established partnerships with 13 major organizations and intuitions. **Publications:** Worked on 32 publications (including papers/abstracts published, accepted, submitted or being prepared - to the best of our knowledge) for research related to the CoE-Cyber. **Meetings and Presentations:** Organized, participated, or presented in 27 major meetings for purposes related to the COE. **Training, Conferences and Professional Development:** Participated in 11 major training courses, conferences, or workshops in cybersecurity.

During this reporting period, a major issue and concern has been that the two new full-time positions for the CoE-Cyber – (a) Research Faculty/Operations Manager, and (b) Cybersecurity Researcher – have not been filled, due to delay in processing by the Human Resources Department at NSU. These two positions were posted on the state job site during February - March, 2016. Candidates for the Cybersecurity Researcher position are being screened and selections are expected to be made by June 2016. The first position will be revised and reposted during 2Q2016. Another recovery plan is to use the unspent funds for these two positions to hire one more, originally unplanned, research faculty so we can catch up more readily.

Another relatively significant event is that the work planned for TSU for Years 2 - 5 is being transitioned to ODU/VMASC with the upcoming move of Dr. Sachin Shetty from TSU to ODU/VAMSC effective July 2016. The impact of this change is expected to be net positive for the CoE-Cyber as it will further facilitate the interaction and coordination among the CoE-Cyber related personnel, plans and activities.

During this reporting period, the COVID-19 pandemic has caused significant impacts on the CoE-Cyber. Impacts were felt across the board on (a) operation of the university and the center; (b) technical research programs especially testbed implementation and operation which require significant physical presence in the labs working with hardware, facilities, and networks; (c) students learning and research; and (d) transition and sustainment.

Despite these impacts and aided by the implementation of various mitigation measures, the CoE-Cyber has kept making strong progress on numerous research activities, especially those that could be performed remotely and individually. The CoE-Cyber was also able to resume the CASE-V Testbed installation, configuration, and enhancements efforts once the lab shutdown policy was starting to be relaxed in November 2020 timeframe. Under the difficult pandemic circumstances, it was a very significant accomplishment to complete the initial installation and configuration of the new high-performance computing cluster and the newly constituted small OpenStack cloud platform.

The CoE-Cyber is very grateful that a no cost extension has been granted and helped the CoE-Cyber to recover from the impacts of the COVID-19 pandemic. Going forward, the CoE-Cyber plans to focus on

- 1) Completing the originally planned activities for the cybersecurity testbed including
 - The high-performance computing subsystem
 - Enhancements to the cloud computing and networking infrastructures
 - Selected user-created projects on cybersecurity and machine learning
- 2) Sustaining the research programs and cybersecurity testbed for the future
 - Securing additional support and resources from NSU
 - Pursuing new funding and partnership opportunities within and without NSU

In summary, the Center of Excellence in Cyber Security at Norfolk State University has established a solid foundation during its first year of operation. The CoE-Cyber has made excellent progress towards its three major objectives and accomplished significant results in all three areas. Going forward, the CoE-Cyber will continue with the established research thrusts and projects; execute the plans for the next phases of the overall mission; and continue to build capacity, en-

hance partnerships, and explore new opportunities. No major adjustments to the original plans are required or envisioned at this moment.

Leveraging the research and education capabilities of the CoE-Cyber center, many new grants have been obtained recently built on the momentum of active research and made the center sustainable. With the strong support of the government agencies especially DOD and our academic and industrial partners, the team is confident that the CoE-Cyber center will further improve its research and education capacity and continue to train students especially underrepresented minorities to be highly qualified workforce and contribute to DOD missions and the nation for years to come.

2.0 INTRODUCTION

The research objective of this project is to *amplify the power of Cyber Intelligence and Cyber Defense Explore innovative theories, methodologies and techniques in cyber intelligence and cyber defense, and leverage big data analytics. Cyber Modeling, Simulation, Analysis & Decision Support. Develop foundational modeling, simulation, analysis, intelligent systems and decision support capabilities, with a focus on human behavior modeling, simulation scalability, and on live-virtual-constructive simulation and kinetic simulation system interoperability. Cyber Training, and Intrusion Detection and Response: Develop intrusion detection, mitigation and forensic analysis capabilities to defend against future advanced threats, and develop integrated training capabilities for cyber defense and response. CASE-V Framework and Testbeds: Develop synthetic framework and testbeds with distributed, virtual and big data-enabled infrastructures capable of interoperability with selected existing large-scale cyber security testbeds.* The CoE-Cyber center's mission is to accelerate research and education in cyber security for science and engineering to transform our ability to effectively address and solve many complex problems posed by cyber threats specifically for military applications.

The impact of cyber-crime has necessitated the government and private-sector organizations across the world to tackle cyber threats. All sectors are now facing similar dilemmas of how to best mitigate against cyber-crime and implement best practices effectively. Extracting actionable and high value intelligence by harvesting public information is emerging rapidly as an important means for cyber defense. As the amount of information available from open sources rapidly increases, countering cyber-crime increasingly depends upon advanced software tools and techniques to collect, process, analyze and leverage the information in an effective and efficient manner.

Open Source Intelligence (OSINT) refers to intelligence that has been derived from publicly available sources such as news articles, blogs, and social media. Using automated OSINT collection and analysis tools and methods, government and private sectors alike can be better prepared to avert cyber-attacks before they hit them or mitigate and respond to incidents more effectively. As an example, Twitter has become an obvious and very important source to gather cybersecurity information, as many professionals in this field use this social media to share in-

formation on threats and attacks as they happen, and to observe trends in the cybersecurity landscape. Twitter makes it possible and convenient for those interested to get on-the-fly notice of cyber events. However, manual methods for cybersecurity OSINT collection and analysis are inadequate due to the volume of information to be analyzed and the speed with which threats arise. Unlike traditional cybersecurity products designed based on known attacks, next-generation security intelligence systems must provide real-time alerts and permit immediate intelligence to be gathered about a threat.

To help automate the cybersecurity OSINT collection and analysis process, we designed and implemented a prototype system named TwitterOSINT for gathering and analyzing cybersecurity related information (e.g., threats, vulnerabilities and attacks) posted on Twitter.

The prototype system was implemented in Java and it used Twitter Streaming API to download relevant tweets based on a set of user provided keywords. To process and analyze the tweets which are unstructured texts, the TwitterOSINT system implemented a Natural Language Processing (NLP) pipeline using the Stanford CoreNLP library and language models for foundational NLP capabilities, and the Stucco cyber domain-specific entity extraction library developed and provided by the Oak Ridge National Lab (ORNL) to extract cyber-domain entities from unstructured texts. Both libraries use advanced machine learning and statistical techniques to provide high performance text analytic capabilities. The processed tweets were then stored in a JavaScript Object Notation (JSON), a lightweight and human readable data-interchange format, file along with all the annotations (e.g., date, username, text, and cyber-domain entities) produced by the NLP pipeline for these tweets.

Machine learning (ML) in general and deep learning (DL) has found rich applications in various domains such as computer vision (CV) and natural language processing (NLP). Motivated by the success in those domains, there has been a growing recent interest in the development of artificial intelligence driven solutions for wireless communications using radio frequency (RF) data [1] [3]. For example, convolutional neural network (CNN) models have been used for modulation recognition [4] and channel decoding [5]. In addition to CNN, feed-forward neural network (FNN) and long short-term memory (LSTM) models have been used for device fingerprinting and identification [6] [9]. On the other hand, various deep neural network (DNN) mod-

els have been proposed for spectrum sensing and prediction [10], wireless resource management [11], beam prediction for initial access [12] and reconfigurable intelligent surfaces [13]. Although the application of DL-based techniques to wireless communications has shown promising results to solve complex problems for which analytical solutions are either unavailable or computationally infeasible, there are concerns in terms of reliability and robustness of such data-driven methods when compared to traditional signal processing methods in real-world deployments. One of such concerns is the presence of adversaries that may target the training process and/or testing (inference) process of DL.

With the increased adoption of DL, it is inevitable that adversaries will explore launching new attacks particularly on DL models, thus leading to a new paradigm called adversarial machine learning (AML). AML studies learning in the presence of adversaries with the goal of understanding the impact of AML attacks, defending against such attacks, and eventually developing secure DL systems. One type of AML attack (namely adversarial attack or evasion attack) occurs when an adversary crafts small perturbations to the original input of a neural network or develops thereby manipulating the operation of the DL system to cause an error in the inference process. These perturbations are not just white noise samples, but they are specifically crafted to produce a vector in the input feature space that is capable of misleading the developed DL model. Classical examples of such attacks include adding small incremental value in the model's gradient direction with respect to the inputs or by solving a constrained optimization problem to produce a vector in the input feature space that can mislead the target DL model.

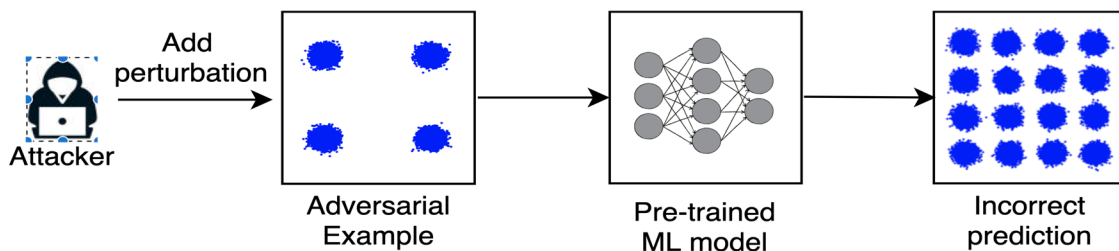


Figure 1a: An example of adversarial attack in wireless communications. An adversary adds perturbation to a test sample to create an adversarial example. Then, the pre-trained ML model wrongly predicts the test sample

AML attacks for CV and NLP have been extensively studied [14] [16]. A canonical example is the misclassification of a panda image as gibbon because of the addition of a specifically crafted perturbation to the panda image so as to mislead the ML classifier [17]. Similarly, in the wireless domain, adversarial attacks can lead to misclassification of signals. As shown in Figure 1, the pre-trained ML model is a classifier for automatic modulation classification. An attacker launches an adversarial attack on the test data in the inference phase by adding a small perturbation to the original QPSK test data sample to form an adversarial example that causes the misclassification of a QPSK modulated signal as 16-QAM. Data-driven, intelligent solutions empowered by ML/DL are considered as the key enablers in 5G and 6G wireless communications [18] - [21]. As seen in [22], [23], the likelihood of jeopardizing ML/DL models deployed to enhance systems is a challenge that raises concerns for the safety of ML adoption. For instance, adversarial attacks causing incorrect ML/DL decisions may harm cyber physical systems such as in autonomous vehicles [24], [25]. As such, *to benefit from the gains of using ML/DL systems, there is an urgent need to ensure security and robustness of ML-based systems in the presence of adversarial attacks.*

There have been multiple research efforts that aim to survey AML attacks as well as recent advances to detect and mitigate them in specific areas where ML/DL has found applications such as CV and NLP [26], [27]. *Different from the existing reviews and surveys, this focuses on approaches to generate, detect and mitigate AML attacks in the wireless communications domain because of the unique characteristics of wireless communications and their effects on AML attacks.* Such unique properties include the effect of the dynamic nature of the communication channel that introduces path loss [28]–[32], the effect of standardized wireless communication blocks to mitigate errors such as forward error correction (FEC) and hardware impairments [33] as well as access to data and heterogeneity in feature representation of the dataset [34], [35]. These unique properties discussed in section III-D necessitate the review of the current attacks and mitigation methods of AML specifically for wireless communications as the methods presented in other domains such as CV and NLP may not be applicable. For instance, the adversary can directly query an API for CV applications without the need to send perturbations over a channel. The design of DL algorithms for wireless communications must consider the new security issues such as identification and mitigation of adversarial signals [36]. Due to the dynamic properties of RF data, it is a non-trivial task to identify adversarial examples in RF data. DL

methods rely on the premise that the training data represents the distribution of the underlying data generation process such that the trained model generalizes well on test data. However, this premise does not necessarily apply in the adversarial scenarios as these distributions change significantly due to adversarial effects that are not straightforward to anticipate or predict [37]–[40]. With such adversarial attacks, wireless communications systems such as 5G and 6G with design components based on DL will be susceptible to disruptions, performance losses, and eventually failures. Therefore, it is critical to analyze the emerging wireless attacks due to AML, reduce and ultimately eliminate the impact of adversaries employing the AML attack techniques. An in-depth study of AML specifically for wireless communication systems is needed given the unique characteristics and emerging use cases of DL in such systems.

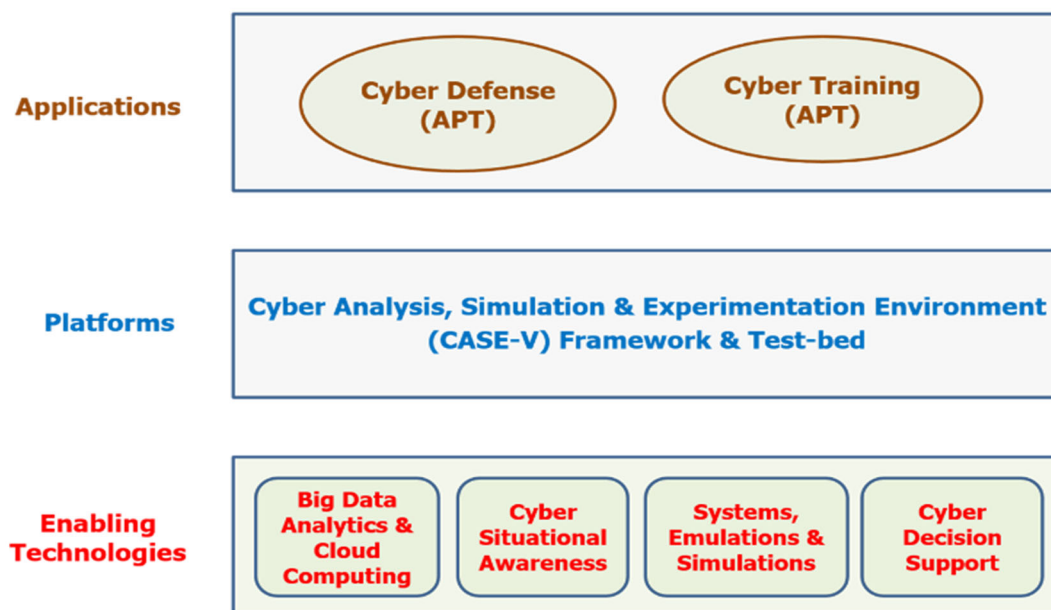


Figure 2. CASE-V architecture

Today's military intelligence analysts are faced with the monumental and escalating task of handling massive volumes of complex data from multiple sources. This includes sensor data, mobile social network data, surveillance data (such as images and videos from UAVs or satellites), and public domain data. An example scenario is given in Figure 1. The data must be aggregated, evaluated, correlated, and ultimately used to support a commander's time-critical decisions and actions. However, currently there is lack of capability to process huge volume of data from heterogeneous sources in military operations [1, 2]. There exist many challenges such as (1) A real-time computing platform for military big data where massive amounts of data are distributed across locations need to be designed and optimized; (2) Heterogeneous data have to be aggregated in a hostile environment and properly stored; (3) Distributed situational awareness and decision making need to be accomplished with minimum delay; (4) Massive datasets and sophisticated results must be presented for easy perception by analysts.

In order to establish and enhance the ability for more effective and efficient cybersecurity research, a multidisciplinary team of researchers (PI: George Hsieh, Sachin Shetty, Jose Padilla, Jonathan Graham, Mary Ann Hoppa) from Norfolk State University (NSU), Old Dominion University (ODU) and Virginia Modeling, Analysis and Simulation Center (VMASC). The CoE-

Cyber center address these fundamental challenges and bring together sensing, perception, and decision support for mission-critical applications of the DOD. Specifically, an integrated computing, communication, and information fusion approach has been proposed and developed. Four research thrusts have been carried out: (1) Cyber Intelligence and Cyber Defense: Explore innovative theories, methodologies and techniques in cyber intelligence and cyber defense, and leverage big data analytics (Lead: Sachin Shetty) (2) Cyber Modeling, Simulation, Analysis & Decision Support: Develop foundational modeling, simulation, analysis, intelligent systems and decision support capabilities, with a focus on human behavior modeling, simulation scalability, and on live-virtual-constructive simulation and kinetic simulation system interoperability (Lead: Jose Padilla); (3) Cyber Training, and Intrusion Detection and Response: Develop intrusion detection, mitigation and forensic analysis capabilities to defend against future advanced threats, and develop integrated training capabilities for cyber defense and response (Lead: Jonathan Graham); (4) CASE-V Framework and Testbeds: Develop synthetic framework and testbeds with distributed, virtual and big data-enabled infrastructures capable of interoperability with selected existing large-scale cyber security testbeds (Lead: George Hsieh)

The global cost of cybercrime exceeded \$600 billion in 2017 [1]. In a report published in 2017 [2], the cost of cybercrime was shown to be accelerating (23 percent increase over the previous year) and costing organizations on average \$11.7 million annually. It's understandable why cybersecurity has become critically important to nations and societies.

To help address this critical need in cybersecurity research and experimentation infrastructures, a research project was initiated at Norfolk State University (NSU) in 2015 to develop a cloud-enabled and big data capable Cyber Analysis, Simulation and Experimentation Environment (CASE-V). A fundamental goal of CASE-V is enhancing situational awareness and decision support, especially for countering new advanced persistent threats (APTs). CASE-V, shown in Figure 1 consists of three major architectural layers: applications, platforms, and enabling technologies including cloud computing and big data analytics.

Cloud computing or virtualization infrastructures are needed for investigating large-scale networks and endpoints to model the real world. Big data analytics also are needed for situational awareness and decision support in cybersecurity research and experimentation, especially for intrusion detection, anomaly detection, incident response, and cyber intelligence. Such investiga-

tions require dealing with big data and can benefit significantly from automated tools, machine learning and data analytic techniques.

The Version 1 CASE-V testbed implemented at NSU uses OpenStack [5] and VMware vCloud as the cloud platforms, and Hadoop [6] as the big data analytics platform. Multiple cloud and big data platforms are configured as clusters with their compute, storage, and networking hardware and resources. They are interconnected through a high-speed core network, an access network, a management network, a remote diagnostic network, and an iSCSI network for shared storage. The testbed is initially operational at the platform and networked system levels. Additional capabilities at the platform and applications levels are being expanded.

The CASE-V testbed is designed with a secure, multi-functional, modular and highly available architecture. The main functions provided by the testbed are big data and cloud/virtualization.

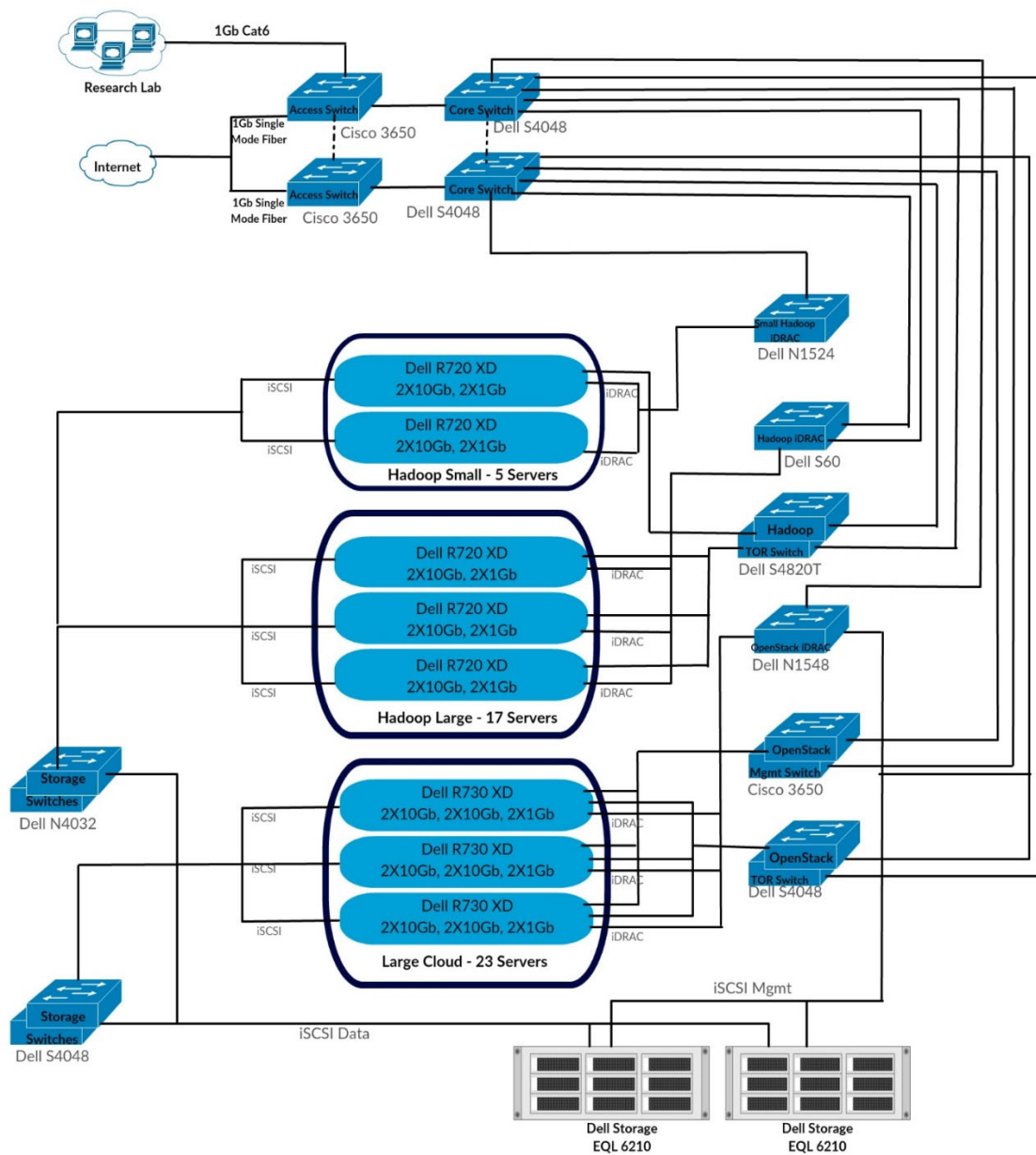


Figure 3: CASE-V testbed connectivity among major components

3.0 METHODS, ASSUMPTIONS, AND PROCEDURES

3.1 Cybersecurity Research and Experimentation Testbed

The CASE-V testbed is designed with a secure, multi-functional, modular and highly available architecture. The main functions provided by the testbed are big data and cloud/virtualization.

The testbed is organized in a modular fashion into clusters such that common functions are grouped together to facilitate management and growth. The clusters are interconnected, when appropriate, to allow communications between them. TABLE 1: CASE-V TESTBED EQUIPMENT is a summary of the major equipment and clusters installed in the CASE-V testbed.

Table 1: CASE-V TESTBED EQUIPMENT

Clusters	Function	Equipment	Qty
Core Infrastructure	Internet Switch	Cisco 3850	2
	Firewall/IPS	Sophos XG-230	2
	Access Switch	Cisco 3650	2
	Core Switch	Dell S4048-ON	2
	iDRAC Switch	Dell N1548	1
		Dell N1524	1
	Fiber Patch Panel	Fiber PP	1
	CAT6 Patch Panel	CAT6 PP	2
	KVM Switch	KVM Switch	2
OpenStack	Server	Dell R730xd	23
	OpenStack ToR Switch	Dell S4048-ON	2
	OpenStack Mgmt Switch	Cisco 3650	2
Hadoop-Large	Master Node	Dell R720xd	5
	Data Node	Dell R720xd	12
	Hadoop ToR Switch	Dell S4820T	2
	Hadoop iDRAC Switch	Dell S60	1
Hadoop-Small	M/D Node	Dell R720xd	5
Shared Storage	iSCSI Disk Arrays	Dell EqualLogic PS6210E	2
	Hadoop iSCSI Switch	Dell N4032	2
	OpenStack iSCSI Switch	Dell S4048-ON	2
Control & Compute	LAN Switch	HP 5130-24AG	2
	Server	Dell R430	6
Cloud-Mid_1	Blade Server	Dell M620 VRTX	1
Cloud-Mid_2	Blade Server	Dell M620 VRTX	1

OpenStack Platform Hardware Components

The OpenStack hardware platform in the CASE-V testbed is composed of the devices and assigned roles listed below. Note the Dell S4048-ON switch is capable of supporting software-defined networking.

Table 2: OpenStack Platform Hardware Components

Role	Device	Qty
MAAS Controller	Dell R430 Server	1
Juju Controller	Dell R430 Server	1
OpenStack Compute Node	Dell R730xd Server	21
OpenStack Network & Storage Node	Dell R730xd Server	1
OpenStack Storage Node	Dell R730xd Server	1
OpenStack ToR Switches	Dell S4048-ON Switch	2
OpenStack Shared Storage Network	Dell S4048-ON Switch	2
OpenStack Shared Storage Array	Dell EqualLogic PS6210E	2
OpenStack iDRAC Switch	Dell N1548 Switch	1
OpenStack Management Switch	Cisco 3650 Switch	2

The computing and storage capacity available for the OpenStack platform is summarized below. Note that each Dell EqualLogic PS6210E storage array has a raw capacity of 96 TB, and 50 percent is allocated for OpenStack shared storage.

Table 3: OpenStack Platform Computing and Storage Capacity

Resource	Capacity	Qty	Total
Server CPU Cores	44	23	1,012
Main RAM Memory	128 GB	23	2,944 GB
Local HDD Storage	9.6 TB	23	220.8 TB
iSCSI Shared Storage Array	48 TB	2	96 TB

High-Availability and Networking

Many components of OpenStack are designed to be highly available. For OpenStack API services, HAProxy [9] was installed and configured to listen on the default service port and to load balance across all service units. To handle hard disk drive failures, RAID 10 was configured on Dell R730xd servers to provide disk mirroring, while RAID 6 was configured on Dell Equal-Logic PS6210E storage arrays to provide protection using double-parity with a small storage penalty. Note there are no backup MAAS or Juju nodes, because they are not needed once the servers and services have been provisioned, deployed, and operational.

OpenStack operation involves several types of network traffic for various purposes, each with different traffic characteristics. A best practice is to use multiple VLANs and IP subnets to isolate the different types of traffic, as follows.

Table 4: OpenStack Platform Network Traffic Types

Network	VLAN	IP Address	Purpose
OpenStack iDRAC	407	10.20.7.0/24	Server management
OpenStack Management	420	10.20.20.0/24	OpenStack management
OpenStack Data Network	450	204.155.188.0/27	Compute traffic between VMs
iSCSI Storage Network	449	10.20.49.0/24	Used by Ceph to access block & object stores

There are six NIC cards (4x10G, 2x1G) on each Dell R730xd server in the OpenStack cluster. Two 10G NICs are used for iSCSI traffic, so they are connected to the two OpenStack iSCSI switches (SW114 and SW115) which are stacked so if one fails the other will take over. The remaining two 10G NICs are used for compute traffic between VMs and are connected to OpenStack TOR switches (SW103 and SW104) which also are stacked. The two 1G NICs are configured for VLAN 420 (10.20.20.2/24) and used for management purpose. Each server also has a built-in 1G iDRAC interface which is connected to the iDRAC switch (SW109) and used for remote access.

VMware vCloud Platforms: Two Dell M620 VRTX blade servers are installed in the CASE-V testbed. Each blade server is equipped with 4 blades, 64 server CPU cores, 1,024 GB RAM, and 30.72 TB shared HDD storage. They are currently running VMware ESXi, vCenter, and vCloud software. To simplify operation, they are scheduled to be migrated to run OpenStack cloud services on top of VMware ESXi and vCenter platforms.

HADOOP BIG DATA PLATFORMS

Apache Hadoop is the most popular open source big data platform software that allows for the distributed processing of large data sets across clusters of computers using simple programming models. Hadoop is designed to scale up from single servers to thousands of machines, each offering local computation and storage.

Hadoop Hardware Components

Two Hadoop clusters were installed and configured in the CASE-V testbed. The Hadoop-Large cluster is composed of 17 nodes, and the Hadoop-Small cluster has 5 nodes. All the nodes use Dell R720xd servers with two different configurations, as shown in the table below.

Table 5: Hadoop Hardware Components

Cluster	Role	#Core	RAM (GB)	HDD (GB)	Qty
Hadoop-Large	Master Node	24	128	3,600	5
	Data Node	24	64	24,576	12
	Total	408	1,408	312,912	
Hadoop-Small	Master Node	24	128	3,600	2
	Data Node	24	128	3,600	3
	Total	120	640	18,000	

The Hadoop clusters use five switches to provide network connectivity for both internal and external communications. Note that the Hadoop clusters can access the shared iSCSI disk arrays through their own backend switches.

Table 6: Hadoop Hardware Switches

Role	Device	Qty
Hadoop ToR Switches	Dell S4820T Switch	2
Hadoop iSCSI Storage Network	Dell N4032 Switch	2
Hadoop iDRAC Switch	Dell S60 Switch	1

Hadoop Software Installation

To simplify Hadoop setup, Cloudera Manager Express (version 5.14.3) [10] was used to install and configure the standard Hadoop packages (e.g., HDFS, Spark, MapReduce, Hive, YARN, HBase) provided by Cloudera Distribution Hadoop (CDH). CentOS (release 7.5) was the Linux OS used on the Hadoop servers.

Hadoop is designed to detect and handle failures at the application layer, so it can deliver a highly available service on top of a cluster of commodity computers. To further enhance the availability of critical services (e.g., NameNode), the master nodes in the Hadoop clusters are

configured with RAID 10 to protect against hard disk failures. Since Hadoop already provides a data replication capability, the slave/data nodes are configured with RAID 0.

zHadoop Networking

The Hadoop clusters are connected to multiple networks for different types of network traffic. Multiple VLANs and IP subnets are configured to isolate the different types of traffic, as follows.

Table 7: Hadoop Hardware Switches

Network	VLAN	IP Address	Purpose
Hadoop iDRAC	407	10.20.7.0/24	Server Management
Hadoop-Large Management	458	204.155.189.0/27	Hadoop-Large management
Hadoop-Large Data Network	426	10.20.26.0/24	Hadoop-Large internal communications
Hadoop-Small Management	459	204.155.189.32/28	Hadoop-Small management
Hadoop-Small Data Network	425	10.20.25.0/24	Hadoop-Small internal communications
iSCSI Storage Network	449	10.20.49.0/24	Access shared iSCSI storage

Note how each Hadoop and OpenStack cluster is assigned two separate VLANs and IP subnets: one for the data network and the other for the management network. This segments the network traffic on a per-cluster basis. On the other hand, the iDRAC or iSCSI network for all Hadoop and OpenStack clusters are assigned the same VLAN and IP subnet to facilitate information sharing and access control on the system basis.

To provide a flexible, feature-rich and user-friendly cybersecurity research and experimentation environment, an automation and UI layer is needed on top of the foundational platforms and their ecosystems. Such a framework is not intended to duplicate natively provided features, but rather to enhance the CASE-V testbed's responsiveness, security, scalability, and support for user-defined projects, diverse platforms, programming languages and API's

The MEAN stack was chosen as the underlying technology for the web-based user interface framework. MEAN is a free and open-source JavaScript software stack used for creating dynamic websites and web applications [6]. It is composed of four components: (1) MongoDB, a NoSQL database; (2) Express.js, a minimal and flexible Node.js web application framework; (3) AngularJS, which extends the HTML vocabulary for declaring dynamic views in web applications; and (4) Node.js, a cross-platform JavaScript run-time environment that executes JavaScript code server-side. In addition to the MEAN stack, Bootstrap [7] was chosen as the front-end component library for the CASE-V testbed. Bootstrap is a free and open-source front-end framework for faster and easier web development.

Figure 3 shows the conceptual architecture of the Automation and UI framework for the CASE-V testbed. A prototype of the framework was implemented on Linux workstations. Python and Bash scripts were developed, along with code for the MEAN Stack, to implement a limited set of capabilities, such as adding a new user/project to the CASE-V testbed, and automatically configuring the CASE-V private OpenStack cloud for this purpose (i.e., without requiring the user or administrator to perform the relevant native OpenStack commands).

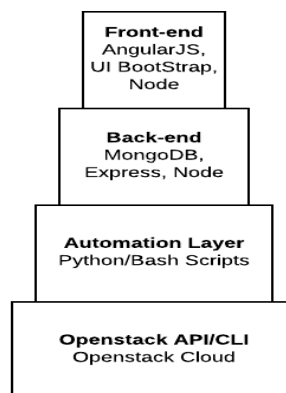


Figure 4: CASE-V automation and UI architecture.

Figure 4 illustrates the steps and components involved in registering and creating a new user for the CASE-V testbed.

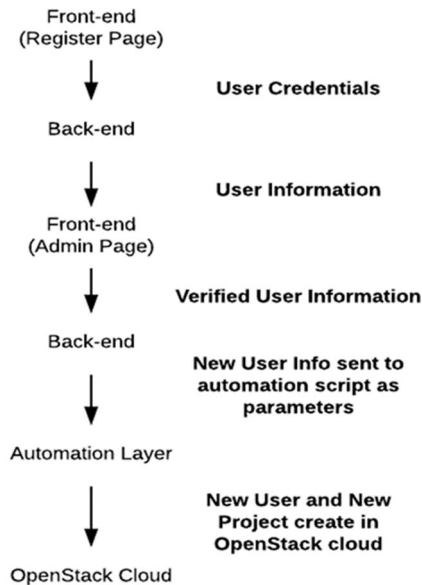


Figure 5: Register and create a new user.

The prototype implementation of the Automation and UI framework is being ported to the Version 1 CASE-V testbed, and ongoing enhancements are underway.

USER-DEFINED PROJECTS

The Version 1 CASE-V testbed does not provide any custom-developed tools for creating, managing and operating user-defined research, experimentation, and training projects. Authorized users need to handle these tasks using the common tools provided by the platforms and their ecosystems. They also may need to use different systems for developing the projects before they are deployed, tested, and operational on the CASE-V testbed.

For example, there are several cybersecurity research and experimentation projects previously developed at NSU on Windows, Linux, VirtualBox, VMware and XenServer platforms for activities such as malware classification [12], cyber intelligence analysis using Twitter data and Elastic Stack [13], and Gh0st APT training [14] [15]. There also is a virtual environment for enterprise APT experimentation [16]. These projects are being migrated and deployed onto the CASE-V testbed as use cases and to support feasibility, performance and scalability assessments.

3.2 Cybersecurity Threat Intelligence Collection and Analysis

Twitter Streaming API

The TwitterOSINT system utilized the Twitter Streaming API [1], as illustrated in Figure 1 [2], which provides low-latency access to Twitter's global stream of tweets and download those that were considered relevant to cybersecurity.

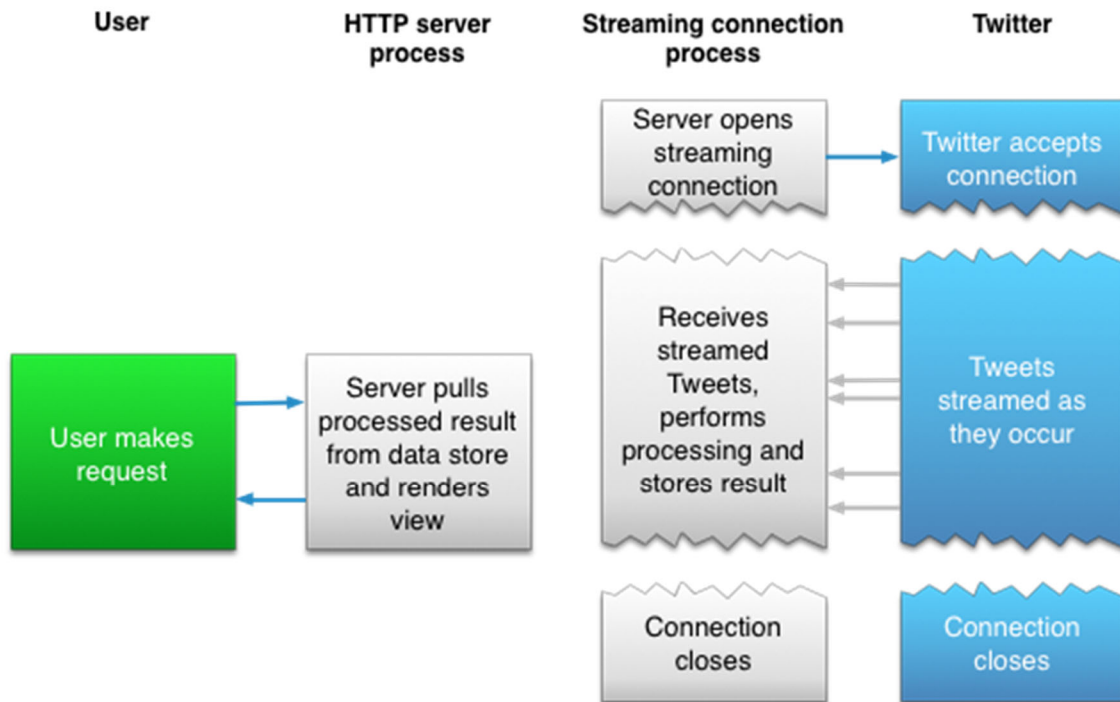


Figure 6: Twitter Streaming API [2].

A streaming client pushes messages indicating tweets and other events have occurred based on user-specified keywords. Connecting to the streaming API requires keeping a persistent HTTP connection open. The code for maintaining the streaming connection is typically run in a process separate from the process that handles HTTP requests. The streaming process gets input tweets and performs any parsing, filtering, and/or aggregation needed before passing the results to a data store. The HTTP handling process queries the data store for results in response to user requests.

Natural Language Processing

According to industry estimates, only 21% of the available data is present in structured form [3]. In order to produce significant and actionable insights from the unstructured text data, Natural Language Processing techniques and tools are developed to analyze, understand, and derive information from the text data in an intelligent and efficient manner. The field of NLP aims to convert human language into a formal representation that is easy for computers to manipulate. NLP has been used to perform numerous automated tasks in a wide variety of applications including machine translation, automatic summarization, named entity recognition, sentiment analysis, etc. [3].

While complete semantic understanding by machine is still a distant goal, researchers have taken a divide and conquer approach and identified several NLP sub-tasks useful for application development and analysis. These range from the syntactic – such as Part-of-Speech (POS) tagging, chunking and parsing – to the semantic – such as word sense disambiguation, semantic-role labeling, named entity extraction and anaphora resolution [4].

Two problems in particular make NLP difficult and require techniques different from those associated with processing artificial languages (e.g., compiler construction). These problems are: the level of ambiguity that exists in natural languages; and the complexity of semantic information contained in even simple sentences. Typically, language processors deal with a large number of words, many of which have alternative usages, and large grammars that allow different phrase types to be formed from the same string of words. NLP is made more complex because of the irregularities of language and the different kinds of ambiguity that can occur [5].

NLP Tasks

The most common tasks in the NLP pipeline, as shown in **Figure 2** [6], are:

- POS Tagging (POS): labels each word with a unique tag that indicates its syntactic role (e.g., plural noun, adverb).
- Chunking: labels segments of a sentence with syntactic constituents such as noun phrase (NP) or verb phrase (VP). Each word is assigned only one unique tag.
- Named Entity Recognition (NER): labels atomic elements in the sentence using categories such as PERSON, COMPANY, or LOCATION.

- Semantic Role Labeling (SRL): assigns a semantic role to each syntactic constituent of a sentence.
- Language Models: traditionally estimate the probability of a word w being the next word in a sequence.
- Semantically Related Words (e.g., synonyms, homonyms, hypernyms): are measured using the WordNet databases as ground truth. This task involves predicting whether two words are semantically related.

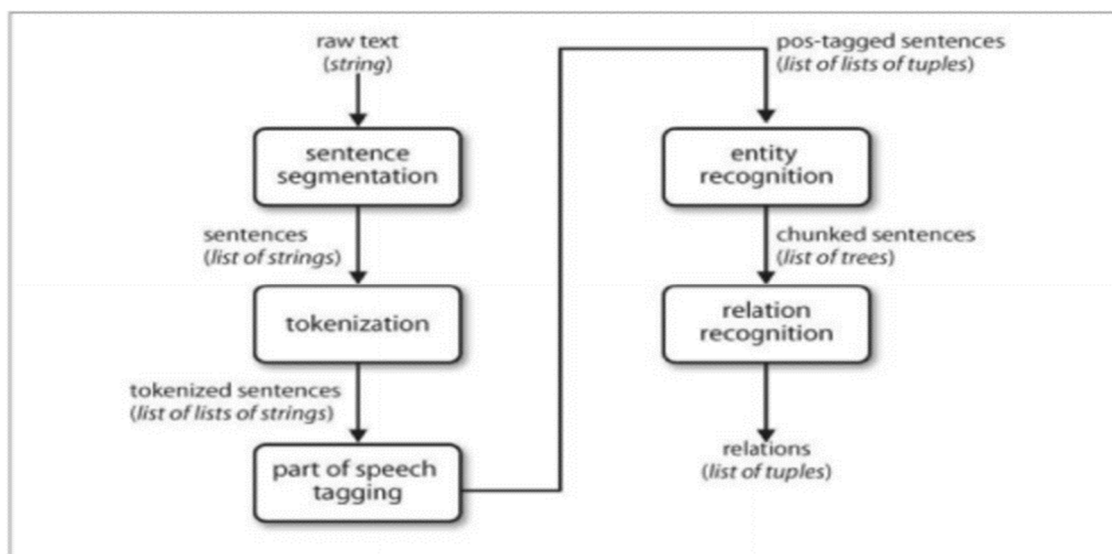


Figure 7: NLP Pipeline [6].

Stanford CoreNLP

The Stanford CoreNLP [7] is one of the most widely used NLP software worldwide. It provides a set of human language technology tools. It can give the base forms of words, their parts of speech, whether they are names of companies, people, etc., normalize dates, times, and numeric quantities, mark up the structure of sentences in terms of phrases and syntactic dependencies, indicate which noun phrases refer to the same entities, indicate sentiment, extract particular or open-class relations between entity mentions, get the quotes people said, etc. [7].

The Stanford CoreNLP is designed to make it very flexible, extensible, and easy to apply a collection of linguistic analysis tools to a piece of text. A tool pipeline can be run on a piece of plain text with just two lines of code. Stanford CoreNLP integrates many of Stanford's NLP tools, including the part-of-speech (POS) tagger, the named entity recognizer (NER), the parser,

the coreference resolution system, sentiment analysis, bootstrapped pattern learning, and the open information extraction tools. Moreover, an annotator pipeline can include additional custom or third-party annotators. CoreNLP's analyses provide the foundational building blocks for higher-level and domain-specific text understanding applications [7].

Entity Extraction for Cyber Domain

Although the Stanford CoreNLP is a state-of-the-art NLP software, it did not provide very good performance for entity extraction from unstructured text in the cyber domain because it was not trained with cyber domain-specific data.

The Stucco Entity-Extractor library, developed by ORNL, was designed to identify and label cyber-domain entities from unstructured text [8] [9] [10]. This library makes use of the Apache OpenNLP [11] library, and the Stanford CoreNLP library to tokenize, part-of-speech tag, and build the parse trees of the unstructured text.

The entity types identified and labelled by the Stucco Entity-Extractor library include [10]:

- Software: Vendor, Product, Version
- File: Name
- Function: Name
- Vulnerability: Name, Description, CVE, MS

For an unstructured text, the Entity-Extractor library produces an Annotation object that represents the unstructured text as a map which includes:

- Text: original raw text
- Sentences: list of sentences
 - Sentence: map representing one sentence
 - Token: word within the sentence
 - POSTag: part-of-speech tag
 - CyberEntity: cyber domain label for the token
 - ParseTree: sentence structure as a tree

As an example, Figure 3 [8] shows a part of the National Vulnerability Database (NVD) text description of the Common Vulnerability Enumeration CVE-2012-0678 with automatically generated labels. Note the cyber-domain annotations such as Software Vendor: Apple and Software Product: Safari.

B:Relevant Term *I:Relevant Term* *B:Relevant Term* *O* *O* *B: Software Vendor*
 Cross-site scripting (XSS) vulnerability in Apple
B: Software Product *B: Version* *I: Version* *B:Relevant Term* *I:Relevant Term* *I:Relevant Term*
 Safari before 6.0 allows remote attackers
B:Relevant Term *B:Relevant Term* *I:Relevant Term* *I:Relevant Term* *B:Relevant Term*
O to inject arbitrary web script *O* or HTML
O *O* *O* *O* *B:Relevant Term*
 via a feed :// URL.

Figure 7: Example text with automatically generated labels [8].

After a tweet was analyzed by the NLP pipeline, the TwitterOSINT system appended the tweet’s text along with its user name, date, and all the annotations to a JSON file, ready for further processing and analysis.

Elastic Stack

The TwitterOSINT system utilized the Elastic Stack [12] which is a scalable, flexible, and feature rich platform enabling users to reliably and securely take data from any source in any format and search, analyze, and visualize it in real time.

Figure 4 shows the key components of the Elastic Stack:

- Logstash: an open source, server-side data processing pipeline that ingests data from a multitude of sources simultaneously, transforms it, and then sends it to a “stash” (which is Elasticsearch if using the Elastic Stack).

- Elasticsearch: a distributed, RESTful search and analytics engine for indexing, storing, and analyzing data.
- Kibana: UI to visualize Elasticsearch data and navigate the Elastic Stack.

The Elastic Stack

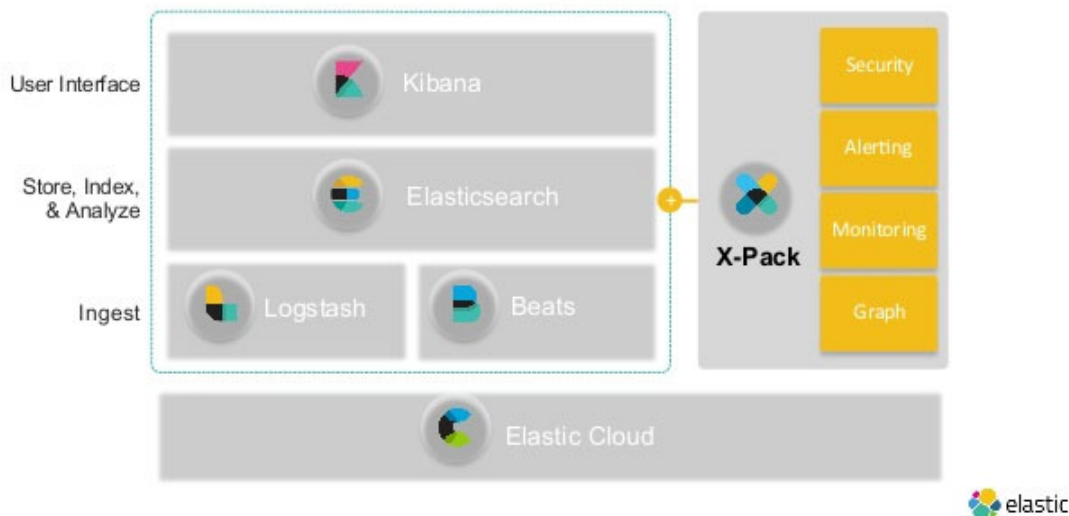


Figure 8: Key components of the Elastic Stack.

The TwitterOSINT system used the Logstash to ingest the JSON file created for the selected tweets and feed the content to Elasticsearch which then indexed and stored the data into the centrally managed data store. To search and analyze the indexed tweets to gain further insights and derive actionable intelligence, a human analyst can use Elasticsearch and Kibana to execute search queries or create visualizations. The X-Pack for the Elastic Stack can be further utilized to provide additional data analytic and task automation capabilities including machine learning, graph analytics, and alerting.

In this section, we describe the design and implementation of the TwitterOSINT prototype system.

System Architecture

TwitterOSINT brought together NLP, machine learning techniques and the Elastic Stack to provide a comprehensive data analytic environment to collect, store, analyze, and visualize tweets that were considered relevant to cybersecurity. The flowchart in Figure 10 at the end of the paper shows the architecture of the TwitterOSINT system and the transformation of Twitter data at each step.

Development Environments and Tools

The TwitterOSINT prototype system was developed using Eclipse Oxygen Release 4.7.0 IDE, and Java Version 8 Update 151. It was initially implemented on a Dell Power Edge 730xd enterprise-grade server running CentOS 7. It was later ported onto Windows 7 workstations. The system was built with the following major application-specific dependencies:

- twitter4j-4.0.4
- stanford-corenlp-3.4.1
- stanford-corenlp-3.4.1-models
- entity-extractor-1.0.0
- opennlp-tools-1.6.0
- jackson-databind-2.7.0

TwitterOSINT Java Application

TwitterOSINT was the main Java application comprised of two main class objects: Stream and annotator.

- **Stream.java** contains the main class that streams the tweets using the Twitter Streaming API, based on the cybersecurity keywords specified in a text file. It removes special characters and URLs and then passes the result onto the annotator class for further NLP and entity extraction. The Stream class is also responsible for creating JSON objects with the tweet date, username, text and annotations returned by the NLP and Entity Extraction pipeline. These JSON objects are next written into a file named twitDB.json.
- **Annotator.java** contains the class that uses the entity-extraction module for tokenizing, POS tagging and adding annotations to each token. The output from the annotator class is a JSON

object with key-value pairs, where keys are the field labels (e.g. token, POS, CyberAnnot) and values are the field values.

3.3. Adversarial Machine Learning in Wireless Communications using RF Data

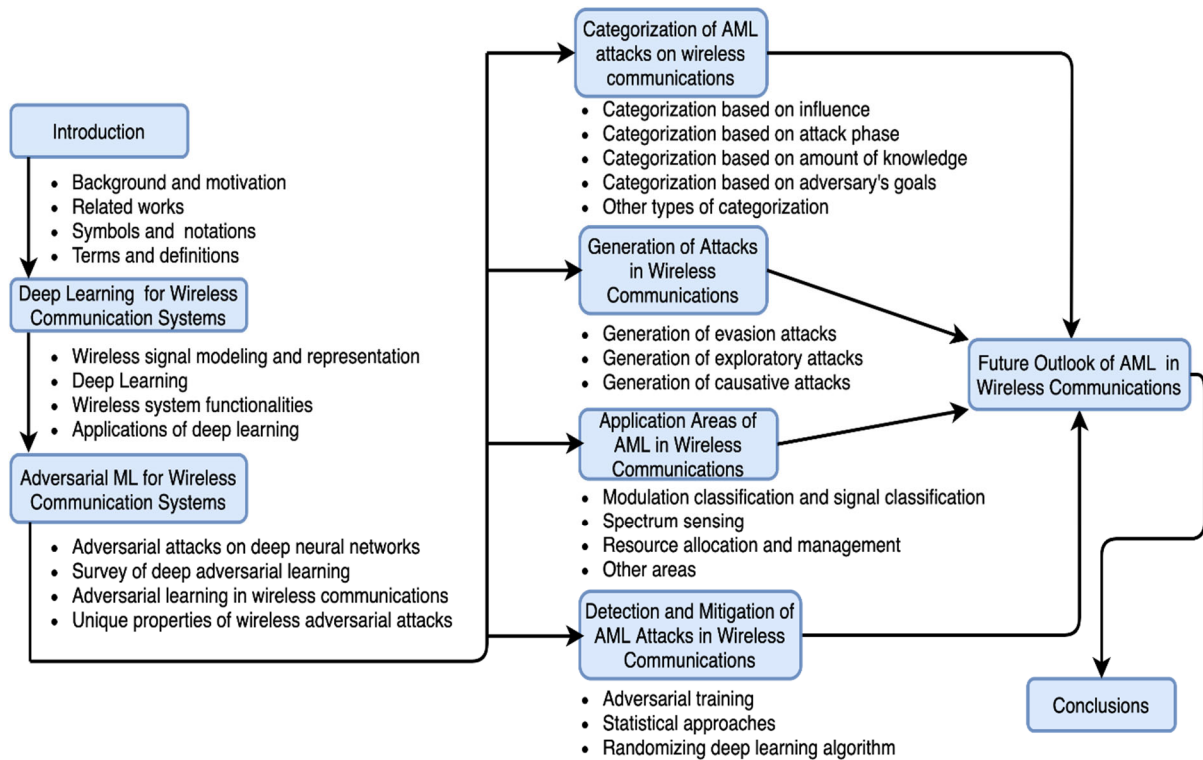


Figure 9: Overview of the arrangement of the materials in this review.

Due to the unique properties of wireless communications, it is essential to understand and quantify the impact of AML attacks on wireless communications and develop defense mechanisms that are tailored for the wireless domain

There is a lot of interest in developing artificial intelligence for wireless communication systems using radio frequency (RF) data. This is because the ability to learn and intelligently respond to dynamic and complex operating conditions will be of utmost importance in future wireless systems. It is envisioned that the knowledge of current operating conditions and en-

environment leveraging on wireless big data analytics will allow communication systems to make the best opportunistic decisions [91]–[94]. To achieve this, work is done to develop efficient algorithms and methods to extract meaningful information from complex and massive datasets from communication systems. Research in DL for wireless communications have gained tremendous attention in recent times, particularly the application of machine learning and deep learning to solve wireless communication problems using RF data. In this section, we present the foundation of wireless signal model, the deep neural network and the application of DL in wireless communication systems.

A. Wireless Signal Modeling and Representation

Traditional wireless communication systems are primarily comprised of a transmitter, a channel and a receiver. A block-based design principle is pursued to split the communication systems into multiple independent blocks. The functional blocks that are typically optimized with the aid of mathematical model and expert knowledge include source encoder (which compresses the input data and removes redundancy), channel encoder (which ensures the system can cope with the effects of the channel by adding redundancy to the source encoder output) and the modulator (which provides a mapping based on the desired received signal level and data rate) [1], [95]–[97]. The sequence of source information bits b_k is transformed by coding to a binary sequence mapped to symbols s_k through modulation as shown in Figure 3. The modulated signal is passed through a pulse filter. The resulting signal is reconstructed to a continuous analog time signal $s_b(t)$ by passing it through a digital-to-analog (D/A) converter. $s_b(t)$ is the baseband signal shifted by the carrier frequency f_c

to obtain the bandpass signal $s(t) = \text{Re} \{ s_b(t)e^{j2\pi f_c t} \}$ [97]

expressed as:

$$s(t) = \text{Re} \{ s_b(t) \} \cos(2\pi f_c t) - \text{Im} \{ s_b(t) \} \sin(2\pi f_c t). \quad (1)$$

At the channel, the transmitted signal encounters impairments such as timing drift, frequency offset, noise. The variations in the channel can be modeled as large-scale fading and small-scale fading. The received signal $r(t)$ is the corrupted form of the signal transmitted due to hardware imperfections and channel impairments. The effect of the channel is modeled as a finite impulse response (FIR) filter $h(t, \tau)$. Therefore the received signal $r(t)$ is the convo-

lution of the signal $s(t)$ and the channel $h(t, \tau)$ given by $s(t) \otimes h(t, \tau)$. $r(t) = r_b(t)e^{2\pi j f_c t}$, where $r_b(t)$ is the baseband complex envelope described.

The received signal $r(t)$ is passed through an amplifier and low-pass filter. The resulting continuous-time signal is converted to its digital version by an analog to digital converter. This is achieved by sampling it at rate $f_s = \frac{1}{T_s}$ samples per second, giving us a discrete representation of the signal made up of the in-phase and the quadrature components and representing the complex raw time-series signals r_k . The signal is demodulated and passed through the decoder to obtain the reconstructed information [97].

Deep Learning

Deep learning can automatically extract relevant features from data with no need for hand-engineering, thus making it more beneficial in most cases than traditional machine learning algorithms. The main types of deep learning models are:

Deep Neural Network (DNN): a type of artificial neural network (ANN) with multiple hidden layers that can extract and transform features automatically [98]. It is also considered a universal function approximator that can identify inherent patterns and discover linear and non-linear associations between features. It is computationally and memory intensive for training due to its large number of parameters. However, recent advances in the field, such as the development of tensor processing units (TPUs), have reduced computation time. A single layer of a neural network is formulated as $y = g(w^T x + b)$ where y is the output of the layer, w are the weights, x is the input data, b is a bias term and g is the activation function.

Convolutional Neural Network (CNN): a type of feed-forward neural network known to perform convolution operations described in Equation 3 for a multi-dimensional data X as input and a multi-dimensional filter K . The layers identify local patterns by using the same geometric transformations (filters) to various spatial locations in the input data. The convolution operation is achieved by taking the dot product of the filter K with dimension m by n and the input data across the input space to get a feature map S of dimension i by j as described in Equation 3. CNNs are well suited for a dataset with a grid-like topology such as images and time-series data. They leverage parameter sharing and sparse interactions to improve efficiency and reduce the memory requirement.

$$S(i, j) = (K * X)(i, j) = \sum_m \sum_n X(i - m, j - n) K(m, n) \quad (3)$$

Recurrent Neural Network (RNN): a type of artificial neural network optimized for sequential input tasks, such as speech recognition, as past observations influence its output. This is because it can learn from long-term dependencies and remember past observations. The dependence of the current output on past output indicates that RNN shares the same weights across several time steps. The dependency is in the form of a feedback loop connected to the output from the past, which distinguishes it from other types of neural networks [98].

DL is primarily using a neural network with multiple layers that perform non-linear transformations on data to provide some output. For example, supervised learning using a DNN is expressed as a mapping f that models the mapping from the input data to a class (label). To learn this input-output relationship, a weight matrix w is estimated. A loss function $l(x, y, w)$ is computed as a point-wise measure of error between the model prediction $\hat{y} = f(x)$ and the observed ground truth y , where $x \in X$ is a sample data point in data X and y is a class in $\mathcal{Y}$. To estimate w , a cost function $J(w)$, which is the average loss over all (n) training data samples in X , is computed

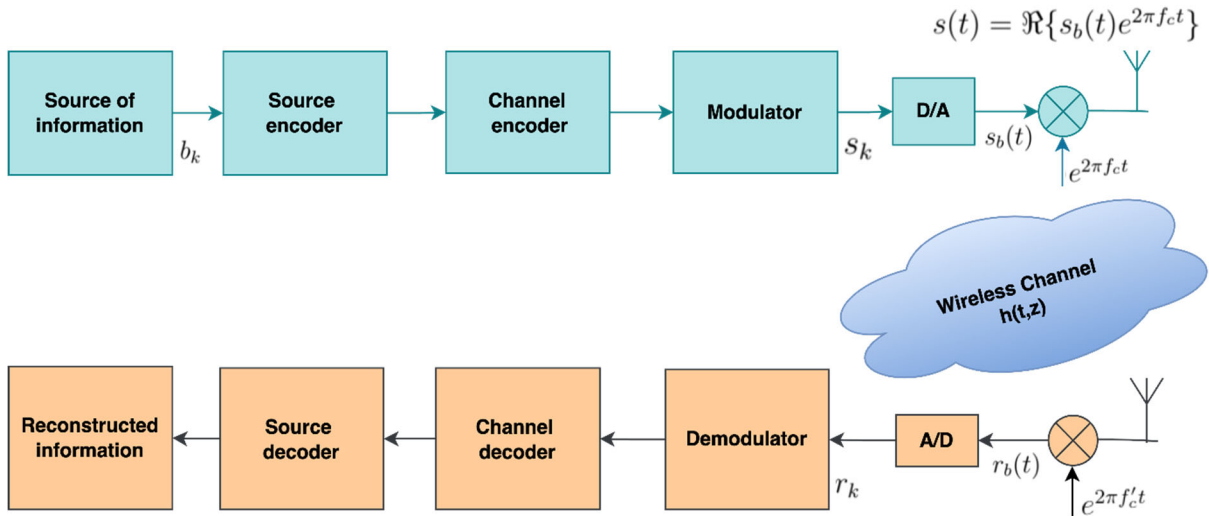


Figure 10: Communication system diagram.

Wireless System Functionalities and Applications of Deep Learning

DL has been extensively used in wireless communications as extensively surveyed before [1] - [3], [92], [99]–[106]. These previous studies highlight various models and applications of DL in wireless communication systems some of which are shown in Figure 4. The authors in [1] gave an overview of deep learning for wireless communications focusing on physical layer communication, spectrum situation awareness and wireless security in deep learning applications. We discuss some of the existing works in various application areas including modulation classification, spectrum sensing, resource allocation and protocol design.

Modulation and signal classification: DL has been applied to identify the modulation type or signal in received signals. The modulation classification problem is an extensively studied problem in the wireless communication domain using DL. While modulation classification seeks to recognize the modulation scheme of a detected signal to aid correct demodulation of the received signal, a signal classification algorithm can help in spectrum monitoring, spectrum management, and secure communications. Numerous approaches have been used to solve this problem using DL.

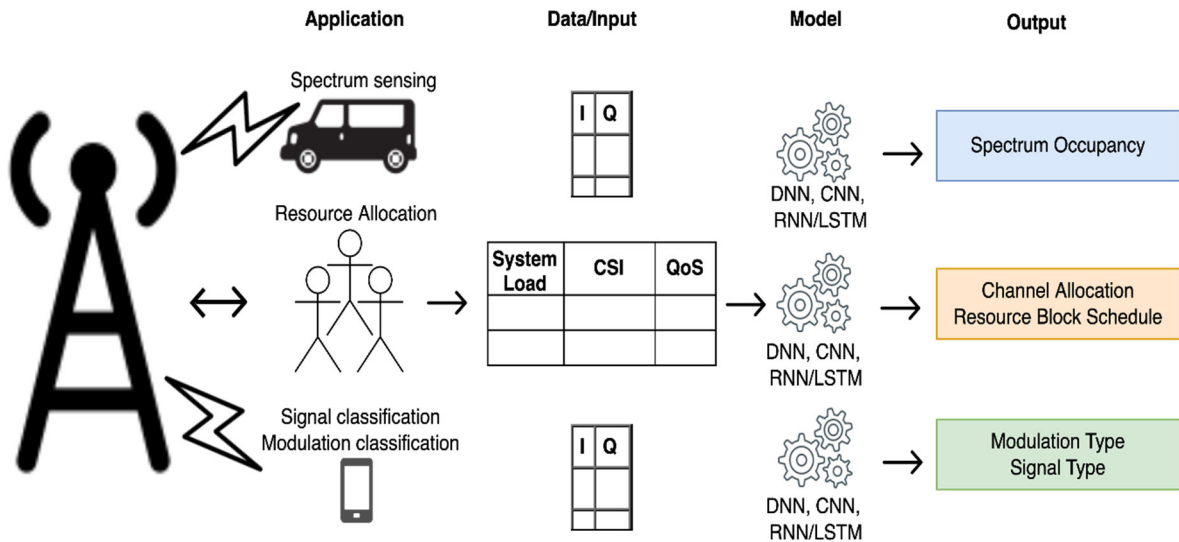


Figure 11: System model for the applications of DL in wireless communications

Table 8: DEEP LEARNING APPLIED TO WIRELESS SYSTEM FUNCTIONALITIES

System functions	Input	Output	DL Models Applied
Spectrum sensing	I/Q signal	Spectrum occupancy	DNN, CNN, RNN/LSTM
Signal classification	I/Q signal	Type of signals or users	DNN, CNN, RNN/LSTM
Modulation classification	I/Q signal	Type of modulations(e.g., QPSK, QAM, etc.)	DNN, CNN, RNN/LSTM
Resource allocation	Channel state information, system load, users QoS requirements, etc.	Channel allocation, resource block schedule, etc.	DNN, CNN, RNN/LSTM

Spectrum sensing: The objective of spectrum sensing is to detect the availability of frequency channels in dynamic spectrum access enabling technology such as cognitive radios. In recent times, machine learning-based sensing has been proposed in previous works [115]–[121]. The spectrum sensing application is typically developed as a classification problem that predicts if the channel is occupied or free. Figure 4 shows spectrum sensing in a vehicular network. Spectrum sensing applications have been surveyed in [122], [123].

Resource allocation and management: In wireless communications, resource allocation and management such as beamformer design and power control plays a central role in the efficient operation of the communication system. With the significant increase in the interaction of different end users, communication technologies, and network operators as indicated in Figure 4 the use efficient resource allocation and management techniques such as DL-based approaches becomes inevitable. Before now, numerical optimization algorithms which are often complex are used for resource allocation and management. Numerous DL-based algorithms are being developed to make resource allocation more efficient [124]–[129]. A large dataset is collected to train and develop DL models for various resource allocation and management problems. The work by [127] assigned resources to non-real time service in advance by learning the resource allocation pattern in a prediction window while DL is used to solve sub-band and power allocation problem

in a multi-cell network [130]. The works in [126], [131] first learn the resource allocation schemes and prompt the appropriate action policy a priori to avoid and alleviate the congestion in an intelligent fashion. Resource allocation to enable coexistence of transceiver pairs and Wi-Fi users in LTE-U networks is proposed in [132].

Adversarial Machine Learning for Wireless Communication Systems

In this section, we discuss the deep adversarial learning starting with a brief introduction on adversarial attacks on deep neural networks and how adversarial attacks work. Then a survey of deep adversarial learning in other domains such as computer vision, graphs, and natural language processing is presented. Also, adversarial learning in wireless communications and how the unique properties of wireless communication systems affect adversarial attacks are discussed.

Adversarial Attack on Deep Neural Network

As an example of AML attacks, we describe next how the evasion or the so-called adversarial attack works. We will discuss other types of AML attacks in next. For every $x \in X$, $f_w(x)$ assigns a label of the form $\hat{y} = \underset{y \in \mathcal{C}}{\arg \max} \|f_w(x, y)\|$, where $f_w(x, y)$ corresponds to the output of f_w for class y in $\mathcal{C}$. An adversarial attack on the classifier f_w aims to modify the classified input x with perturbation r_x such that $f_w(x + r_x) \neq f_w(x)$. For the stealth nature of this attack, it is required that the perturbation r_x should be small such that the difference of $x + r_x$ from the unperturbed input data x remains small. Typically, this requirement is relaxed to constraint $\|r_x\|_p \leq c$, where c is a small positive number and $\|\cdot\|_p$ represents the l_p norm with $p = 1, 2$, denoting the type of norm. The adversarial perturbation r_x for x and f_w is determined by solving the following optimization problem:

$$\begin{aligned} \min_{r_x} \quad & \|r_x\|_p \\ \text{s.t.} \quad & f_w(x) \neq f_w(x + r_x). \end{aligned} \tag{6}$$

In the context of wireless communications, the l_2 -norm is a natural choice for r_x as it accounts for the perturbation (signal) power [139] [90] compared to other norms such as the l_1 -norm that has been used in the CV domain to deter the variation from human perception [90].

Survey of Deep Adversarial Learning

DL has found numerous applications in various areas and has provided major advancements to solving problems that have appeared unsolvable in the past. DL has achieved unprecedented results at scale in solving hard scientific problems. Although DL has achieved remarkable success in various areas, DL is susceptible to attacks such as small perturbations which can significantly change the prediction of the DL algorithm as described. Next, we discuss the recent work in deep adversarial learning under computer vision, graphs and natural language processing.

Adversarial Learning in Wireless Communications

The application of DL to the wireless communications domain differs significantly from other application areas such as CV and NLP because specialized domain knowledge of wireless communications is needed for proper data representation, preprocessing, and result interpretation. In addition, channel, interference, and traffic effects shape both legitimate communications and others. These properties make the RF data unique and possibly have a different response to the adversarial attacks. Some unique properties of RF data are identified and discussed in relation to adversarial example generation and its effect. Various areas in wireless communications where AML occur, and various DL models used. These works establish the detrimental effect of adversarial attacks on wireless communication systems. The different types of attacks are categorized as detailed summarize the previous work on AML attacks in wireless communication detailing the attack type, categorization of attack, a description of the attack and the relevant publications. The body of knowledge on adversarial learning in wireless communications is explored and that include the uniqueness of wireless adversarial attacks, categorization of AML attacks, attack generation methods, some application areas and AML attack detection and proposed mitigation schemes.

Unique Properties of Wireless Adversarial Attacks

DL has a strong potential to assist network decisions in achieving optimal resource management and future networks are envisioned to rely on DL-driven agents for functions such as network management automation and optimization of radio interfaces. The evaluation of the extent to which an adversary can affect such a complex system becomes critical in designing future communication systems with DL. This evaluation becomes critical as there are unique properties in the wireless communication domain as discussed in this section.

- *Effect of the communication channel:* The wireless communication channel has a major impact on adversarial attacks as they need to be launched over the air to reach the target receivers that host the victim DL models. One aspect is that the channel will impose path loss and phase on adversarial perturbations and could weaken and/or change the direction of adversarial signals as they travel through wireless channels before reaching the target receivers. For an adversarial attack to succeed over the air, the adversary needs to account for the dynamic nature of channels when crafting the adversarial perturbations [28]–[32]. In addition, the adversary may not have access to the DL model at the target receiver and the process of data gathering by the adversary for training a surrogate model is typically performed through a channel, which implies that the training data used by the adversary is imperfect by default and the effectiveness of the surrogate model trained by the adversary strongly depends on channel effects [59]. Furthermore, it has been shown in [235] that adversarial attacks crafted to fool DL models trained on time-domain features do not necessarily transfer to DL models trained using frequency-domain features. By leveraging differences in channels, it is also viable to seek the objective of correct signal classification at one receiver while fooling the signal classifier at another receiver [28], [31], [189], [236], [237].
- *Exploitation and mitigation of channel effects:* The adversarial perturbation may aim to fool a wireless signal classifier at one receiver while the perturbed signal can be still decoded at the intended receiver with minimal loss of reliability [236]. This paradigm can be used to hide wireless signals (ranging from simple modulated signals to complex 5G signals) from eavesdroppers [189], [238]. While balancing the objectives of covertness and communication performance, the adversary may aim to preserve not only the bit error rate (BER) as in [31] but also the spectral shape of the perturbed signal as in [58]. In addition, forward error correction (FEC) - a building block in communication systems to detect and correct errors encountered due to noise, channel effects, and hardware impairments - may reduce the effect of adversarial attack while maintaining communication performance with the intended receiver [33]. On the other hand, the adversary may use multiple antennas to transmit perturbations to increase the effect of adversarial attacks [30].
- *Indirect influence on training and test data:* A wireless adversary cannot directly manipulate the training or testing data input to a classifier. It also cannot directly query a transmitter's classifier and obtain its classification results as opposed to the typical CV and NLP

applications that often rely on API queries. In an over-the-air attack, the wireless adversary needs to monitor the actions in wireless communications and indirectly try to manipulate/influence the outcome of the DL model or initiate actions such as jamming the channel during sensing and data transmissions [34].

- *Heterogeneity in feature representation:* The coexistence of various communication systems introduces heterogeneity as a more diverse and complex feature representation in radio data [35]. This significantly affects the effectiveness of crafted perturbations. In this context, the adversary may also aim to fool the wireless signal classifier at multiple receivers (each possibly belonging to a different network) by transmitting a signal perturbation and relying on the broadcast nature of omnidirectional transmissions to reach and jointly affect multiple classifiers. Then, the perturbation needs to be determined by accounting for multiple channels to different receivers [28]. In addition to dynamic channels, it is also possible that network traffic is of stochastic nature or traffic needs to be adapted to dynamic channel effects. Then, the generation of perturbations needs to be jointly considered with queue stability with the extended goal for the adversary to reduce the stable throughput (the maximum achievable throughput while keeping the packet queues stable [239]).

CATEGORIZATION OF ADVERSARIAL MACHINE LEARNING ATTACKS ON WIRELESS COMMUNICATIONS

We present a categorization of AML attacks to aid the understanding of key concepts of AML and provide an overview of AML attacks from various perspectives borrowed from [29], [35], [51]–[53], [55], [244], [245] for wireless communications and [23], [65], [67]–[69], [86], [246], [247] for other domains, especially CV. In particular, we discuss adversarial attacks in wireless communications under three main categories (see Fig. 5) that are defined according to the influence - the type of attack, the attack phase, and the amount of knowledge the adversary has about the victim model. Table IV lists previous works of AML attacks on wireless communications under these categories.

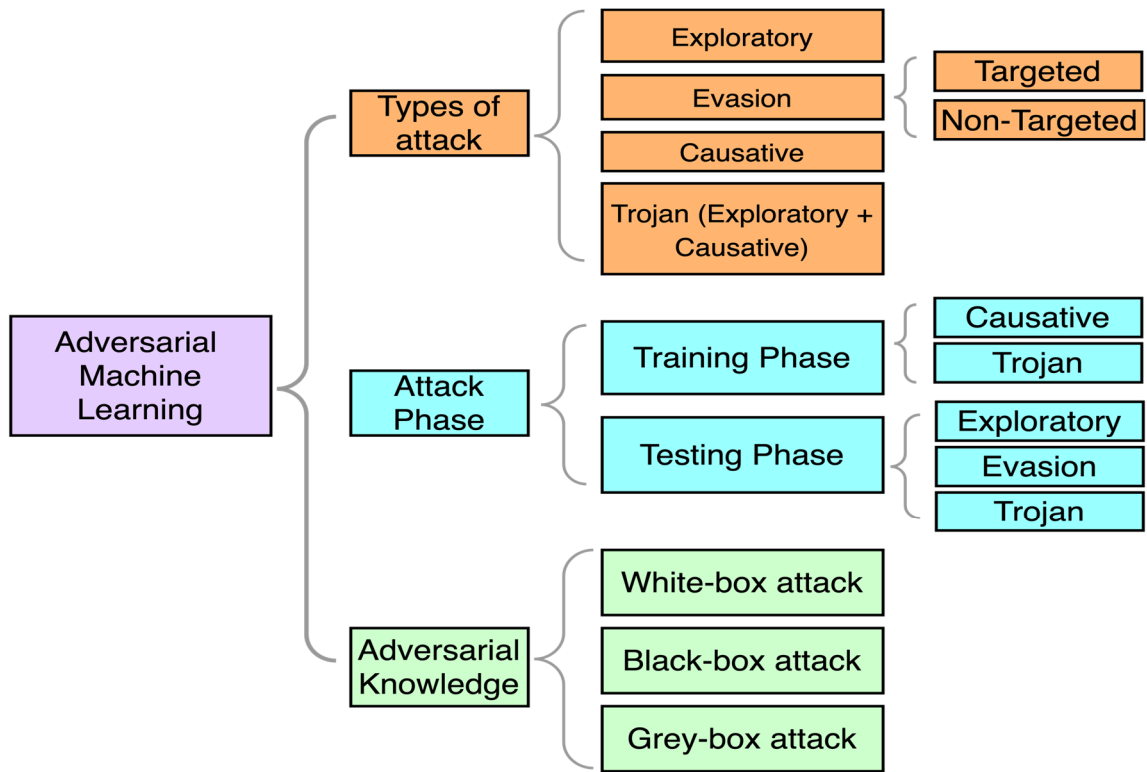


Figure 12: Categorization of attacks based on adversarial machine learning.

Categorization Based on Influence - Types of Attack

Exploratory attacks: Also called inference attacks, exploratory attacks seek to fathom the inner workings of ML algorithms/models by collecting training data and imitating the ML model functionally with similar types of inputs and outputs, namely building a surrogate (shadow) model [248], [249]. The exploratory attack is typically the leading step prior to subsequent attacks as it aims to “explore” the victim model using techniques such as active learning [250], [251] or augment the limited information with generative adversarial networks (GANs) [252]–[255]. Adversarial attacks designed with the surrogate model are known to transfer to the target model (known as the transferability property [256]). In a wireless communications scenario, the adversary can learn the transmission patterns of the victim communication system by observing the spectrum over the air. The effect of wireless channels on surrogate models was studied in [59].

Evasion attack: Also called adversarial attacks, evasion attacks aim to fool ML models into making wrong decisions by manipulating the input test data [14], [17]. Evasion attacks have been

applied to wireless communications in terms of fooling classifiers used for spectrum sensing [34], modulation recognition [139], autoencoder- based end-to-end communication systems [54], channel state information (CSI) feedback for massive MIMO [90], [257], channel estimation [258], [259], and initial access in directional communications [260]. One challenge in the wireless domain is that the adversarial perturbations observed by the target classifier are received subject to channel effects such that these perturbations need to be crafted by accounting for channel effects [28], [31], [32], [56]. These attacks are typically stealthier and more energy-efficient than traditional jamming attacks (such as the one that solely aim to cause interference to data transmission, e.g., [77], [78], [261]) as they only need to transmit low-power signals over a short period of time to confuse the ML algorithms in their decision making.

Causative attack: Also called poisoning attacks, causative attacks aim to manipulate the training process of ML models by injecting vulnerabilities such as false training data to the ML models [262], [263]. Causative attacks are effective against DNNs as they are highly susceptible to inaccuracies in training data. An example of a causative attack in wireless communications is spectrum poisoning where the adversary exploits the (re)training process of the ML classifier so that the ML model is poorly (re)trained. Examples of causative attacks in wireless communications include attacks on spectrum sensing [34], cooperative spectrum sensing [241], and IoT systems [242], [264].

Trojan attacks: Also called backdoor attacks are combinations of evasion and causative attacks, where the adversary injects triggers (backdoors) to training data and then activates them for some input samples in test time [266], [267]. Trojan attack has been considered against wireless signal classifiers in [180] by formulating controlled phase shifts as backdoors and adding them into the transmitted and received RF data samples. Figure 6 gives a pictorial representation of AML attack models and shows how the different attacks work, as well as the general procedure for attack formation.

Categorization Based on Attack Phase.

- *Training phase:* The causative and trojan attacks occur in the training phase of the model development. See Section IV-A for discussion on these attacks.
- *Testing phase:* The exploratory attack, evasion attack and trojan attack occur in the test time. See Section IV-A for discussion on these attacks.

Categorization Based on The Amount of Knowledge The Adversary Has About The Victim Model

White-box attack: The adversary knows the training data, architecture, algorithm and/or optimization techniques such that it has full access to the trained model f and knows the input at the classifier. For an over-the-air white box attack in wireless communications, the adversary also needs to know the channel between the adversary and the receiver, since the channel affects the perturbation perceived at the receiver [28], [31], [32], [59].

Black-box attack: The black-box attack is a more realistic and more relaxed model for many security threats [268] as the adversary has neither knowledge about nor access to the training data and/or the trained model f . In that case, the adversary tries to deduce information from the returned results of the model. Black-box attacks typically use a surrogate (shadow) model, which is trained to perform an identical task as the target network during the inference attack. In wireless communications, the surrogate model strongly depends on the channel characteristics with respect to the adversary and the victim system [59].

Gray-box attack: In a gray-box attack, the adversary has some knowledge of the data/algorithm as well as limited access to the model f . One example from the wireless domain is not knowing the exact channel but knowing an estimate, the distribution or some statistics of the channel, e.g., Rayleigh vs. Rician fading, the value of pathloss exponent or the value of signal-to-noise-ratio (SNR), etc. [28].

Categorization Based on Type of Adversary's Goals

Targeted attack: The adversary tries to produce the perturbation r_x that will cause specific errors in the output of a receiver by targeting a signal x to generate errors such that $f_w(x + r_x) = y$ for a particular target class y . For example, consider a 4-class modulation classification problem with BPSK, QPSK, QAM 16 and QAM 64 as classes and consider the adversary that seeks to fool the modulation classifier into classifying a signal modulated with QAM 16 as a QAM 64.

Non-targeted attack: The adversary tries to produce the perturbation r_x that will cause errors in the output of the algorithm independent of the classes thereby reducing the confidence in the algorithm due to decreased accuracy. In our previous example of a 4-class modulation classification problem with BPSK, QPSK, QAM 16 and QAM 64 as classes, the adversary seeks to fool

the modulation classifier into classifying a signal modulated with QAM 16 as any other available class such as BPSK, QPSK, or QAM 64, i.e., $f_w(x + r_x)$ for any $y = f(x)$

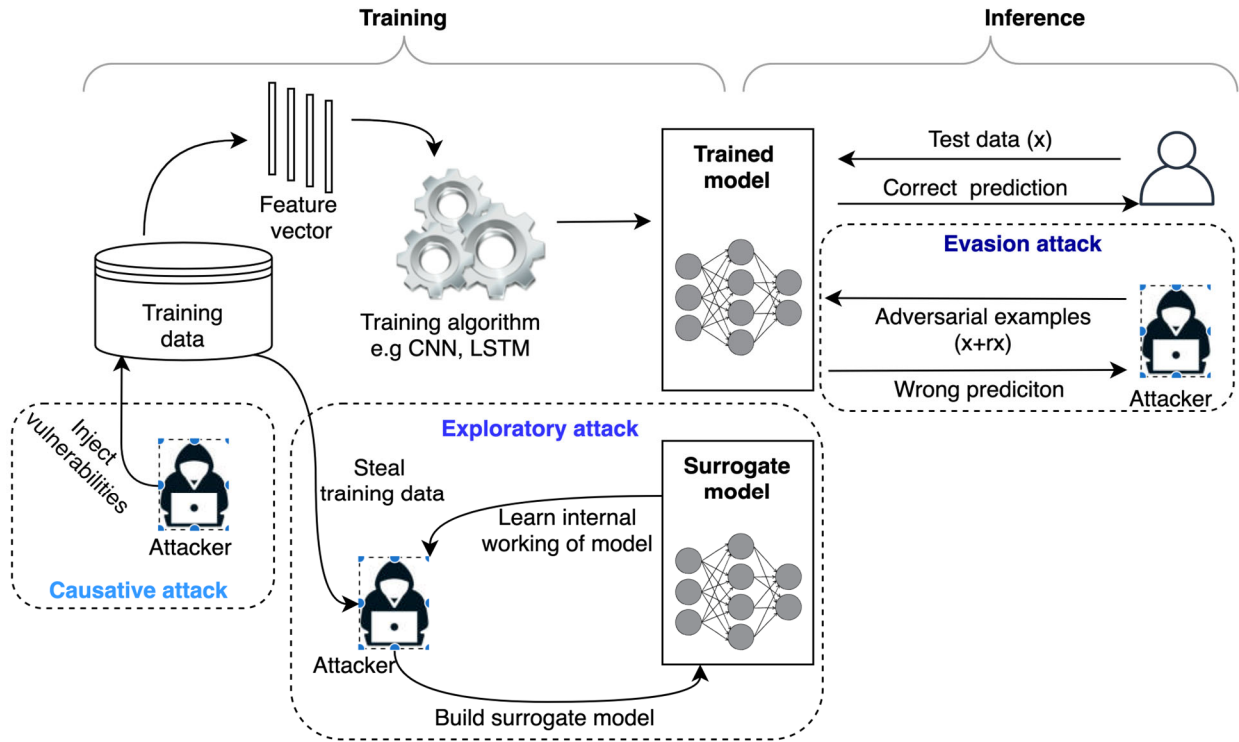


Figure 13: Attacks in adversarial machine learning.

Other Types of Categorization

Some other types of categorizations for AML attacks include categorization based on the frequency of attack and categorization based on the degree of freedom the adversary has in accessing the system's input.

In the attack frequency category [35], there is a one-step attack that computes the gradient of the loss function at a time to create perturbations. There is also an iterative attack that performs the same computation as a one-step attack multiple times to create the perturbations. This is enabled by perpetual access to the model and can be computationally intensive and costly.

Based on the adversary's degree of freedom to the input of the system, there may be physical and digital attacks. In a physical attack, there is an indirect application of the input to the model

which is different from the digital attack where the adversary can explicitly design the input of the model [54] [268]. We can also remove the requirement for the adversary to know the input test samples and craft input-agnostic adversarial attacks, also called Universal Adversarial Perturbation (UAP) [269]. UAP attacks can be built against wireless signal classifiers, where the adversary does not need to know about the transmitted signal [28], [139]. It is also possible to limit attacks to certain users, e.g., evasion attacks to fool classifiers only at a subset of receivers [28] and causative attacks launched from a selected subset of IoT devices to manipulate the training process at a fusion center [242]. These attack types may overlap in the description of an attack. For example, the authors in [29] studied both targeted, white-box evasion attack and non-targeted white-box evasion attacks for DL-based modulation classification.

GENERATION OF ATTACKS IN WIRELESS COMMUNICATIONS

We describe the approaches to generate attacks in AML, including evasion attacks, exploratory attacks, and causative attacks.

Generation of Evasion Attacks: Evasion attacks to generate adversarial examples are characterized as subtly crafted imperceptible perturbations r_x added to the original inputs x of the ML model such that the input to the classifier is $x = x + r_x$ by solving the optimization problem (6) for r_x . In the wireless domain, the problem is changed by setting with $x = h_{t,r}x + h_{t,a}r_x + n$, where x is the received signal, x is the transmitted signal, r_x is the transmitted perturbation, $h_{t,r}$ is the channel gain from the transmitter to the receiver, $h_{t,a}$ is the channel gain from the transmitter to the adversary, and n is the receiver noise. It is difficult to solve (6) due to the non-linearity of the DNN mapping. Instead, we can approximate the solution. This is achieved by maximizing the loss function $L(w, x, y)$ used in the training process of the classifier f_w (e.g., the cross-entropy loss function is $L(w, x, y) = \log(1 + \exp(f_w(x, y)))$) subject to the condition that the perturbation is upper bounded. Thus, the adversarial perturbation r_x can be found by solving

$$\begin{aligned} \max \quad & L(w, x + r_x, y) \\ \text{s.t.} \quad & \min \|r_x\|_p \leq c, \end{aligned} \tag{7}$$

where c is the upper bound on the adversarial perturbation. Various methods have been proposed to approximately solve (7), thus generating adversarial examples. These methods include but are not limited to one-step gradient-based method such as Fast Gradient Sign Method (FGSM) [17],

and its iterative variants such as Basic Iterative Method (BIM) [268], Projected Gradient Descent (PGD) [198], Momentum Iterative Method [270], Box Constrained Limited-Memory Broyden-FletcherGoldfarbShanno (L-BFGS) [14], and Carlini & Wagner (C&W). We describe some of these methods that have been used to generate adversarial attacks on wireless communications.

Fast Gradient Sign Method (FGSM) [17]: One popular method that has been used to generate adversarial examples in the wireless domain is the FGSM. The FGSM is a one-step gradient-based method developed on finding the scaled sign of the gradient of the cost function and aims at minimizing the strength of the perturbation. It is a computationally effective method for adversarial attack generation based on the postulation that the DNNs are prone to linear-type adversarial attacks because of their use of linear techniques for easier optimization. The FGSM-generated perturbation is achieved by linearizing the model's cost function $J(w, x, y)$ around the current value of x . FGSM generates an adversarial example that significantly resembles the original input x by starting with an original input to the ML model and adjusts it in the direction of the gradient of the loss function with respect to the input by an amount limited by a parameter c stated as:

$$x^* = x + c \cdot \text{sign}(\nabla_x J(x, y, w)), \quad (8)$$

where x^* is the crafted adversarial example, x is the original input, $J(x, y, w)$ is the cost function of the DNN with parameters w , y is the class associated with x , and c is a small number to limit the perturbation [243]. Note that the Fast Gradient Method (FGM) is a generalization of FGSM to meet the L_2 norm bound $\|x^* - x\|_2 < c$ as

$$x^* = x + \epsilon \cdot \frac{\nabla_x J(x, y, w)}{\|\nabla_x J(x, y, w)\|_2}. \quad (9)$$

FGSM is also used in iterative methods such as the following.

Basic iterative method (BIM) [268]: The BIM can be viewed as multiple steps of FGSM with a small step size as well as a clipping of results after each iteration to establish that the results are in an c -neighborhood of the original input data. The iteration for the BIM is given by

$$x_i^* = \text{clip}_{x, \epsilon} (x_{i-1}^* + \epsilon \cdot \text{sign}(\nabla_x J(x_{i-1}^*, y, w))) , \quad (10)$$

where $\text{clip}_{x, \epsilon}$ clips the values of the adversarial sample to an ϵ -neighborhood of the original sample x .

Projected Gradient Descent (PGD) [198]: PGD is a multi-step variant of FGSM on the negative loss function that helps initiate adversarial attacks with the aim of examining or analyzing the performance of a neural network from an optimization point of view. It puts a limit on the total perturbation in L_∞ -norm from $x^*0 = x + U_n(\epsilon, \epsilon)$, where $U_n(\epsilon, \epsilon)$ is the uniform random perturbation. In t iterations, x is determined as

$$x_t^* = \Pi_{B_\epsilon(x)} (x_{t-1}^* + \epsilon \cdot \text{sign}(\nabla_x J(x_{t-1}^*, y, w))) , \quad (11)$$

where $\Pi_{B_\epsilon(x)}$ is the projection operator. The procedure of adversarial example generation includes adding noise, computing gradient, stepping, and projecting back.

Momentum Iterative Method (MIM) [270]: The MIM was developed to solve the problems related to over-fitting and local minima. It helps speed up the gradient descent iterations. In scenarios where the performance of FGSM is inhibited due to the presence of strong interference and distortion, the MIM applies momentum to the iterative attacks so as to stabilize and maintain the correct direction of the update with generalization of adversarial examples.

L-BFGS Method [14]: Box-constrained L-BFGS is an example of optimization-based method for generating adversarial examples. For an instance of x , L-BFGS finds a different x^* that is very much alike to x based on the L_2 distance but is labeled differently by the classifier. This problem is laborious and is modeled as the following constrained minimization:

$$\min \|x - x^*\|_2^2, \quad (12)$$

$$s.t. \quad f_w(x + r_x) = y,$$

$$(x + r_x) \in [0, 1]^n,$$

where y is the target class. To solve (12), it is transformed to

$$\min \lambda \cdot \|x - x^*\|_2^2 + J(x, y, w) \quad (13)$$

$$s.t. \quad x^* \in [0, 1]^n,$$

where $J(x, y, w)$ is a function mapping an instance of x to a positive real number using the loss function. In (13), the cross-entropy loss function is typically used as the loss function J and line search is used to find the positive constant λ that generates an adversarial example of minimum distance.

Generation of Exploratory Attack: In an exploratory attack, an adversary seeks to understand the internal operations on an ML model and leverage on such knowledge to attack the ML model and ultimately undermine the effectiveness of the ML-based system. This is achieved by developing a surrogate (or shadow) model that can mimic the operation of the original ML model. The approach is to steal training data and build models that have the same (or similar) output as the original model. With the information on how the original model functions, the attacker can make informed decisions as to how to attack and alter the overall performance of the ML-based system. For example, in [240], an exploratory attack is launched on the cognitive radio as a preliminary step to learn when and how to jam the cognitive radio transmissions. The adversary senses a channel, captures the transmitters decisions, and learns the pattern of successful transmissions by tracking acknowledgments from the receiver, and launches an exploratory attack to build a DL classifier (surrogate model) that is functionally equivalent to the one at that transmitter. In another example [51], an adversary observes the spectrum and develops a DNN to infer the channel access algorithm used by an IoT transmitter and predict the outcome of the transmissions before jamming it. Exploratory attacks are also launched to infer reinforcement learning mechanisms used in wireless channel access and jam the underlying communications. [60], [61] designed a DRL- based jamming attacker that employs a dynamic policy and aims to minimize the channel access accuracy of a DRL- based dynamic channel access user. The DRL attacker is able to observe the victim's interaction with its environment for a period of time and learns the activity pattern. The attacker has no prior information about the channel switching pattern on a victim's action policy and both DRL-based systems can interact with each other, retrain their models and adapt to the opponent's policy. Also, [62], [63] introduced a reinforcement learning (RL) algorithm to disrupt 5G radio access network slicing (based on another RL algorithm [271], [272]) by observing the spectrum and developing an RL-based surrogate model. This model decides which resource blocks to jam in order to attain a high number of failed network slicing requests. Jamming a resource block causes a reduction in the RL algorithm's reward thus, affecting the model's performance as the reward is used to update the RL algorithm. Although the exploratory at-

tacks described above follow the same paradigm, there are no standard attack generation methods as we have in evasion attacks. All the methods use a common approach based on learning the internal workings of the original model, building a surrogate model, and attacking the ML-based model using the knowledge from the surrogate model. Table IV shows examples of previous works on exploratory attacks.

Generation of Causative Attack: The causative attack seeks to downgrade the ability of a ML system to perform optimally by injecting vulnerabilities into the training process. These injected vulnerabilities can be because of the adversary by intentionally meddling with the label of the data or by crafting and embedding unique learnable features. The model learns these unique features which causes it to deviate from its original objective, thus, causing considerable degradation in the ability of the model to make good predictions. The injection of vulnerabilities might be during the initial model development or the model re-training process. Since the attack occurs before model training, the contamination affects any type of ML model and efforts on parameter tuning yields little or no improvement on the model's ability to make good predictions [265]. As we have observed in the exploratory attack, there are no standard attack generation methods but the paradigm of poisoning the training data is consistent in all forms of causative attacks. Table IV shows some previous work on causative attack.

APPLICATION AREAS OF ADVERSARIAL MACHINE LEARNING IN WIRELESS COMMUNICATIONS

Grouped according to the type of DL problem, we discuss and review various studies on AML in wireless communication and present them in Tables V (for modulation classification), Table VI (for spectrum sensing), Table VIII (for signal classification), and Table IX (for other areas in the wireless communication domain).

Table 9: Review of Previous Works On AML Attacks In Modulation Classification

Paper	Attack Type	Objective	Attack Method	DL Problem	Data used
[29]	White-box, Blackbox, Targeted, Untargeted, Evasion	Explore how to launch a realistic evasion attack by taking into consideration the channel effect as well as the power constraint at the adversary.	FGM, UAP	Modulation classification	RML2016.10A [275]
[182]	Evasion	Test the attack efficiency of two types of adversarial attacks.	FGSM, Box constrained L-BFGS	Modulation classification	RML2016.04C [4]
[35]	Evasion, Black-box, White-box	Explore the feasibility and effectiveness of adversarial attacks and determine the optimal perturbation level for attack invisibility and success.	FGSM, PGD, BIM, MIM	Modulation classification	RML2016.10A [275]

Table 10: Review of Previous Works On AML Attacks In Modulation Classification(a)

[180]	Trojan attack	Introduce a new attack that slightly manipulates training data by inserting Trojans (i.e., triggers) to a few training data samples and then activate them later in test time.	Embed Trojans in training data and trigger it in test time	Modulation classification	RML2016.10A [275]
[55]	Evasion	Evaluate the vulnerabilities of raw I/Q based modulation classification.	FGSM	Modulation classification	RML2016.10A [275], three synthetic datasets using GNU radio
[31],[236],[276]	Evasion	Minimize the accuracy of the intruder in determining the modulation scheme used by a transmitter by perturbing the channel input symbols.	PGD	Modulation classification	Locally generated data
[179]	Gray-box	Mitigate adversarial examples in RF classifiers using autoencoder preprocessing.	FGSM	Modulation classification	RML2018A [277]
[30]	Evasion White-box Targeted	Investigate the use of multiple antennas to generate multiple concurrent perturbations over different channel effects (subject to total power budget) to the input of DNN-based modulation classifier at a wireless receiver.	FGM	Modulation classification	RML2016.10A [275]

Table 11: Review of Previous Works On AML Attacks In Modulation Classification(b)

[139]	Evasion, White-box, Black-box	Show the susceptibility of DL models to adversarial attacks and present practical methods for crafting adversarial examples in modulation classification.	PCA-based UAP, FGM-modified bi-section search	Modulation classification	RML2016.10A [275]
[189]	Evasion, Targeted, White-box	Fool a DL-based eavesdropper by a cooperative jammer transmitting adversarial perturbation so as to hide 5G communications from the eavesdropper.	FGM	Modulation classification and 5G system	Locally generated data using MATLAB 5G tool-box
[33]	Evasion	Extend the use of communication-aware evasion attack through the use of forward error correction (FEC).	Perturbation created by Adversarial Mutation Network	Modulation classification	Synthetic data generated using Liquid DSP
[181]	Evasion, Untargeted, White-box, Targeted	Examine how the DL-based automatic modulation classifier breaks down in the presence of an adversary with direct access to its inputs.	FGSM, MI-FGSM	Modulation classification	RML2016.10A [275].
[28]	Evasion, Black-box, Untargeted, White-box, Targeted	Explore how to design realistic adversarial attacks in the presence of realistic channel effects and multiple classifiers at different receivers.	FGM, UAP	Modulation classification	RML2016.10A [275]

Table 12: Review of Previous Works On AML Attacks In Modulation Classification (c)

[32]	White-box, Untargeted Targeted	Provide a generalized wireless AML problem (GWAP) and experimental evaluation of AML attacks launched against wireless DL systems	AML waveform jamming, AML waveform synthesis	Modulation classification, Radio Fingerprinting	RML2018.01A [277], 1000-device dataset of WiFi and ADS-B transmissions
[58]	Evasion	Introduce spectral deception loss metric implemented during training to make the spectral shape more in-line with the original signal.	Adversarial Mutation Network (AMN)	Modulation classification	Locally generated data
[183]	Evasion	Minimize the accuracy of an intruder while still ensuring the successful decoding of a signal by the intended receiver.	PGD, uniform random noise of $L2$ -norm to block of 128 I/Q symbols	Defense against modulation detection	Locally generated data
[278]	Evasion	Study the effect of adversarial attacks on ML-based AMC model using different data-driven subsampling strategies	Carlini-Wagner	Modulation classification	RML2016.10B [279]
[280]	white-box, blackbox	Present an input-agnostic adversarial attack that is undetectable and robust to removal	UAP	Modulation classification	RML2016.10A [275]

Modulation Classification and Signal Classification

Modulation classification – the process between the detection and demodulation of a signal – is an important step towards developing an intelligent radio receiver and has been extensively studied [273]. Recently, the research in this area has gained new attention with the use of DL. [274] showed that DL algorithms for modulation classification using RF data outperform the use of expert features and higher order moments. AML for modulation recognition has been extensively using evasion and/or trojan attacks. [35], [139], [182] showed the susceptibility of DL models to adversarial attacks on modulation classification under different scenarios such as data with noise and variation in SNR level. [55], [139] pointed out that the adversarial examples for modulation classification are more effective than perturbations crafted by adding Gaussian noise. Table V shows AML studies focused on modulation classification. Like the modulation classification problem, AML has also been applied to wireless signal classification. Signal classification aims to identify the transmitting signal in a spectrum. Table VIII highlights relevant work on AML for signal classification.

Spectrum Sensing: Most of the AML studies focused on spectrum sensing and dynamic spectrum access (DSA) use exploratory attacks in which the attacker learns the operational pattern of the DNN model and builds a surrogate model which gives the same or similar output as the original classifier when given the same input. Spectrum sensing and DSA are key components of cognitive radio systems for efficient discovery and use of spectrum to achieve situational awareness [53]. Given the broadcast and over-the-air nature of wireless communications, there are various attacks that attempt to undermine the spectrum sensing results. This can include causative attack to change the sensed features as well as priority violation attacks to violate priorities in channel access by pretending to have higher priority and transmitting during the sensing phase [51]. AML attacks on spectrum sensing typically involve an adversary that develops a surrogate model and can correctly predict the outcome of the original DL model. Such an adversary jams the successful transmissions, thus reducing the success rate of the original model [52], [60], [240]. As shown in [34], [52], [240], [241], these attacks are more energy-efficient and more effective than random jamming. Table VI highlight the studies on AML for spectrum sensing and DSA applications.

Resource Allocation and Management: The need to satisfy the demand for faster response time, increased reliability of wireless transmissions as well as larger bandwidth has made the use of efficient algorithm and approach for resource allocation and control such as power control, and beamformer design increasingly important. ML/DL techniques applied to achieve optimal resource usage are also susceptible to adversarial attacks. Table VII shows some studies on AML for resource allocation and management.

Table 13: Review of Previous Works On AML Attacks In Spectrum Sensing.

Paper	Attack Type	Objective	Attack Method	DL Problem	Data used
[52], [240]	Exploratory	Launch an exploratory attack on cognitive radio as a preliminary step before jamming.	Jam legitimate user transmissions	Cognitive radio channel access	Locally generated data
[53]	Spectrum poisoning, Exploratory	Launch an over-the-air attack to infer the transmitter's behavior and falsify the spectrum sensing data by transmitting perturbations over the air.	Inject perturbation to the channel	Spectrum sensing	Locally generated data
[51]	Exploratory, Evasion, Causative, Black box	Build a DNN classifier to infer the channel access algorithm of a transmitter and predict the transmission result and develop a defense mechanism to confuse the adversary in its attack strategy.	Jamming, Spectrum poisoning, Priority violation	Spectrum sensing	Locally generated data
[241]	Black-box, Exploratory	Propose Learning-Evaluation-Beating(LEB) attack against cooperative spectrum sensing (extension of spectrum sensing data falsification attack on cognitive radio).	LEB attack	Cooperative spectrum sensing	Dataset collected from 5282 locations [281], Data from 5 locations using USRP N210

[61]	Black box Gray box	Develop deep reinforcement learning (DRL)-based dynamic channel access system and DRL-based jamming attack to study the sensitivity of DRL-based wireless communications to adversarial attacks.	DRL-based attacker uses dynamic policy to minimize the victim's channel access	Dynamic spectrum access (DSA)	Locally generated data.
[34]	Exploratory, Evasion, Causative	Launch over-the-air spectrum poisoning attacks by learning first the transmitter's behavior.	Attacker jams the transmitter's sensing period	Spectrum sensing, DSA	Locally generated data
[60]	Exploratory	Compare the performances of two adversarial policies: feed forward neural network (FNN) and DRL policies on the performance of a DRL-based dynamic channel access agent.	Jamming of a victim model	Dynamic spectrum access	Locally generated data
[282]	Exploratory, Evasion	Present AML attack on spectrum sharing of 5G communications with incumbent users.	Manipulate inputs to the classifier by transmitting during spectrum sensing	Spectrum sharing	Locally generated data

Detection and Mitigation of AML Attacks in Wireless Communications

We present various methods that have been used to defend against AML attacks on wireless communication systems.

Train with Adversarial Perturbations: Training the DNNs with adversarial examples is a popular method of mitigating the effect of adversarial attacks and ensuring robustness [198]. Adversarial training consists of generating adversarial examples according to one or more attack methods and retraining the DNNs with labeled adversarial examples. The goal is to prevent an adversary from identifying the structure of the signals transmitted to prevent it from disrupting the communication system. This defense mechanism has been applied to protect wireless signal classifiers against evasion attacks [28], where randomized smoothing [293] used in training time increases the robustness of classifiers against adversarial attacks later in test time. There are some challenges that pertain to adversarial training; some of these include a scenario where the adversary uses a different attack from the one used in training. In addition to that, an adversary can create new perturbations using a model trained with adversarial examples. However, training with adversarial examples typically reduces the performance of DL models on unperturbed signals. In trying to overcome these challenges, training can be performed with multiple adversarial examples [198]. In addition, certified defense [294] can be built in train and test times to ensure statistical significance of classification results in the presence of adversarial attacks in test time. Certified defense has been used in [28] to provide performance guarantees for wireless signal classifiers in the presence of adversarial examples. Another way to mitigate adversarial attacks during training is pre-training of the DL based wireless signal classifiers by an autoencoder such that the trained model becomes robust to the deceiving effect of adversarial examples [179].

Statistical Approaches

Peak-to-average power ratio (PAPR) of RF samples: One way of detecting adversarial examples is to leverage the properties of radio frequency (RF) data and apply a statistical test by taking advantage of the PAPR of digitized RF samples of the received signals [243]. The PAPR distribution is used in the wireless domain to depict a specific class of signal's footprint such that if the inferred output of an input signal is categorized to a particular class and its PAPR statistic contradicts with high level of confidence, further analysis and assessment should be carried out to reach a true conclusion. To determine if an adversarial example is present in a wireless communication environment, the Kolmogorov- Smirnov (KS) two-sample test (performed by collecting a sample sized input and then computing and evaluating the PAPR distributions for the samples) can be used to ascertain if the PAPR is similar to the statistic of an adversarial example or a legitimate example [243].

Use the softmax outputs of the DNN classifier: This is also a KS statistical test for adversarial example detection. The softmax output of the DNN classifier is used to determine if adversarial examples have brought about a change in distribution using the statistics of the final layer of a neural network. Typically, the statistical size of the input data is a good determination of the quality of the statistical test. The effectiveness of the method depends on the type of waveform and propagation channel, as discussed in [243]

Statistical methods to detect adversarial triggers: Adversarial triggers such as Trojans (or backdoors) inserted to training data can be detected by statistical methods such as clustering and median absolute deviation (MAD) algorithm. In particular, the MAD algorithm computes the median of the absolute deviations from the data's median, namely $\text{median}(x_i - X)$, where $X = \text{median}(X)$, discover outliers, and is shown in [180] to be effective against Trojan attacks on wireless signal classifiers

Randomizing the DL Algorithm by Adding Small Variations to Classifier Outputs

This approach adds slight variations to the output of the DNN classifier and deliberately performs some (typically a small number of) erroneous actions so as to deceive the adversary and defend against exploratory attacks. For example, the transmitter can intentionally utilize a busy channel for transmitting or skip a transmission opportunity in an idle channel such that the adversary that observes the spectrum cannot train a reliable surrogate model. This defense mechanism has been applied to protect wireless signal classifiers against inference attacks (before launching jamming attacks) in [240] as well as against evasion and causative attacks in [34]. Note that network uncertainty (e.g., due to dynamic channel and traffic effects) and randomization or obfuscation-based defense methods have also been leveraged before against traditional network attacks such as jamming and traffic inference to confuse the adversaries [77], [78], [261], [295]. Defense against AML attacks on reinforcement learning follows a similar diversification approach to protect wireless communications [61], [186]. The defender may alternate among different channel access decisions (derived through proportional-integral-derivative (PID) controller and imitation learning) to increase the uncertainty at the adversary. Another defense approach is to adopt orthogonal policies in reinforcement learning to prevent the adversary from quickly switching between its imitated policies. On top of these mitigation policies, it is also viable to detect the adversaries by monitoring changes in rewards and distinguishing the

spectrum environment changes from adversarial jamming attacks [61], [186]. Table X reviews the defense strategies that have been used by previous works in the wireless domain.

Future Outlook of Adversarial Machine Learning in Wireless Communications

As ML/DL becomes the core of current and emerging communications systems such as 5G and 6G, ML/DL itself becomes susceptible to adversarial effects. To build on the promise of intelligent operations and efficient resource management with the application of ML in dynamic wireless environments, it becomes imperative to develop ML models that are secure, resilient and robust to attacks by taking into account the unique properties of wireless communications. An AML attack model that is carefully characterized by considering wireless properties is expected to be the foundation of ML driven wireless security research and development for the existing and emerging communication systems. In this section, we discuss some important ideas that can help in realizing the promise of robust models that are effective even in the presence of adversaries.

Wireless communication dataset: The limitation of standardized real-world datasets that adequately represent real life scenarios in the wireless communication domain is a challenge that needs to be solved. Compared to other domains such as CV and NLP, there are just a handful of publicly available ML/DL datasets in the wireless domain to work with such as [4], [111], [296]–[299]. Typically, these datasets do not include adversarial effects recorded. It is important for the research community to develop more publicly available datasets, representing different scenarios including not only variations in channel, interference type and waveform, but also AML attacks. This approach will help with benchmarking solutions for robust development and evaluation of the DL models for wireless versions of AML attacks. From literature reviewed, a large ratio of the work on AML in wireless communication is in the area of modulation classification. This is largely because modulation classification datasets (e.g. [4], [111]) are the seemingly standardized datasets for ML/DL studies in the wireless communications domain. Although more recent datasets are available such as [296]–[299] as well as testbed efforts such as [300] expand datasets, it is critical to implement the AML attacks and corresponding defenses with embedded radio platforms to assess the real channel and radio hardware effects.

Robust features: The need for robust features in developing ML/DL models in wireless communication is also an important consideration for future systems. [301] laid the analytical founda-

tion and showed via experimental data that adversarial attacks are very effective because most existing ML/DL models are not developed using robust features. Although the discussion of feature engineering and identifying robust features are on-going especially in CV and NLP domains, new techniques must be developed to identify the most important features which are robust to the effect of adversarial attacks in wireless communication systems. An important research direction for AML in wireless communications is to identify robust features in datasets, disentangle non-robust features from robust features, train models that generalize well, and are robust to adversarial attacks in wireless communications.

Certifiable defense: The design and development of certifiable defense mechanisms for AML is another important component of the work that needs to be achieved to realize the full potential of ML/DL in wireless communications domain. Various defense mechanisms have been proposed, including adversarial training and use of statistical methods but most of these defenses become ineffective with more powerful adversaries. This has led to the need for certifiable defense mechanisms, where prediction on the test data can be verified to be a constant within a small region around the input data. Although there are some efforts that present certifiable defense using randomized smoothing [28] with reasonable performance, more rigorous evaluation of such approaches as well as development of other methods are needed to provide security guarantees for ML models used in wireless communications.

Although important areas of development to ensure good performance from robust ML/DL models have been discussed in this section, for the safe adoption of ML/DL-based models into future communication systems, there is also a need to move from black-box models to interpretable models. Such paradigm as seen today includes physics-guided ML [302], where the fundamental laws of physics is infused into the learning process to improve generalizability and explainability, and ensure robustness to attacks, as needed in wireless communication systems.

To conclude, the recent trends observed in this survey show that the AML attacks can effectively disrupt wireless communications. Compared to the traditional jamming attacks or the addition of noise/interference, adversarial attacks are more difficult to detect operating with small spectrum footprint, more energy efficient, and overall, more effective. Furthermore, as wireless communication evolves into smarter systems as in 5G and 6G, the AML attacks also become smarter. The attacks have evolved from adversaries that send jamming signals randomly to overwhelm a

communication system to those that develop surrogate models to know the internal workings of the original model and send out small perturbation signals to fool the ML/DL models and disrupt wireless communications. The goal of AML is to support safe adoption of ML/DL solutions to the emerging applications in the presence of adversaries. Wireless communications strongly benefit from ML/DL applications in terms of learning from spectrum data and solving complex tasks. Various studies reviewed in this paper have established that AML attacks on the ML/DL-driven wireless systems are effective and pose serious threats leading to major performance degradation. This review attempts to present the current state of the art in the area of AML attacks specific to the wireless communications domain and help the research community identify the research accomplishments as we forge ahead to investigate methods by which efficient data driven intelligent systems can be built while taking into account the unique nature of wireless communications, robustness, resilience, and security of ML/DL based methods. We summarized the latest research efforts by reviewing the attack types, adversarial example generation, and defense strategies, established a taxonomy across different attack categories, and discussed the future outlook to AML in wireless communications. This is an emerging research area and there is an urgent need for further investigations to quantify the impact of AML attacks and detect and mitigate them to enable the safe adoption of the promising ML/DL solutions for wireless communications.

4.0 RESULTS AND DISCUSSION

In this section, we first discuss how to run the TwitterOSINT system along with a few examples to show how to use the system for cybersecurity operations. We then discuss the preliminary experiment conducted to gain operational experiences and insights about the system and its capabilities, and our observations.

Running TwitterOSINT System:

To run the TwitterOSINT system, start the Stream.java application, Logstash, Elasticsearch and Kibana services. Keep them running. The Stream.java application maintains a connection to Twitter, to download the relevant tweets. It then uses the NLP pipeline to analyze the tweets, and store the texts with the automatically generated annotations in the `twitDB.json` file.

The Logstash service is configured to digest the `twitDB.json` file and forward the data to Elasticsearch which then indexes and stores the tweets along with their annotations in the centrally managed data store. The UI provided by the Kibana service is used to create visualizations of the data stored in Elasticsearch, navigate the Elastic Stack, and execute commands.

Sample Data:

The textbox below shows a sample tweet (with the date, username, and the text of the tweet) which was processed by the TwitterOSINT system.

► September 4th 2017, 03:55:11.078 cyberinbox

Google removed almost 300 Android apps involved in

DDoS attack ps://t.co/UpZNV96PRK

A section of the annotations generated for this tweet's text is shown below using the JSON format. The text is broken down into tokens. Each token is tagged with three fields: `CyberAnnot` (which denotes the cyber entity: vulnerability, software vendor, software product, etc. or "O" if not a cyber-related entity), `POS` (which denotes the part-of-speech), and the token itself.

```

“annotations”:[{“CyberAnnot”：“Sw.vendor”,“POS”：“NNP”, “to-
ken”：“Google”},{“CyberAnnot”：“O”,“POS”：“VBD”,“token”: “re-
moved”},{“CyberAnnot”：“O”,“POS”：“RB”,“token”：“almost”},{“CyberAnnot”：“Sw.product”,“P
OS”：“JJ”,“token”：“Android”}
...
{“CyberAn-
not”：“vuln.description”,“POS”：“NN”,“token”：“DDoS”},{“CyberAnnot”：“vuln.description”,“PO
S”：“NN”,“token”：“attack”}
...

```

Data Visualization

Manually examining a large number of indexed tweets is not practical for an analyst. To illustrate how the power and ease-of-use of the Elastic Stack can be utilized to help analysts define and execute tasks of high interest and value to them in an automated manner, the Elastic Stack was used to create five visualizations to make it easier to gather and view the aggregated or related information. Elasticsearch provides a feature named *Significant Terms Aggregation* [13] which is an aggregation that returns interesting or unusual occurrences of terms in a set. The terms selected are not simply the most popular terms in a set. They are the terms that have undergone a significant change in popularity when measured between a *foreground* (e.g., tweets indexed today) set and a *background* (tweets indexed until but not including today) set.

For example, suppose the term *Ransomware* existed in only 5 out of 10 million tweets indexed previously, and yet it is found in 4 of the 100 tweets that were indexed today. Such a big shift in frequency (5/10,000,000 or 0.000005% versus 4/100 or 4%) is significant, and it likely indicates a need for further investigation [13]. It is common that a current threat, or a software product affected by a threat, would appear in the day’s tweets with a higher-than-normal frequency. We built a pie chart [14], as shown in Figure 5 to show the distribution of the token words in the tweets aggregating on the Most Significant Terms of a day.

This chart can help analysts to more quickly spot significant increases in frequencies of cyber threat and vulnerability terms appearing in tweets, e.g., on a daily basis.

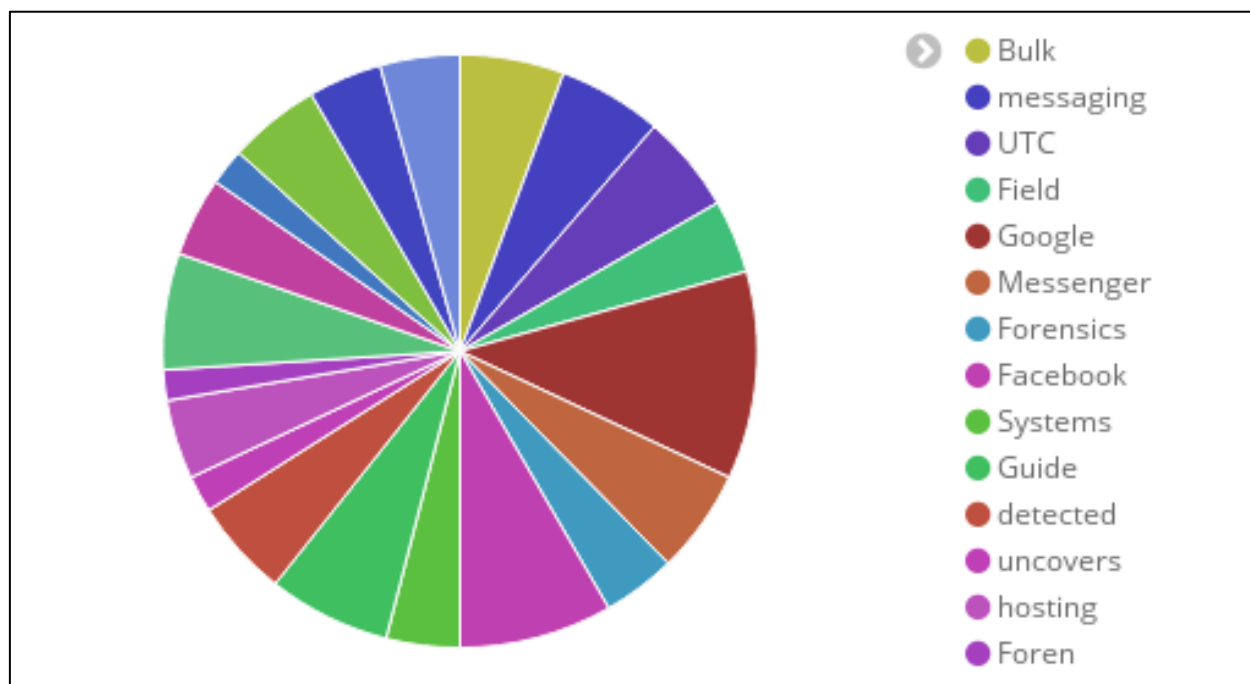


Figure 14: Significant terms distribution (pie chart).

A *tag cloud* is a visual representation of free-form text [14]. Tags are usually single words, and the importance of each tag is shown with font size and/or color as determined by the *metrics* aggregation. Significant Term Aggregation was used to build a tag cloud for TwitterOSINT. The purpose of this tag cloud is to represent the most significant words of the day, and to provide one-click access to all tweets that contain a particular word of interest. For example, in the tag cloud shown in

Figure 15: Significant terms aggregation (tag cloud).

, clicking on *Google* would display all the indexed tweets that contained the word *Google* on that day.



Figure 15: Significant terms aggregation (tag cloud).

A *count* aggregation returns a raw count of the elements in the selected index pattern. This visualization provides a quick view of the number of relevant tweets of the day. When a word is clicked on the tag cloud to display the tweets containing that word, the Count matrix displays the number of tweets returned by the query, as illustrated in figure 11.



Figure 16: Total count of tweets with relevant information.

The last two Kibana-generated visualizations that we built were displayed as tables. The first table, as shown in figure 12, displays the tweets containing the top 20 significant terms of the day. The second table, as shown in **Figure 13**, contains all the relevant tweets of the day. These tables provide a quick view of the day's tweets. The tables could be modified to show the tweets of interest, based on clicking a word in the tag cloud.

"Bulk messaging #malware in #Facebook Messenger" https://t.co/SzcEktil09 #InfoSec #Security #CyberSecurity #CyberAttack #CyberWar #Hacker: Tweet	
Most Significant words of the day ↕	Count ↕
Bulk	2
messaging	2
Messenger	2
Export: Raw  Formatted 	
1 2 »	
https://t.co/MAuIXRvYiq detected 102 new websites hosting #phishing - [03:05 UTC] #infosec #cybersecurity #email #scam #tech #info #data: Tweet	
Most Significant words of the day ↕	Count ↕
UTC	1
detected	1
hosting	1
Export: Raw  Formatted 	
1 2 »	
Malware Forensics Field Guide for Linux Systems: Digital Foren https://t.co/4MHPkveBFz #Cybersecurity #CyberAttack https://t.co/g8AEe6H2Ky: Tweet	

Figure 17: Tweets with top 20 significant terms.

All Relevant tweets 	Count 
"Bulk messaging #malware in #Facebook Messenger" https://t.co/SzcEktiL09 #InfoSec #Security #CyberSecurity #CyberAttack #CyberWar #Hacker	2
"VMware Puts Security at the Heart of its Software" https://t.co/qawGaGXlbK #ransomware #malware #security... https://t.co/iaRiXjsXAd	1
"Why the US Government Shouldn't Ban Kaspersky Security Software" https://t.co/E2iz1ecZBy #tech #malwaretech #CyberSecurity #virus #Malware	1
#Avira Antivirus Pro 2017 is the best virus protection https://t.co/D3RP5LGDv	1
#Canadian university loses \$10m in #phishing scam - BBC News https://t.co/XubkE6OjFP #CyberSecurity #dataprotection https://t.co/5xrH3oxQWm	1
	101
Export: Raw  Formatted 	
1 2 3 4 5 ...20 »	

Figure 18: All relevant tweets.

Experiment

We conducted a very preliminary experiment to assess the efficacy of the TwitterOSINT system in terms of tagging tweets with information relevant to the cyber domain. During this experiment, three groups consisting of 500 randomly chosen tweets each, were chosen. The tweets were manually studied to identify the properties of a tweet with relevant data. The properties are depicted by the presence or absence of cybersecurity entities in the form of annotations within the JSON objects. Results are summarized in Table 1.

Table 14: TWEET SAMPLING RESULTS

	<i>No of tweets</i>	<i>tweets with relevant information (a)</i>	<i>Relevant tweets with all 3 Cyber entities (b)</i>	<i>Relevant tweets without one or more Cyber entities (c)</i>	<i>No of tweets in (c) where the information is covered in (a)</i>
	500	18	14	4	2
	500	12	9	3	1
	500	21	17	4	2
Total	1500	51	40	11	5
Percentages	100%	3.4% - of sample 100% - of relevant tweets	78.43%	21.56%	9.8%

The experiment observed that 51 of the 1500 (3.4%) sampled tweets were relevant, meaning they contained some interesting information related to a vulnerability. 78% (40/51) of those tweets had a vulnerability description, a software vendor and/or a software product which is affected by the vulnerability. In the remaining 22% of the tweets with relevant information, one or more of these fields was missing; however, about half the time the information they did contain was already covered in the other 78% of relevant tweets.

For this experiment, the observed average daily unfiltered tweet count was about 10,000 tweets. We developed a rule in Kibana to filter down daily tweets to show only those tweets that contain the three entities: *vulnerability*, *software vendor* and/or *software product*. This rule removed most of the irrelevant tweets, so that on average 130 tweets remained on the filtered-down list. The rule also removed a small percentage (~12%) of tweets that did in fact contain relevant information. Also, there was a small percentage of *false-positive tweets* in the filtered down list. False-positives are tweets that do not actually contain information pertaining to cybersecurity, or the information they contain is simply informational and not about a recent cybersecurity threat.

5.0 CONCLUSIONS

5.1 Cybersecurity Research and Experimentation Testbed

This research presented the architectural design and implementation of the Version 1 CASE-V testbed for cybersecurity research and experimentation using the OpenStack cloud platform, Hadoop big data platform, iSCSI SAN, and network connectivity. The testbed is initially operational at the platform and networked system levels. Additional capabilities at the applications and platform levels are being expanded.

Users of the testbed are currently limited to developers who continue to build up the testbed. Future plans for major enhancements to the testbed include: (1) Add a graphics processing unit-enabled and all solid state drive based cluster for high performance computing needs and to support machine learning and big data analytics. (2) Enhance the automation and UI framework to add more features and functionalities to support both OpenStack and Hadoop platforms. (3) Port, migrate, enhance, and add existing and new application scenarios (e.g., blockchain) to the testbed. (4) Support additional users from selected research partners. (5) Develop tools to facilitate the creation, management and operation of user-defined projects

5.2 Cybersecurity Threat Intelligence Collection and Analysis

This paper described the design and implementation of a TwitterOSINT prototype system that can help analysts gather and derive cybersecurity intelligence using open source data from Twitter. The TwitterOSINT prototype system was implemented using the Twitter Streaming API to collect tweets that contain specified keywords related to cybersecurity. It used the Stanford CoreNLP library and the Stucco cyber-domain specific entity extraction library from Oak Ridge National Lab to analyze the unstructured tweet text and automatically classify and label cyber related entities from the texts. The system also used the Elastic Stack to store, search and visualize the annotated tweets.

One major challenge in using a data source like Twitter for cyber intelligence lies in the accurate classification and labeling of domain-specific entities from unstructured text, since most NLP tools are trained to handle the general English language. With the aid of ORNL's cyber-domain specific entity extraction library to address this challenge, the TwitterOSINT prototype system performed well.

Overall, the combination of NLP, machine learning and Elastic Stack provided a very powerful, extensible, and easy to use platform. The TwitterOSINT framework shows promise as a tool for helping cybersecurity analysts monitor the latest cyber threats being uncovered around the globe, empowering them to be proactive in protecting their organizations' assets from the threats, with a high degree of automation and accuracy.

In the future, we plan to incorporate additional open sources into the OSINT prototype system, such as security blogs, news websites, and other online venues where users discuss and debate cybersecurity threats, attacks and vulnerabilities. We also plan to make use of additional capabilities provided by the Elastic Stack, such as its built-in machine learning and alerting features, to facilitate further automation of data analytic and cybersecurity operation tasks. In addition, we like to trial our TwitterOSINT prototype system in actual cyber operation environments so we can get input and feedback from people who work on cyber operations to help guide our future planning, design, implementation, and experimentation efforts.

5.3. Adversarial Machine Learning in Wireless Communications using RF Data

In this research work, we developed multi-layered deep learning approaches for classifying APTs. Procedures of this research including data analytics, image conversion, deep learning models, and simulation setups were described and rationalized in detail. 8 APTs with about 3000 PE samples were adopted into this research. In total, we tested 21 simulations with various parameter setups as described in this paper. Our results indicated that our research goal is reasonable and achievable. We successfully classified 8 APTs with training accuracy above 90%. In the future, we plan to collect more sample files from more types of APTs. These sample files will include PE files, log files, and other tangible evidence related to APTs. We also plan to study more simulation setups for each developed model to optimize our solutions. In addition, we would like to apply more advanced data analytic techniques to build an adaptive model that could automatically learn and train from its surrounding environments in real time.

6.0 REFERENCES

- [1] T. Erpek, T. J. O'Shea, Y. E. Sagduyu, Y. Shi, and T. C. Clancy, "Deep learning for wireless communications," in *Development and Analysis of Deep Learning Architectures*. Springer, 2020, pp. 223–266.
- [2] O. Simeone, "A very brief introduction to machine learning with applications to communication systems," *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 4, pp. 648–664, 2018.
- [3] L. J. Wong, W. H. Clark IV, B. Flowers, R. M. Buehrer, A. J. Michaels, and W. C. Headley, "The RFML ecosystem: A look at the unique challenges of applying deep learning to radio frequency applications," *arXiv preprint arXiv:2010.00432*, 2020.
- [4] T. J. O'Shea, J. Corgan, and T. C. Clancy, "Convolutional radio modulation recognition networks," in *Engineering Applications of Neural Networks*, C. Jayne and L. Iliadis, Eds. Cham: Springer International Publishing, 2016, pp. 213–226.
- [5] F. Liang, C. Shen, and F. Wu, "An iterative BP-CNN architecture for channel decoding," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 144–159, 2018.
- [6] H. Jafari, O. Omotere, D. Adesina, H. Wu, and L. Qian, "IoT devices fingerprinting using deep learning," in *IEEE Military Communications Conference (MILCOM)*, 2018.
- [7] T. Jian, B. C. Rendon, E. Ojuba, N. Soltani, Z. Wang, K. Sankhe, Gritsenko, J. Dy, K. Chowdhury, and S. Ioannidis, "Deep learning for RF fingerprinting: A massive experimental study," *IEEE Internet of Things Magazine*, vol. 3, no. 1, pp. 50–57, 2020.
- [8] K. Merchant, S. Revay, G. Stantchev, and B. Noursain, "Deep learning for RF device fingerprinting in cognitive communication networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 160–167, 2018.
- [9] K. Davaslioglu, S. Soltani, T. Erpek, and Y. E. Sagduyu, "DeepWiFi: Cognitive WiFi with deep learning," *IEEE Transactions on Mobile Computing*, 2019.
- [10] O. Omotere, J. Fuller, L. Qian, and Z. Han, "Spectrum occupancy prediction in coexisting wireless systems using deep learning," in *IEEE Vehicular Technology Conference (VTC-Fall)*, 2018.

- [11] H. Sun, X. Chen, Q. Shi, M. Hong, X. Fu, and N. D. Sidiropoulos, "Learning to optimize: Training deep neural networks for wireless resource management," in IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), 2017.
- [12] T. S. Cousik, V. K. Shah, J. H. Reed, T. Erpek, and Y. E. Sagduyu, "Fast initial access with deep learning for beam prediction in 5G mmwave networks," arXiv preprint arXiv:2006.12653, 2020.
- [13] N. Abuzainab, M. Alrabeiah, A. Alkhateeb, and Y. E. Sagduyu, "Deep learning for thz drones with flying intelligent surfaces: Beam and hand- off prediction," in IEEE International Conference on Communications Workshops (ICC Workshops), 2021.
- [14] C. Szegedy, W. Zaremba, I. Sutskever, J. B. Estrach, D. Erhan, I. Good- fellow, and R. Fergus, "Intriguing properties of neural networks," in International Conference on Learning Representations (ICLR), 2014.
- [15] Y. Vorobeychik and M. Kantarcioglu, "Adversarial machine learning," Synthesis Lectures on Artificial Intelligence and Machine Learning, vol. 12, no. 3, pp. 1–169, December 2017.
- [16] Y. Shi, Y. E. Sagduyu, K. Davaslioglu, and R. Levy, "Vulnerability detection and analysis in adversarial deep learning," in Guide to Vulnerability Analysis for Computer Networks and Systems. Springer, 2018, pp. 211–234.
- [17] I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," CoRR, vol. abs/1412.6572, 2015.
- [18] N. H. Mahmood, H. Alves, O. A. Lpez, M. Shehab, D. P. M. Osorio, and M. Latva-Aho, "Six key features of machine type communication in 6g," in 2020 2nd 6G Wireless Summit (6G SUMMIT), 2020, pp. 1–5.
- [19] M. G. Kibria, K. Nguyen, G. P. Villardi, O. Zhao, K. Ishizu, and F. Ko- jima, "Big data analytics, machine learning, and artificial intelligence in next-generation wireless networks," IEEE Access, vol. 6, pp. 32 328– 32 338, 2018.
- [20] G. Wikstrm, J. Peisa, P. Rugeland, N. Johansson, S. Parkvall, M. Girnyk, G. Mildh, and I. L. Da Silva, "Challenges and technologies for 6g," in 2020 2nd 6G Wireless Summit (6G SUMMIT), 2020, pp. 1–5.

- [21] L. Zhang, Y.-C. Liang, and D. Niyato, “6G visions: Mobile ultra- broadband, super inter- net-of-things, and artificial intelligence,” *China Communications*, vol. 16, no. 8, pp. 1–14, 2019.
- [22] X. Liao, L. Ding, and Y. Wang, “Secure machine learning, a brief overview,” in 2011 Fifth International Conference on Secure Software Integration and Reliability Improvement - Companion, 2011, pp. 26– 29.
- [23] M. Barreno, B. Nelson, R. Sears, A. D. Joseph, and J. D. Tygar, “Can machine learning be secure?” in *ACM Symposium on Information, computer and communications security*, 2006, pp. 16–25.
- [24] [24] F. O. Olowononi, D. B. Rawat, and C. Liu, “Resilient machine learning for net- worked cyber physical systems: A survey for machine learning security to securing machine learning for cps,” *IEEE Communications Surveys Tutorials*, vol. 23, no. 1, pp. 524–552, 2021.
- [25] A. Qayyum, M. Usama, J. Qadir, and A. Al-Fuqaha, “Securing connected autonomous vehicles: Challenges posed by adversarial ma- chine learning and the way forward,” *IEEE Communications Surveys Tutorials*, vol. 22, no. 2, pp. 998–1026, 2020.
- [26] N. Akhtar and A. Mian, “Threat of adversarial attacks on deep learning in computer vi- sion: A survey,” *IEEE Access*, vol. 6, pp. 14 410–14 430, 2018.
- [27] H. Xu, Y. Ma, H. Liu, D. Deb, H. Liu, J. Tang, and A. Jain, “Adversarial attacks and de- fenses in images, graphs and text: A review,” *International Journal of Automation and Com- puting*, vol. 17, pp. 151– 178, 2020.
- [28] B. Kim, Y. E. Sagduyu, K. Davaslioglu, T. Erpek, and S. Ulukus, “Channel-aware adver- sarial attacks against deep learning-based wire- less signal classifiers,” *arXiv preprint arXiv:2005.05321*, 2020.
- [29] ———, “Over-the-air adversarial attacks on deep learning based modulation classifier over wireless channels,” in *2020 54th Annual Conference on Information Sciences and Systems (CISS)*, 2020.

- [30] B. Kim, Y. E. Sagduyu, T. Erpek, K. Davaslioglu, and S. Ulukus, "Adversarial attacks with multiple antennas against deep learning-based modulation classifiers," in IEEE Globecom Workshops, 2020.
- [31] M. Z. Hameed, A. Gyorgy, and D. Gunduz, "The best defense is a good offense: Adversarial attacks to avoid modulation detection," IEEE Transactions on Information Forensics and Security, vol. 16, pp. 1074–1087, 2020.
- [32] F. Restuccia, S. D'Oro, A. Al-Shawabka, B. C. Rendon, K. Chowdhury, S. Ioannidis, and T. Melodia, "Generalized wireless adversarial deep learning," in ACM Workshop on Wireless Security and Machine Learning (WiseML), 2020.
- [33] M. DelVecchio, B. Flowers, and W. C. Headley, "Effects of forward error correction on communications aware evasion attacks," arXiv preprint arXiv:2005.13123, 2020.
- [34] Y. E. Sagduyu, Y. Shi, and T. Erpek, "Adversarial deep learning for over-the-air spectrum poisoning attacks," IEEE Transactions on Mobile Computing, vol. 20, no. 2, pp. 306–319, 2021.
- [35] Y. Lin, H. Zhao, Y. Tu, S. Mao, and Z. Dou, "Threats of adversarial attacks in DNN-based modulation recognition," in IEEE Conference on Computer Communications (INFOCOM), 2020.
- [36] Y. E. Sagduyu, Y. Shi, T. Erpek, W. Headley, B. Flowers, G. Stantchev, and Z. Lu, "When wireless security meets machine learning: Motivation, challenges, and research directions," arXiv preprint arXiv:2001.08883, 2020.
- [37] D. Roy, T. Mukherjee, and M. Chatterjee, "Machine learning in adversarial RF environments," IEEE Communications Magazine, vol. 57, no. 5, pp. 82–87, 2019.
- [38] D. Adesina, J. Bassey, and L. Qian, "Robust deep radio frequency spectrum learning for future wireless communications systems," IEEE Access, vol. 8, pp. 148 528–148 540, 2020.
- [39] D. Adesina., J. Bassey, and L. Qian, "Practical radio frequency learning for future wireless communication systems," in IEEE Military Communications Conference (MILCOM), 2019.

- [40] Y. Shi, K. Davaslioglu, Y. E. Sagduyu, W. C. Headley, M. Fowler, and G. Green, "Deep learning for RF signal classification in unknown and dynamic spectrum environments," in IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN), 2019.
- [41] J. Deogirikar and A. Vidhate, "Security attacks in IoT: A survey," in 2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2017, pp. 32–37.
- [42] F. Shahzad, M. Pasha, and A. Ahmad, "A survey of active attacks on wireless sensor networks and their countermeasures," 2017.
- [43] S. Jaitly, H. Malhotra, and B. Bhushan, "Security vulnerabilities and countermeasures against jamming attacks in wireless sensor networks: A survey," in 2017 International Conference on Computer, Communications and Electronics (Comptelix), 2017, pp. 559–564.
- [44] S. Vadlamani, B. Eksioglu, H. Medal, and A. Nandi, "Jamming attacks on wireless networks: A taxonomic survey," International Journal of Production Economics, vol. 172, pp. 76–94, 2016.
- [45] Y. Wu, A. Khisti, C. Xiao, G. Caire, K.-K. Wong, and X. Gao, "A survey of physical layer security techniques for 5G wireless networks and challenges ahead," IEEE Journal on Selected Areas in Communications, vol. 36, no. 4, pp. 679–695, 2018.
- [46] I. Stellios, P. Kotzanikolaou, M. Psarakis, C. Alcaraz, and J. Lopez, "A survey of iot-enabled cyberattacks: Assessing attack paths to critical infrastructures and services," IEEE Communications Surveys Tutorials, vol. 20, no. 4, pp. 3453–3495, 2018.
- [47] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances, and future trends," Proceedings of the IEEE, vol. 104, no. 9, pp. 1727–1765, 2016.
- [48] M. Arif, G. Wang, M. Zakirul Alam Bhuiyan, T. Wang, and J. Chen, "A survey on security attacks in vanets: Communication, applications and challenges," Vehicular Communications, vol. 19, 2019.
- [49] P. Dewal, G. S. Narula, V. Jain, and A. Baliyan, "Security attacks in wireless sensor networks: A survey," in Cyber Security, M. U. Bokhari, N. Agrawal, and D. Saini, Eds. Singapore: Springer Singapore, 2018, pp. 47–58.

- [50] M. Parashar, A. Poonia, and K. Satish, "A survey of attacks and their mitigations in software defined networks," in 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2019, pp. 1–8.
- [51] Y. E. Sagduyu, Y. Shi, and T. Erpek, "IoT network security from the perspective of adversarial deep learning," in IEEE International Conference on Sensing, Communication, and Networking (SECON), 2019.
- [52] Y. Shi, Y. E. Sagduyu, T. Erpek, K. Davaslioglu, Z. Lu, and J. H. Li, "Adversarial deep learning for cognitive radio security: Jamming attack and defense strategies," in IEEE International Conference on Communications Workshops (ICC Workshops), 2018.
- [53] Y. Shi, T. Erpek, Y. E. Sagduyu, and J. H. Li, "Spectrum data poisoning with adversarial deep learning," in IEEE Military Communications Conference (MILCOM), 2018.
- [54] M. Sadeghi. and E. G. Larsson, "Physical adversarial attacks against end-to-end autoencoder communication systems," IEEE Communications Letters, vol. 23, no. 5, pp. 847–850, 2019.
- [55] B. Flowers, R. M. Buehrer, and W. C. Headley, "Evaluating adversarial evasion attacks in the context of wireless communications," IEEE Transactions on Information Forensics and Security, vol. 15, pp. 1102– 1113, 2020.
- [56] F. Restuccia, S. D'Oro, A. Al-Shawabka, B. C. Rendon, K. Chowdhury, S. Ioannidis, and T. Melodia, "Hacking the waveform: Generalized wireless adversarial deep learning," arXiv preprint arXiv:2005.02270, 2020.
- [57] S. Kokalj-Filipovic, R. Miller, and J. Morman, "Targeted adversarial examples against RF deep classifiers," in ACM Workshop on Wireless Security and Machine Learning (WiseML), 2019.
- [58] M. DelVecchio, V. Arndorfer, and W. C. Headley, "Investigating a spectral deception loss metric for training machine learning-based evasion attacks," arXiv preprint arXiv:2005.13124, 2020.
- [59] B. Kim, Y. E. Sagduyu, T. Erpek, K. Davaslioglu, and S. Ulukus, "Channel effects on surrogate models of adversarial attacks against wireless signal classifiers," in IEEE International Conference on Communications, 2021.

- [60] C. Zhong, F. Wang, M. C. Gursoy, and S. Velipasalar, “Adversarial jamming attacks on deep reinforcement learning based dynamic multichannel access,” in IEEE Wireless Communications and Networking Conference (WCNC), 2020.
- [61] F. Wang, C. Zhong, M. C. Gursoy, and S. Velipasalar, “Adversarial jamming attacks and defense strategies via adaptive deep reinforcement learning,” arXiv preprint arXiv:2007.06055, 2020.
- [62] Y. Shi, Y. E. Sagduyu, T. Erpek, and M. C. Gursoy, “How to attack and defend 5G radio access network slicing with reinforcement learning,” arXiv preprint arXiv:2101.05768, 2021.
- [63] Y. Shi and Y. E. Sagduyu, “Adversarial machine learning for flooding attacks on 5G radio access network slicing,” in IEEE International Conference on Communications Workshops (ICC Workshops), 2021.
- [64] S. Bhambri, S. Muku, A. Tulasi, and A. B. Buduru, “A survey of black-box adversarial attacks on computer vision models,” 2020.
- [65] [65] X. Yuan, P. He, Q. Zhu, and X. Li, “Adversarial examples: Attacks and defenses for deep learning,” IEEE Transactions on Neural Networks and Learning Systems, vol. 30, no. 9, pp. 2805–2824, 2019.
- [66] I. Tomi and J. A. McCann, “A survey of potential security issues in existing wireless sensor network protocols,” IEEE Internet of Things Journal, vol. 4, no. 6, pp. 1910–1923, 2017.
- [67] B. Biggio and F. Roli, “Wild patterns: Ten years after the rise of adversarial machine learning,” Pattern Recognition, vol. 84, p. 317331, Dec 2018.
- [68] L. Huang, A. D. Joseph, B. Nelson, B. I. Rubinstein, and J. D. Tygar, “Adversarial machine learning,” in Proceedings of the 4th ACM Workshop on Security and Artificial Intelligence, ser. AISec ’11. New York, NY, USA: Association for Computing Machinery, 2011, p. 4358.
- [69] A. Chakraborty, M. Alam, V. Dey, A. Chattopadhyay, and D. Mukhopadhyay, “Adversarial attacks and defences: A survey,” 2018.
- [70] E. Nowroozi, A. Dehghantanha, R. M. Parizi, and K.-K. R. Choo, “A survey of machine learning techniques in adversarial image forensics,” Computers and Security, vol. 100, 2021.

- [71] G. R. Machado, E. Silva, and R. R. Goldschmidt, “Adversarial machine learning in image classification: A survey towards the defender’s perspective,” 2020.
- [72] A. Serban, E. Poll, and J. Visser, “Adversarial examples on object recognition: A comprehensive survey,” *ACM Comput. Surv.*, vol. 53, no. 3, Jun. 2020.
- [73] L. Sun, Y. Dou, C. Yang, J. Wang, P. S. Yu, L. He, and B. Li, “Adversarial attack and defense on graph data: A survey,” 2020.
- [74] Y. Zou, J. Zhu, L. Yang, Y. Liang, and Y. Yao, “Securing physical-layer communications for cognitive radio networks,” *IEEE Communications Magazine*, vol. 53, no. 9, pp. 48–54, 2015.
- [75] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, “A survey on wireless security: Technical challenges, recent advances, and future trends,” *Proceedings of the IEEE*, vol. 104, no. 9, pp. 1727–1765, 2016.
- [76] Y. E. Sagduyu, “Securing cognitive radio networks with dynamic trust against spectrum sensing data falsification,” in *IEEE Military Communications Conference*, 2014.
- [77] Y. E. Sagduyu, R. A. Berry, and A. Ephremides, “Jamming games in wireless networks with incomplete information,” *IEEE Communications Magazine*, vol. 49, no. 8, pp. 112–118, 2011.
- [78] ———, “Wireless jamming attacks under dynamic traffic uncertainty,” in *International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, 2010.
- [79] S. Rajendran, W. Meert, V. Lenders, and S. Pollin, “Saife: Unsupervised wireless spectrum anomaly detection with interpretable features,” in *IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, 2018.
- [80] N. Van Huynh, D. N. Nguyen, D. T. Hoang, and E. Dutkiewicz, “Jam me if you can: Defeating jammer with deep dueling neural network architecture and ambient backscattering augmented communications,” *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 11, pp. 2603–2620, 2019.

- [81] N. Abuzainab, T. Erpek, K. Davaslioglu, Y. E. Sagduyu, Y. Shi, S. J. Mackey, M. Patel, F. Panettieri, M. A. Qureshi, V. Isler et al., “QoS and jamming-aware wireless networking using deep reinforcement learning,” in IEEE Military Communications Conference (MIL-COM), 2019.
- [82] Y. Zhou, M. Kantarcioglu, and B. Xi, “A survey of game theoretic approach for adversarial machine learning,” Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 9, no. 3, p. e1259, 2019.
- [83] X. Wang, J. Li, X. Kuang, Y. an Tan, and J. Li, “The security of machine learning in an adversarial setting: A survey,” Journal of Parallel and Distributed Computing, vol. 130, pp. 12–23, 2019.
- [84] G. Li, P. Zhu, J. Li, Z. Yang, N. Cao, and Z. Chen, “Security matters: A survey on adversarial machine learning,” arXiv preprint arXiv:1810.07339, 2018.
- [85] M. Ozdag, “Adversarial attacks and defenses against deep neural networks: A survey,” Procedia Computer Science, vol. 140, pp. 152–161, 2018, cyber Physical Systems and Deep Learning Chicago, Illinois November 5-7, 2018.
- [86] Q. Liu, P. Li, W. Zhao, W. Cai, S. Yu, and V. C. M. Leung, “A survey on security threats and defensive techniques of machine learning: A data driven view,” IEEE Access, vol. 6, pp. 12 103–12 117, 2018.
- [87] N. Pitropakis, E. Panaousis, T. Giannetsos, E. Anastasiadis, and G. Loukas, “A taxonomy and survey of attacks against machine learning,” Computer Science Review, vol. 34, p. 100199, 2019.
- [88] S. H. Silva and P. Najafirad, “Opportunities and challenges in deep learning adversarial robustness: A survey,” 2020.
- [89] R. R. Wiyatno, A. Xu, O. Dia, and A. de Berker, “Adversarial examples in modern machine learning: A review,” 2019.
- [90] Q. Liu, J. Guo, C. Wen, and S. Jin, “Adversarial attack on DL- based massive MIMO CSI feedback,” Journal of Communications and Networks, vol. 22, no. 3, pp. 230–235, 2020.

- [91] M. G. Kibria, K. Nguyen, G. P. Villardi, O. Zhao, K. Ishizu, and F. Kojima, "Big data analytics, machine learning, and artificial intelligence in next-generation wireless networks," *IEEE Access*, vol. 6, pp. 32 328–32 338, 2018.
- [92] M. Kulin, T. Kazaz, E. De Poorter, and I. Moerman, "A survey on machine learning-based performance improvement of wireless networks: Phy, mac and network layer," *Electronics*, vol. 10, no. 3, 2021.
- [93] S. Zheng, S. Chen, L. Yang, J. Zhu, Z. Luo, J. Hu, and X. Yang, "Big data processing architecture for radio signals empowered by deep learning: Concept, experiment, applications and challenges," *IEEE Access*, vol. 6, pp. 55 907–55 922, 2018.
- [94] L. Qian, J. Zhu, and S. Zhang, "Survey of wireless big data," *Journal of Communications and Information Networks*, vol. 2, no. 1, pp. 1–18, 2017. [Online]. Available: <http://dx.doi.org/10.1007/s41650-017-0001-2>
- [95] L. Dai, R. Jiao, F. Adachi, H. V. Poor, and L. Hanzo, "Deep learning for wireless communications: An emerging interdisciplinary paradigm," *IEEE Communications Magazine*, vol. PP, 05 2020.
- [96] Proakis, *Digital Communications 5th Edition*. McGraw Hill, 2007.
- [97] M. Kulin, T. Kazaz, I. Moerman, and E. D. Poorter, "End-to-end learning from spectrum data: A deep learning approach for wireless signal identification in spectrum monitoring applications," *IEEE Access*, vol. 6, pp. 18 484–18 501, 2018.
- [98] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, <http://www.deeplearningbook.org>.
- [99] Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wireless networks: Key techniques and open issues," *IEEE Communications Surveys Tutorials*, vol. 21, no. 4, pp. 3072–3108, 2019.
- [100] L. Dai, R. Jiao, F. Adachi, H. V. Poor, and L. Hanzo, "Deep learning for wireless communications: An emerging interdisciplinary paradigm," *IEEE Wireless Communications*, vol. 27, no. 4, pp. 133–139, 2020.

- [101] P. V. Klaine, M. A. Imran, O. Onireti, and R. D. Souza, "A survey of machine learning techniques applied to self-organizing cellular networks," *IEEE Communications Surveys Tutorials*, vol. 19, no. 4, pp. 2392–2431, 2017.
- [102] D. C. Nguyen, P. Cheng, M. Ding, D. Lopez-Perez, P. N. Pathirana, J. Li, A. Seneviratne, Y. Li, and H. V. Poor, "Enabling ai in future wire- less networks: A data life cycle perspective," *IEEE Communications Surveys Tutorials*, vol. 23, no. 1, pp. 553–595, 2021.
- [103] J. Jagannath, N. Polosky, A. Jagannath, F. Restuccia, and T. Melodia, "Machine learning for wireless communications in the internet of things: A comprehensive survey," *Ad Hoc Networks*, vol. 93, 2019.
- [104] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Artificial neural networks-based machine learning for wireless networks: A tutorial," *IEEE Communications Surveys Tutorials*, vol. 21, no. 4, pp. 3039–3071, 2019.
- [105] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Communications*, vol. 24, no. 2, pp. 98–105, 2017.
- [106] G. Y. Li, W. Saad, A. Ozgur, P. Kairouz, Z. Qin, J. Hoydis, Z. Han, D. Gunduz, and J. Elmirghani, "Series editorial: The second issue of the series on machine learning in communications and networks," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 7, pp. 1855– 1857, 2021.
- [107] T. Huynh-The, C.-H. Hua, V.-S. Doan, Q.-V. Pham, T.-V. Nguyen, and D.-S. Kim, "Deep learning for constellation-based modulation classification under multipath fading channels," in *2020 International Conference on Information and Communication Technology Convergence (ICTC)*, 2020, pp. 300–304.
- [108] Y. Kumar, M. Sheoran, G. Jajoo, and S. K. Yadav, "Automatic modulation classification based on constellation density using deep learning," *IEEE Communications Letters*, vol. 24, no. 6, pp. 1275–1278, 2020.
- [109] S. Peng, H. Jiang, H. Wang, H. Alwageed, Y. Zhou, M. M. Sebdani, and Y.-D. Yao, "Modulation classification based on signal constellation diagrams and deep learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 3, pp. 718–727, 2019.

- [110] T. J. O'Shea, J. Corgan, and T. C. Clancy, "Convolutional radio modulation recognition networks," 2016.
- [111] T. J. O'Shea, T. Roy, and T. C. Clancy, "Over-the-air deep learning based radio signal classification," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 168–179, Feb 2018.
- [112] S. Rajendran, W. Meert, D. Giustiniano, V. Lenders, and S. Pollin, "Deep learning models for wireless signal classification with distributed low-cost spectrum sensors," *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 3, pp. 433–445, 2018.
- [113] S. Peng, S. Sun, and Y.-D. Yao, "A survey of modulation classification using deep learning: Signal representation and data preprocessing," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–19, 2021.
- [114] R. Zhou, F. Liu, and C. W. Gravelle, "Deep learning for modulation recognition: A survey with a demonstration," *IEEE Access*, vol. 8, pp. 67 366–67 376, 2020.
- [115] S. Zheng, S. Chen, P. Qi, H. Zhou, and X. Yang, "Spectrum sensing based on deep learning classification for cognitive radios," *China Communications*, vol. 17, no. 2, pp. 138–148, 2020.
- [116] J. Xie, J. Fang, C. Liu, and X. Li, "Deep learning-based spectrum sensing in cognitive radio: A cnn-lstm approach," *IEEE Communications Letters*, vol. 24, no. 10, pp. 2196–2200, 2020.
- [117] W. Lee, M. Kim, and D.-H. Cho, "Deep cooperative sensing: Cooperative spectrum sensing based on convolutional neural networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 3, pp. 3005–3009, 2019.
- [118] W. M. Lees, A. Wunderlich, P. J. Jeavons, P. D. Hale, and M. R. Souryal, "Deep learning classification of 3.5-ghz band spectrograms with applications to spectrum sensing," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 2, pp. 224–236, 2019.

- [119] C. Liu, J. Wang, X. Liu, and Y.-C. Liang, “Deep cm-cnn for spectrum sensing in cognitive radio,” *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 10, pp. 2306–2321, 2019.
- [120] J. Gao, X. Yi, C. Zhong, X. Chen, and Z. Zhang, “Deep learning for spectrum sensing,” *IEEE Wireless Communications Letters*, vol. 8, no. 6, pp. 1727–1730, 2019.
- [121] J. Xie, C. Liu, Y.-C. Liang, and J. Fang, “Activity pattern aware spectrum sensing: A cnn-based deep learning approach,” *IEEE Communications Letters*, vol. 23, no. 6, pp. 1025–1028, 2019.
- [122] Y. Arjoune and N. Kaabouch, “A comprehensive survey on spectrum sensing in cognitive radio networks: Recent advances, new challenges, and future research directions,” *Sensors*, vol. 19, no. 1, 2019. [Online]. Available: <https://www.mdpi.com/1424-8220/19/1/126>
- [123] V. Ramani and S. K. Sharma, “Cognitive radios: A survey on spectrum sensing, security and spectrum handoff,” *China Communications*, vol. 14, no. 11, pp. 185–208, 2017.
- [124] M. Eisen and A. Ribeiro, “Optimal wireless resource allocation with random edge graph neural networks,” *IEEE Transactions on Signal Processing*, vol. 68, pp. 2977–2991, 2020.
- [125] H. Sun, X. Chen, Q. Shi, M. Hong, X. Fu, and N. D. Sidiropoulos, “Learning to optimize: Training deep neural networks for wireless resource management,” in *2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2017, pp. 1–6.
- [126] Y. Zhou, Z. M. Fadlullah, B. Mao, and N. Kato, “A deep-learning-based radio resource assignment technique for 5G ultra dense networks,” *IEEE Network*, vol. 32, no. 6, pp. 28–34, 2018.
- [127] J. Guo and C. Yang, “Predictive resource allocation with deep learning,” in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*, 2018, pp. 1–7.
- [128] J. Gao, M. R. A. Khandaker, F. Tariq, K.-K. Wong, and R. T. Khan, “Deep neural network based resource allocation for v2x communications,” in *2019 IEEE 90th Vehicular Technology Conference (VTC2019- Fall)*, 2019, pp. 1–5.

- [129] Z. Xu, Y. Wang, J. Tang, J. Wang, and M. C. Gursoy, "A deep reinforcement learning based framework for power-efficient resource allocation in cloud rans," in 2017 IEEE International Conference on Communications (ICC), 2017, pp. 1–6.
- [130] K. I. Ahmed, H. Tabassum, and E. Hossain, "Deep learning for radio resource allocation in multi-cell networks," *IEEE Network*, vol. 33, no. 6, pp. 188–195, 2019.
- [131] H. Zhang, M. Feng, K. Long, G. K. Karagiannidis, and A. Nallanathan, "Artificial intelligence-based resource allocation in ultradense networks: Applying event-triggered q-learning algorithms," *IEEE Vehicular Technology Magazine*, vol. 14, no. 4, pp. 56–63, 2019.
- [132] U. Challita, L. Dong, and W. Saad, "Deep learning for proactive resource allocation in lte-u networks," in *European Wireless 2017; 23th European Wireless Conference*, 2017, pp. 1–6.
- [133] D. C. Nguyen, P. Cheng, M. Ding, D. Lopez-Perez, P. N. Pathirana, J. Li, A. Seneviratne, Y. Li, and H. V. Poor, "Enabling ai in future wire- less networks: A data life cycle perspective," *IEEE Communications Surveys Tutorials*, vol. 23, no. 1, pp. 553–595, 2021.
- [134] Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wireless networks: Key techniques and open issues," *IEEE Communications Surveys Tutorials*, vol. 21, no. 4, pp. 3072–3108, 2019.
- [135] I. Ahmad, S. Shahabuddin, H. Malik, E. Harjula, T. Leppnen, L. Lovn, Anttonen, A. H. Sodhro, M. Mahtab Alam, M. Juntti, A. YlIjski, T. Sauter, A. Gurtov, M. Ylianttila, and J. Riekkki, "Machine learning meets communication networks: Current trends and future challenges," *IEEE Access*, vol. 8, pp. 223 418–223 460, 2020.
- [136] L. Zhang, Y.-C. Liang, and D. Niyato, "6g visions: Mobile ultra- broadband, super internet-of-things, and artificial intelligence," *China Communications*, vol. 16, no. 8, pp. 1–14, 2019.
- [137] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Artificial neural networks-based machine learning for wireless networks: A tutorial," *IEEE Communications Surveys Tutorials*, vol. 21, no. 4, pp. 3039–3071, 2019.

- [138] J. Wang, C. Jiang, H. Zhang, Y. Ren, K.-C. Chen, and L. Hanzo, “Thirty years of machine learning: The road to pareto-optimal wireless networks,” *IEEE Communications Surveys Tutorials*, vol. 22, no. 3, pp. 1472–1514, 2020.
- [139] M. Sadeghi and E. G. Larsson, “Adversarial attacks on deep-learning based radio signal classification,” *IEEE Wireless Communications Letters*, vol. 8, no. 1, pp. 213–216, 2019.
- [140] X. Wu, D. Sahoo, and S. C. Hoi, “Recent advances in deep learning for object detection,” *Neurocomputing*, vol. 396, pp. 39–64, 2020.
- [141] S. S. Islam, S. Rahman, M. M. Rahman, E. K. Dey, and M. Shoyaib, “Application of deep learning to computer vision: A comprehensive study,” in *2016 5th international conference on informatics, electronics and vision (ICIEV)*. IEEE, 2016, pp. 592–597.
- [142] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, “Deep learning for computer vision: A brief review,” *Computational intelligence and neuroscience*, vol. 2018, 2018.
- [143] M. Hassaballah and A. I. Awad, *Deep learning in computer vision: principles and applications*. CRC Press, 2020.
- [144] J. Ni, T. Young, V. Pandealea, F. Xue, V. Adiga, and E. Cambria, “Recent advances in deep learning based dialogue systems: A systematic survey,” 2021.
- [145] G. Weisz, P. Budzianowski, P.-H. Su, and M. Gašić, “Sample efficient deep reinforcement learning for dialogue systems with large action spaces,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 11, pp. 2083–2097, 2018.
- [146] M. Korpusik and J. Glass, “Deep learning for database mapping and asking clarification questions in dialogue systems,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 8, pp. 1321–1334, 2019.
- [147] Y.-N. Chen, A. Celikyilmaz, and D. Hakkani-Tur, “Deep learning for dialogue systems,” in *Proceedings of the 27th International Conference on Computational Linguistics: Tutorial Abstracts*, 2018, pp. 25–31.
- [148] P. Beckmann, M. Kegler, H. Saltini, and M. Cernak, “Speech-vgg: A deep feature extractor for speech processing,” *arXiv preprint arXiv:1910.09909*, 2019.

- [149] S. Watanabe, T. Hori, S. Karita, T. Hayashi, J. Nishitoba, Y. Unno, N. E. Y. Soplin, J. Heymann, M. Wiesner, N. Chen et al., “Espnet: End- to-end speech processing toolkit,” arXiv preprint arXiv:1804.00015, 2018.
- [150] H. M. Fayek, M. Lech, and L. Cavedon, “Evaluating deep learning ar- chitectures for speech emotion recognition,” *Neural Networks*, vol. 92, pp. 60–68, 2017.
- [151] R. Haeb-Umbach, S. Watanabe, T. Nakatani, M. Bacchiani, Hoffmeister, M. L. Seltzer, H. Zen, and M. Souden, “Speech processing for digital home assistants: Combining signal processing with deep-learning techniques,” *IEEE Signal processing magazine*, vol. 36, no. 6, pp. 111–124, 2019.
- [152] D. Zu“gner, A. Akbarnejad, and S. Gu“nnemann, “Adversarial attacks on neural networks for graph data,” in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery Data Mining*, ser. KDD ’18. New York, NY, USA: Association for Computing Machinery, 2018, p. 28472856. [Online]. Available: <https://doi.org/10.1145/3219819.3220078>
- [153] L. Chen, J. Li, J. Peng, T. Xie, Z. Cao, K. Xu, X. He, Z. Zheng, and Wu, “A survey of adversarial learning on graphs,” 2020. [Online]. Available: <https://arxiv.org/abs/2003.05730>
- [154] X. Zhang and M. Zitnik, “Gnnguard: Defending graph neural networks against adversarial attacks,” in *Proceedings of Neural Information Processing Systems, NeurIPS*, 2020.
- [155] J. Xu, J. Chen, S. You, Z. Xiao, Y. Yang, and J. Lu, “Robustness of deep learning models on graphs: A survey,” *AI Open*, vol. 2, pp. 69–78, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2666651021000139>
- [156] B. Tang, Z. Pan, K. Yin, and A. Khateeb, “Recent advances of deep learning in bioinformatics and computational biology,” *Frontiers in Genetics*, vol. 10, p. 214, 2019. [Online]. Available: <https://www.frontiersin.org/article/10.3389/fgene.2019.00214>
- [157] A. Yazdani, L. Lu, M. Raissi, and G. E. Karniadakis, “Systems biology informed deep learning for inferring parameters and hidden dynamics,” *PLoS computational biology*, vol. 16, no. 11, p. e1007575, 2020.

- [158] G. Eraslan, Z. Avsec, J. Gagneur, and F. J. Theis, “Deep learning: new computational modelling techniques for genomics,” *Nature Reviews Genetics*, vol. 20, no. 7, pp. 389–403, 2019.
- [159] L. Rampasek and A. Goldenberg, “Tensorflow: biologys gateway to deep learning?” *Cell systems*, vol. 2, no. 1, pp. 12–14, 2016.
- [160] B. Shickel, P. J. Tighe, A. Bihorac, and P. Rashidi, “Deep ehr: A survey of recent advances in deep learning techniques for electronic health record (ehr) analysis,” *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 5, pp. 1589–1604, 2018.
- [161] A. Rajkomar, E. Oren, K. Chen, A. M. Dai, N. Hajaj, M. Hardt, P. J. Liu, X. Liu, J. Marcus, M. Sun et al., “Scalable and accurate deep learning with electronic health records,” *NPJ Digital Medicine*, vol. 1, no. 1, pp. 1–10, 2018.
- [162] B. Norgeot, B. S. Glicksberg, L. Trupin, D. Lituiev, M. Gianfrancesco, Oskotsky, G. Schmajuk, J. Yazdany, and A. J. Butte, “Assessment of a deep learning model based on electronic health record data to forecast clinical outcomes in patients with rheumatoid arthritis,” *JAMA network open*, vol. 2, no. 3, pp. e190 606–e190 606, 2019.
- [163] D. Liu, T. Miller, R. Sayeed, and K. D. Mandl, “Fadl: Federated- autonomous deep learning for distributed electronic health record,” *arXiv preprint arXiv:1811.11400*, 2018.
- [164] A. B. Levine, C. Schlosser, J. Grewal, R. Coope, S. J. Jones, and S. Yip, “Rise of the machines: Advances in deep learning for cancer diagnosis,” *Trends in Cancer*, vol. 5, no. 3, pp. 157–169, 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2405803319300184>
- [165] A. G. C. Pacheco and R. A. Krohling, “Recent advances in deep learning applied to skin cancer detection,” 2019.
- [166] Y. Xiao, J. Wu, and Z. Lin, “Cancer diagnosis using generative adversarial networks based on deep learning from imbalanced data,” *Computers in Biology and Medicine*, p. 104540, 2021.
- [167] U. Kose and J. Alzubi, *Deep learning for cancer diagnosis*. Springer, 2021.

- [168] K. Yu, L. Tan, L. Lin, X. Cheng, Z. Yi, and T. Sato, “Deep-learning- empowered breast cancer auxiliary diagnosis for 5GB remote e-health,” *IEEE Wireless Communications*, vol. 28, no. 3, pp. 54–61, 2021.
- [169] W. Sun, B. Zheng, and W. Qian, “Computer aided lung cancer diagnosis with deep learning algorithms,” in *Medical imaging 2016: computer- aided diagnosis*, vol. 9785. International Society for Optics and Photonics, 2016, p. 97850Z.
- [170] A. Ioannidou, E. Chatzilari, S. Nikolopoulos, and I. Kompatsiaris, “Deep learning advances in computer vision with 3d data: A survey,” *ACM Comput. Surv.*, vol. 50, no. 2, apr 2017. [Online]. Available: <https://doi.org/10.1145/3042064>
- [171] C. R. Qi, “Deep learning on 3d data,” in *3D Imaging, Analysis and Applications*. Springer, 2020, pp. 513–566.
- [172] X. Chai, G. Tang, S. Wang, K. Lin, and R. Peng, “Deep learning for irregularly and regularly missing 3-d data reconstruction,” *IEEE Transactions on Geoscience and Remote Sensing*, 2020.
- [173] Q.-Y. Zhou, J. Park, and V. Koltun, “Open3d: A modern library for 3d data processing,” *arXiv preprint arXiv:1801.09847*, 2018.
- [174] X. Zhou and S. Prasad, “Advances in deep learning for hyperspectral image analysis addressing challenges arising in practical imaging scenarios,” *Advances in Computer Vision and Pattern Recognition*, p. 117140, 2020.
- [175] A. Signoroni, M. Savardi, A. Baronio, and S. Benini, “Deep learning meets hyperspectral image analysis: A multidisciplinary review,” *Journal of Imaging*, vol. 5, no. 5, p. 52, 2019.
- [176] S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi, and J. A. Benediktsson, “Deep learning for hyperspectral image classification: An overview,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 6690–6709, 2019.
- [177] X. Yang, Y. Ye, X. Li, R. Y. Lau, X. Zhang, and X. Huang, “Hyperspectral image classification with deep learning models,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 9, pp. 5408–5423, 2018.

- [178] Z. Gong, P. Zhong, and W. Hu, "Statistical loss and analysis for deep learning in hyperspectral image classification," *IEEE transactions on neural networks and learning systems*, vol. 32, no. 1, pp. 322–333, 2020.
- [179] S. Kokalj-Filipovic, R. Miller, N. Chang, and C. L. Lau, "Mitigation of adversarial examples in RF deep classifiers utilizing autoencoder pre-training," in *International Conference on Military Communications and Information Systems (ICMCIS)*, 2019.
- [180] K. Davaslioglu and Y. E. Sagduyu, "Trojan attacks on wireless signal classification with adversarial machine learning," in *IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, 2019.
- [181] S. Bair, M. DelVecchio, B. Flowers, A. J. Michaels, and W. C. Headley, "On the limitations of targeted adversarial evasion attacks against deep learning enabled modulation recognition," in *ACM Workshop on Wireless Security and Machine Learning (WiseML)*, 2019.
- [182] D. Ke, Z. Huang, X. Wang, and L. Sun, "Application of adversarial examples in communication modulation classification," in *International Conference on Data Mining Workshops (ICDMW)*, 2019.
- [183] M. Z. Hameed, A. Gyoergy, and D. Gu"ndu"z, "Communication without interception: Defense against modulation detection," *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2019.
- [184] S. Kokalj-Filipovic and R. Miller, "Adversarial examples in RF learning: detection of the attack and its physical robustness," *arXiv preprint arXiv:1902.06044*, 2019.
- [185] S. Karunaratne, E. Krijestorac, and D. Cabric, "Penetrating RF fingerprinting-based authentication with a generative adversarial attack," *arXiv preprint arXiv:2011.01538*, 2020.
- [186] F. Wang, C. Zhong, M. C. Gursoy, and S. Velipasalar, "Defense strategies against adversarial jamming attacks via deep reinforcement learning," in *Conference on Information Sciences and Systems (CISS)*, 2020.
- [187] Y. Shi, K. Davaslioglu, and Y. E. Sagduyu, "Over-the-air membership inference attacks as privacy threats for deep learning-based wireless signal classifiers," in *ACM Workshop on Wireless Security and Machine Learning (WiseML)*, 2020.

- [188] Y. Shi and Y. E. Sagduyu, “Membership inference attack and defense for wireless signal classifiers with deep learning,” arXiv preprint arXiv:2107.12173, 2021.
- [189] B. Kim, Y. E. Sagduyu, K. Davaslioglu, T. Erpek, and S. Ulukus, “How to make 5G communications” invisible”: Adversarial machine learning for wireless privacy,” arXiv preprint arXiv:2005.07675, 2020.
- [190] N. Akhtar and A. Mian, “Threat of adversarial attacks on deep learning in computer vision: A survey,” IEEE Access, vol. 6, pp. 14 410–14 430, 2018.
- [191] Y.-J. Cao, L.-L. Jia, Y.-X. Chen, N. Lin, C. Yang, B. Zhang, Z. Liu, X.-X. Li, and H.-H. Dai, “Recent advances of generative adversarial networks in computer vision,” IEEE Access, vol. 7, pp. 14 985–15 006, 2019.
- [192] N. Akhtar, A. Mian, N. Kardan, and M. Shah, “Advances in adversarial attacks and defenses in computer vision: A survey,” IEEE Access, vol. 9, pp. 155 161–155 196, 2021.
- [193] Z. Zou, S. Lei, T. Shi, Z. Shi, and J. Ye, “Deep adversarial decomposition: A unified framework for separating superimposed images,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2020.
- [194] N. Sarafianos, X. Xu, and I. A. Kakadiaris, “Adversarial representation learning for text-to-image matching,” in Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), October 2019.
- [195] H. Wang, D. Gong, Z. Li, and W. Liu, “Decorrelated adversarial learning for age-invariant face recognition,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2019.
- [196] Y. Zhong and W. Deng, “Adversarial learning with margin-based triplet embedding regularization,” in Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), October 2019.
- [197] K. Eykholt, I. Evtimov, E. Fernandes, B. Li, A. Rahmati, C. Xiao, Prakash, T. Kohno, and D. Song, “Robust physical-world attacks on deep learning models,” 2018.
- [198] A. Madry, A. Makelov, L. Schmidt, D. Tsipras, and A. Vladu, “Towards deep learning models resistant to adversarial attacks,” arXiv preprint arXiv:1706.06083, 2017.

- [199] T. Bai, J. Luo, J. Zhao, B. Wen, and Q. Wang, “Recent advances in adversarial training for adversarial robustness,” 2021.
- [200] A. Kurakin, D. Boneh, F. Tramr, I. Goodfellow, N. Papernot, and P. McDaniel, “Ensemble adversarial training: Attacks and defenses,” 2018. [Online]. Available: <https://openreview.net/pdf?id=rkZvSe-RZ>
- [201] T. Bai, J. Luo, J. Zhao, B. Wen, and Q. Wang, “Recent advances in adversarial training for adversarial robustness,” in Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21, Z.-H. Zhou, Ed. International Joint Conferences on
- [202] Artificial Intelligence Organization, 8 2021, pp. 4312–4321, survey Track. [Online]. Available: <https://doi.org/10.24963/ijcai.2021/591>
- [203] M. Qiu and H. Qiu, “Review on image processing based adversarial example defenses in computer vision,” in 2020 IEEE 6th Intl Conference on Big Data Security on Cloud (Big Data Security), IEEE Intl Conference on High Performance and Smart Computing, (HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS), 2020, pp. 94–99.
- [204] F. Feng, X. He, J. Tang, and T.-S. Chua, “Graph adversarial training: Dynamically regularizing based on graph structure,” IEEE Transactions on Knowledge and Data Engineering, vol. 33, no. 6, pp. 2493–2504, 2021.
- [205] Y. Qiao, X. Luo, C. Li, H. Tian, and J. Ma, “Heterogeneous graph- based joint representation learning for users and pois in location-based social network,” Information Processing and Management, vol. 57, p. 102151, 03 2020.
- [206] Q. Tan, N. Liu, and X. Hu, “Deep representation learning for social network analysis,” Frontiers in Big Data, vol. 2, 2019. [Online]. Available: <https://www.frontiersin.org/article/10.3389/fdata.2019.00002>
- [207] W. L. Hamilton, R. Ying, and J. Leskovec, “Inductive representation learning on large graphs,” 2018.
- [208] Abusnaina, A. Khormali, H. Alasmary, J. Park, A. Anwar, and Mohaisen, “Adversarial learning attacks on graph-based iot malware detection systems,” in 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS), 2019, pp. 1296–1305.

- [209] H. Dai, H. Li, T. Tian, X. Huang, L. Wang, J. Zhu, and L. Song, “Adversarial attack on graph structured data,” in International conference on machine learning. PMLR, 2018, pp. 1115–1124.
- [210] Z. Zhang, Z. Zhang, Y. Zhou, Y. Shen, R. Jin, and D. Dou, “Adversarial attacks on deep graph matching,” *Advances in Neural Information Processing Systems*, vol. 33, 2020.
- [211] D. Zügner, A. Akbarnejad, and S. Gunnemann, “Adversarial attacks on neural networks for graph data,” in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2018, pp. 2847–2856.
- [212] S. Pan, R. Hu, S.-F. Fung, G. Long, J. Jiang, and C. Zhang, “Learning graph embedding with adversarial training methods,” *IEEE Transactions on Cybernetics*, vol. 50, no. 6, pp. 2475–2487, 2020.
- [213] H. Wang, J. Wang, J. Wang, M. Zhao, W. Zhang, F. Zhang, X. Xie, and M. Guo, “Graphgan: Graph representation learning with generative adversarial nets,” in *Proceedings of the AAAI conference on artificial intelligence*, vol. 32, no. 1, 2018.
- [214] L. Chen, J. Li, J. Peng, T. Xie, Z. Cao, K. Xu, X. He, and Z. Zheng, “A survey of adversarial learning on graphs,” *arXiv preprint arXiv:2003.05730*, 2020.
- [215] L. Sun, Y. Dou, C. Yang, J. Wang, P. S. Yu, L. He, and B. Li, “Adversarial attack and defense on graph data: A survey,” *arXiv preprint arXiv:1812.10528*, 2018.
- [216] Y. Zhang, S. Khan, and M. Coates, “Comparing and detecting adversarial attacks for graph deep learning,” in *Proc. Representation Learning on Graphs and Manifolds Workshop, Int. Conf. Learning Representations*, New Orleans, LA, USA, 2019.
- [217] S. Minaee, N. Kalchbrenner, E. Cambria, N. Nikzad, M. Chenaghlu, and J. Gao, “Deep learning–based text classification: A comprehensive review,” *ACM Computing Survey*, vol. 54, no. 3, Apr 2021. [Online]. Available: <https://doi.org/10.1145/3439726>
- [218] J. Kim, S. Jang, E. Park, and S. Choi, “Text classification using capsules,” *Neuro-computing*, vol. 376, pp. 214–221, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0925231219314092>

- [219] N. Carlini and D. Wagner, “Audio adversarial examples: Targeted attacks on speech-to-text,” 2018.
- [220] A. Hannun, C. Case, J. Casper, B. Catanzaro, G. Diamos, E. Elsen, R. Prenger, S. Satheesh, S. Sengupta, A. Coates, and A. Y. Ng, “Deep Speech: Scaling up end-to-end Speech Recognition,” 2014.
- [221] T. Miyato, A. M. Dai, and I. Goodfellow, “Adversarial training methods for semi-supervised text classification,” 2021.
- [222] J. Ebrahimi, A. Rao, D. Lowd, and D. Dou, “Hotflip: White-box adversarial examples for text classification,” 2018.
- [223] Liang, H. Li, M. Su, P. Bian, X. Li, and W. Shi, “Deep text classification can be fooled,” in Proceedings of the 27th International Joint Conference on Artificial Intelligence, ser. IJCAI’18. AAAI Press,

APPENDIX A – PUBLICATIONS AND PRESENTATIONS

Publications:

- 1) S. M. Debb, and M. K. McClellan, “Factors Impacting Perceived Vulnerability as a Determinant of Increased Risk For Cyber Security Risk Behavior,” *Cyberpsychology, Behavior, & Social Networking*. (In press).
- 2) D. Adesina, C. Hsieh, Y. E. Sagduyu, and L. Qian, “Adversarial Machine Learning in Wireless Communications using RF Data: A Review,” submitted to *IEEE Communications Surveys & Tutorials*, arXiv:2012.14392. (In revision).
- 3) S. M. Debb, K.J. Haschke, and M. K. McClellan, “Fear of Missing Out Manifests Differently for African American and Caucasian College Students,” Paper to be presented at the 128th annual convention of the American Psychological Association, Virtual Conference, August 12-14, 2021.
- 4) S. M. Debb, K. J. Haschke, and M. K. McClellan, “Assessment of Fear of Missing Out: Comparing Survey Instruments For Use With Diverse Groups,” Poster to be presented at the 128th annual convention of the American Psychological Association, Virtual Conference, August 12-14, 2021.
- 5) K. J. Haschke, M. K. McClellan, and S. M. Debb, “Assessing Fear of Missing Out From a Cross Cultural Lens: Concurrent Validity of FoMO-10 and FoMO-SF,” Poster to be presented at the 128th annual convention of the American Psychological Association, Virtual Conference, August 12-14, 2021.
- 6) M. K. McClellan, K. J. Haschke, and S. M. Debb, “Fueling FoMO: The Relationship Between Anxiety and the Need to Belong,” Poster to be presented at the 128th annual convention of the American Psychological Association, Virtual Conference, August 12-14, 2021.
- 7) Eranga Bandara, Xueping Liang, Peter Foytik, Sachin Shetty, Crissie Hall, Daniel Bowden, Nalin Ranasinghe, and Kasun De Zoysa, "A blockchain empowered and privacy preserving digital contact tracing platform," *Information Processing & Management*, Volume 58, Issue 4, 2021, 102572, ISSN 0306-4573, July 2021.

- 8) Yen-Hung Hu, and Chung-Chu Hsieh, "A Study of Classifying Advanced Persistent Threats With Multi-Layered Deep Learning Approaches," submitted to 2021 Intl Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS2021), July 19-22, 2021, George Mason Univ., Fairfax, VA, July 19-22, 2021.
- 9) Eranga Herath, Deepak Tosh, Pete Foytik, Sachin Shetty, Nalin Ranasinghe, and Kasun De Zoysa, "Tikiri: Towards a Lightweight Blockchain for IoT," Future Generation Computer Systems, Volume 119, June 2021, Pages 154-165.
- 10) Eranga Herath, Xueping Liang, Sachin Shetty, and Pete Foytik, "Rahasak - Scalable Blockchain Architecture for Enterprise Applications," Journal of Systems Architecture, Volume 116, June 2021, 102061.
- 11) Danielle D. Godsey, Yen-Hung (Frank) Hu, and Mary Ann Hoppa, "A Multi-Layered Approach to Fake News Identification, Measurement and Mitigation," Future of Information and Communication Conference (FICC) 2021, Vancouver, Canada, Virtual Conference, April 29-30, 2021.
- 12) M. K. McClellan, K. J. Haschke, and S. M. Debb, "FoMO factors: How personality, emotion, and the need to belong impact the fear of missing out," Research and innovation Symposium and Exhibition (RISE 21), Norfolk State University, Norfolk, VA, Virtual Conference, April 19-22, 2021.
- 13) K. J. Haschke, M. K. McClellan, and S. M. Debb, "Mental health disparities among various cultural groups: Differences in anxiety, depression, affect, and resilience," 14th Xavier University Annual Health Disparities Conference (poster), New Orleans, LA, Virtual Conference, April 7-8, 2021.
- 14) M. K. McClellan, M.K., *Haschke, and S. M. Debb, "The role of gender on perceived stress and psychological well-being," 14th Xavier University Annual Health Disparities Conference, New Orleans, LA, Virtual Conference, April 7-8, 2021.
- 15) M. A. R. A. Amin, Sachin Shetty, Laurent Njilla, Deepak Tosh, and Charles Kamhoua, "Hidden Markov Model and Cyber Deception for the Prevention of Adversarial Lateral Movement," IEEE Access, vol. 9, pp. 49662-49682, 2021, doi: 10.1109/ACCESS.2021.3069105, March 2021.

- 16) Hamdi Kava, Jose Padilla et al., "Simulation for Cybersecurity: State of the Art and Future Directions," *Journal of Cybersecurity*, 7(1), March 2021.
- 17) S. M. Debb, "Enhancing Online Graduate Education Through Interactive Asynchronous Instruction: Utilizing Cyberpsychology To Teach CyberPsychology," *Southeastern Teaching of Psychology Conf.*, Atlanta, GA, Virtual Conference, February 19-20, 2021.
- 18) Luiz A. DaSilva, Jeffrey H. Reed, Sachin Shetty, Jerry Park, Duminda Wijesekera, and Haining Wang, *Securing 5G: A NATO's Role in Collaborative Risk Assessment and Mitigation, Cyber Threats and NATO 2030: Horizon Scanning and Analysis*, 2021.
- 19) Mohammad Wazid, Ashok Kumar Das, Sachin Shetty, and Jo Minho, "A Tutorial and Future Research for Building a Blockchain-Based Secure Communication Scheme for Internet of Intelligent Things," *IEEE Access* 8: 88700-88716 (2020).
- 20) Abhinav Kommula, Yen-Hung (Frank) Hu, Mary Ann Hoppa, and Samuel Olatunbosun, "Machine Learning Techniques to Enhance Container Network Security," in *Proceedings of the 2020 Intl Conf. on Computational Science and Computational Intelligence CSCI'20*, December 16-18, 2020, Las Vegas, NV.
- 21) Charles Hu, and Yen-Hung (Frank) Hu, "Data Poisoning on Deep Learning Models," in *Proceedings of the 2020 Intl Conf. on Computational Science and Computational Intelligence CSCI'20*, December 16-18, 2020, Las Vegas, NV.
- 22) Girish Nandakumar and Jose Padilla, "A Game-Transformation-based Framework to Understand Initial Conditions and Outcomes in the context of Cyber-enabled Influence Operations (CIOs)," *Int'l Conf on Social Computing, Behavior-Cultural Modeling & Prediction and Behavior Representations in Modeling and Simulation*, Washington DC, October 19-22, 2020.
- 23) S. P. Gochhayat, S. Shetty, R. Mukkamala, P. Foytik, G. A. Kamhoua, and L. Njilla, "Measuring Decentrality in Blockchain based Systems," in *IEEE Access*, September 2020.
- 24) Dominicia Williams, Yen-Hung (Frank) Hu, and Mary Ann Hoppa, "Follow the Money Through Apple Pay," *CISSE 2020*, Baltimore, MD, June 2020.
- 25) Dar'rell Pope, Yen-Hung (Frank) Hu, and Mary Ann Hoppa, "A Survey on Securing Personally Identifiable Information on Smartphones," *Virginia Journal of Science* 2020.

- 26) S. M. Debb, D. R. Schaffer, and D. G. Colson, "A reverse digital divide. Comparing information security behaviors of Generation Y and Generation Z adults," *Intl Journal of Cybersecurity Intelligence & Cybercrime*, 3(1), 2020, Pages 42-55. Retrieved from <https://vc.bridgew.edu/ijcic/vol3/iss1/4/>.
- 27) Kjartan Palsson, Steinn Guðmundsson, and Sachin Shetty, "Analysis of the impact of cyber events for cyber insurance," *The Geneva Papers on Risk and Insurance - Issues and Practice*, 45, Pages 564-579, June 2020.
- 28) M. Wazid, A.K. Das, Sachin Shetty, and M. Jo, "A Tutorial and Future Research for Building a Blockchain-Based Secure Communication Scheme for Internet of Intelligent Things," in *IEEE Access*, Vol. 8, pp. 1-17, 2020.
- 29) Muhammad Saad, Jeffrey Spaulding, Laurent Njilla, Charles Kamhoua, Sachin Shetty, DaeHun Nyang, and David Mohaisen, "Exploring the Attack Surface of Blockchain: A Comprehensive Survey," *IEEE Communications Surveys & Tutorials*, 2020.
- 30) Scott M. Debb, "Fear of Missing Out: Correlates and Contributors," 25th Annual CyberPsychology, CyberTherapy & Social Networking Conf., Milan, Italy, June 2020.
- 31) Scott M. Debb, and Daniel R. Shaffer, "Cognitive dissonance, frontal cortex asymmetry, and its effect on cybersecurity password creation," 25th Annual CyberPsychology, CyberTherapy & Social Networking Conf., Milan, Italy, June 2020.
- 32) Charles Kamhoua, Alexander Kott, Laurent Njilla and Sachin Shetty (editors), *Modeling and Design of Secure Internet of Things*, John Wiley & Sons, 1st Edition, June 2020, ISBN 978-1-119-59336-2.
- 1) Charles Kamhoua, Alexander Kott, Laurent Njilla and Sachin Shetty (editors), *Modeling and Design of Secure Internet of Things*, John Wiley & Sons, 1st Edition, June 2020, ISBN 978-1-119-59336-2.
- 2) Girish Nandakumar and Jose Padilla, "A Game-Transformation-based Framework to Understand Initial Conditions and Outcomes in the context of Cyber-enabled Influence Operations," accepted for Int'l Conference on Social Computing, Behavior-Cultural Modeling & Prediction and Behavior Representations in Modeling and Simulation, Washington DC, October 18-21, 2020.

- 3) Jose Padilla et al., "Simulation for Cybersecurity: State of the Art and Future Directions," accepted for publication by Journal of Cybersecurity. Working on revisions.
- 4) Scott M. Debb, "Fear of Missing Out: Correlates and Contributors," accepted for 25th Annual CyberPsychology, CyberTherapy & Social Networking Conference, Milan, Italy, June 2020.
- 5) Scott M. Debb and Daniel R. Shaffer, "Cognitive dissonance, frontal cortex asymmetry, and its effect on cybersecurity password creation," accepted for 25th Annual CyberPsychology, CyberTherapy & Social Networking Conference, Milan, Italy, June 2020.
- 6) Scott Debb, Daniel R. Schaffer and Darlene G. Colson, "A reverse digital divide. Comparing information security behaviors of Generation Y and Generation Z adults," International Journal of Cybersecurity Intelligence & Cybercrime, 3(1), 42-55, February 2020. Retrieved from <https://vc.bridgew.edu/ijcic/vol3/iss1/4/>.
- 7) Daniel R. Schaffer and Scott M. Debb, "Assessing Instagram use across cultures - A confirmatory factor analysis," Cyberpsychology, Behavior & Social Networking, 23(2), February 2020. doi: 10.1089/cyber.2019.0247.
- 8) Sharif Ullah, Sachin Shetty, Amin Hassanzadeh, Anup Nayak and Kamrul Hasan, "On the Effectiveness of Intrusion Response Systems against Persistent Threats," IEEE ICNC, Hawaii, February 17-20, 2020.
- 9) Muhammad Saad, Jeffrey Spaulding, Laurent Njilla, Charles Kamhoua, Sachin Shetty, DaeHun Nyang and David Mohaisen, "Exploring the Attack Surface of Blockchain: A Comprehensive Survey," IEEE Communications Surveys & Tutorials, 2020.
- 10) Mohammad Wazid, Ashok Kumar Das, Joel J. P. C. Rodrigues, Sachin Shetty and YoungHo Park, "IoMT Malware Detection Approaches: Analysis and Research Challenges," IEEE Access, Vol: 7, December 17, 2019.
- 11) Mohammad Wazid, Ashok Kumar Das, Sachin Shetty, Joel J. P. C. Rodrigues and YoungHo Park, "LDAKM-ElIoT: Lightweight Device Authentication and Key Management Mechanism for Edge-Based IoT," Sensors, 19(24), December 14, 2019.

- 12) Kamrul Hasan, Sachin Shetty, Sharif Ullah, Amin Hassanzadeh and Ethan Hadar, "Towards Optimal Cyber Defense Remediation in Energy Delivery Systems," IEEE Globecom, Hawaii, December 9-13, 2019.
- 13) Marco Gamarra, Sachin Shetty, Oscar Gonzalez, David Nicol and Laurent Njilla, "Modeling Stepping Stone Attacks with Constraints in Cyber Infrastructure," IEEE Globecom, Hawaii, December 9-13, 2019.
- 14) J. Padilla, D. Shuttleworth and K. O'Brien, "Agent-based model characterization using Natural Language Processing," Winter Simulation Conference, National Harbor, MD, December 8-11, 2019.
- 15) Neil Balram, George Hsieh and Christian McFall, "Static Malware Analysis using Machine Learning Algorithms on APT1 Dataset with String and PE Header Features," 6th Annual Conf. on Computational Science & Computational Intelligence (CSCI'19), Las Vegas, NV, December 5-7, 2019.
- 16) Jules Chenou, George Hsieh and Tonya L. Fields, "Radial Basis Function Network: Its Robustness and Ability to Mitigate Adversarial Examples," 6th Annual Conf. on Computational Science & Computational Intelligence (CSCI'19), Las Vegas, NV, December 5-7, 2019.
- 17) Tonya Fields, George Hsieh and Jules Chenou, "Mitigating Drift in Time Series Data with Noise Augmentation," 6th Annual Conf. on Computational Science & Computational Intelligence (CSCI'19), Las Vegas, NV, December 5-7, 2019.
- 18) Jonathan Graham, Kwesi Elliot, Yusef Yassin, Trenton Ward, Johan Caldwell and Tawab Attie, "A Comparison of Machine Learning Algorithms in Keystroke Dynamics," 6th Annual Conf. on Computational Science & Computational Intelligence (CSCI'19), Las Vegas, NV, December 5-7, 2019.
- 19) Ariful Haque, Sachin Shetty and Bheshaj Krishnappa, "CPS resilience: Frameworks, metrics, complexities, challenges and future directions," Complexity Challenges in Cyber Physical Systems: Using Modeling and Simulation (M&S) to Support Intelligence, Adaptation and Autonomy, Eds.: Saurabh Mittal and Andreas Tolk, Wiley, November 2019.

- 20) Ariful Haque, Sachin Shetty and Bheshaj Krishnappa, "Modeling Cyber Resilience for Energy Delivery Systems Using Critical System Functionality," Resilience Week, San Antonio, TX, November 4-7, 2019.
- 21) Mohammad Wazid, Ashok Kumar Das, Sachin Shetty, Joel J. P. C. Rodrigues and YoungHo Park, "AKM-IoV: Authenticated Key Management Protocol in Fog Computing-Based Internet of Vehicles Deployment," IEEE Internet of Things Journal, 6(5), October 2019.
- 22) Md Ali Reza Al Amin, Sachin Shetty, Laurent Njilla, Deepak Tosh and Charles Kamhoua, "Online Cyber Deception System using Partially Observable Monte-Carlo Planning Framework," SecureComm 2019 Conf., Orlando, FL, October 23-25, 2019.
- 23) Sharif Ullah, Sachin Shetty, Anup Nayak, Amin Hassanzadeh and Kamrul Hasan "Cyber Threat Analysis based on Characterizing Adversarial Behavior for Energy Deliver System," SecureComm 2019 Conf. (short paper), Orlando, FL, October 23-25, 2019.
- 24) Scott M. Debb, A.T. Shappie, C.A. Dawson and D.T. Shaffer, "Attitudes and personality factors associated with cybersecurity and social media for college students," 126th Annual Convention of the American Psychological Association, Chicago, IL, August 8-11, 2019.
- 25) Jules Chenou, George Hsieh and Tonya L. Fields, "Increasing the Robustness of Deep Learning with Colored Noise Augmentation," Security and Management 2019 Conf., Las Vegas, NV, July 29 - August 1, 2019.
- 26) Deepak Tosh, Sachin Shetty, Xueping Liang, Charles Kamhoua and Laurent Njilla, "Data Provenance in the Cloud: A Blockchain-Based Approach," IEEE Consumer Electronics Magazine, vol. 8, no. 4, pp. 38-44, July 2019.
- 27) Sarada Prasad Gochhayat, Eranga Herath, Sachin Shetty and Peter Foytik, "Yugala: Blockchain based Encrypted Cloud Storage for IoT Data," IEEE Blockchain 2019 Conf., Atlanta, GA, July 14-17, 2019.
- 28) Charlotte Dawson, Alex Shappie and Scott Debb, "Personality and cybersecurity behaviors: The role of the 'Big Five'," CYPSY24 2019 Conf., Norfolk State University, Norfolk, VA, June 24-26, 2019.

- 29) Daniel Schaffer and Scott Debb, "Confirmatory factor analysis to analyze Instagram Questionnaire factor structure fit across cultural settings," CYPSY24 2019 Conf., Norfolk State University, Norfolk, VA, June 24-26, 2019.
- 30) Mary Ann Hoppa, George Hsieh, Bigyan Kc and Scott Debb, "Twitter Open Source Intelligence (OSINT): automated open source intelligence collection, analysis and visualization tool," CYPSY24 2019 Conf., Norfolk State University, Norfolk, VA, June 24-26, 2019.
- 1) Sachin Shetty, Val Red, David Satterfield, Charles Kamhoua, Kevin Kwiat and Laurent Njilla, "Data Provenance Assurance in Cloud Using Blockchain," SPIE Defense + Security Conference, Anaheim, CA, April 9-13, 2017.
- 2) Ali, A., Hsieh, G. and Khan, M., "Security of Anonymous Networks," Society of Industrial and Applied Mathematics and Old Dominion University Mathematics Awareness Conference, Norfolk, VA, April 15, 2017.
- 3) Kavak, H., Padilla, J., Vernon-Bido, D., Diallo, S., and Gore, G. "The Spread of Wi-Fi Router Malware Revisited," Communications and Network Symposium at the Spring Simulation Conference, Virginia Beach, VA, April 23-26, 2017.
- 4) Ali, A., Hsieh, G. and Khan, M., "Application of Cycle-Free Graphs for Data Security in Large Distributed Systems," Mathematical Association of America, Maryland-District of Columbia-Virginia Section, Frostburg, MD, April 28-29, 2017.
- 5) Yen-Hung Hu and Shawnee L. Parker, "Big Data Content Mining Techniques for Detecting Cyberbullying in Social Media," 2017 Virginia Academy of Science Annual Meeting, May 17-19, 2017, Richmond, VA.
- 6) Yen-Hung Hu and Thomas Kofi Annan, "Review of secured mobile application development process in the industry," 2017 Virginia Academy of Science Annual Meeting, May 17-19, 2017, Richmond, VA.
- 7) Deepak Tosh, Sachin Shetty, Shamik Sengupta, Jay Keshan and Charles Kamhoua, "Risk Management using Cyber-Threat Information Sharing and Cyber-Insurance," EAI International Conference on Game Theory for Networks, May 2017, Knoxville, TN.
- 8) Xueping Liang, Sachin Shetty, Deepak Tosh, Charles Kamhoua, Kevin Kwiat and Laurent Njilla, "ProvChain: A Blockchain-based Data Provenance Architecture in Cloud Environ-

ment with Enhanced Privacy and Availability," IEEE CCGRiD, pp. 468-477, May 2017, Madrid, Spain.

9) Deepak Tosh, Sachin Shetty, Xueping Liang, Charles Kamhoua, Kevin Kwiat and Laurent Njilla, "Security Implications of Blockchain Cloud with Analysis of Block Withholding Attack," IEEE CCGRiD, pp. 458-467, May 2017, Madrid, Spain.

10) Uttam Ghosh, Pushpita Chatterjee and Sachin Shetty, "A Security Framework for SDN Enabled Power Grid," IEEE International workshop on Communication, Computing, and Networking in Cyber Physical Systems, In association with IEEE the 37th International Conference on Distributed Computing Systems (IEEE ICDCS 2017), June 5 - 8, 2017, Atlanta, GA.

11) Xueping Liang, Sachin Shetty, Lingchen Zhang, Charles Kamhoua and Kevin Kwiat, "Man in the Cloud (MITC) Defender: SGX-based User Credential Protection for Synchronization applications in Cloud Computing Platform," IEEE Cloud, June 25- 30, 2017, Honolulu, Hawaii.

12) Uttam Ghosh, Pushpita Chatterjee, Deepak Tosh, Sachin Shetty, Kaiqi Xiong and Charles Kamhoua, "A SDN based Framework for Guaranteeing Security and Performance in Information Centric Cloud Networks," IEEE Cloud (short paper), June 25-30, 2017, Honolulu, Hawaii.

13) Debb, S.M. and Parker, T., "Examining the culture of information security awareness for post-millennial African-American undergraduate students," 22nd CyberPsychology, CyberTherapy and Social Networking Conference, Wolverhampton, U.K, June 2017. (Poster)

14) Boinapally, V., Hsieh, G. and Nauer, K., "Building a Gh0st Malware Experimentation Environment," 2017 International Conference on Security and Management, Las Vegas, NV, July 17-20, 2017.

15) Perry, K., Hsieh, G., Galvin, S. and Nauer, K., "Building a Virtual Enterprise Network Environment for APT Experimentation," 2017 International Conference on Security and Management, Las Vegas, NV, July 17-20, 2017.

16) Hinds, Cheryl and Sherfield, Joshua, "Thematic Graphical Authorization: Graphical Authentication using Thematic Images on Mobile Devices," 2017 International Conference on Security and Management, Las Vegas, NV, July 17-20, 2017.

- 17) Xueping Liang, Juan Zhao, Sachin Shetty, Jihong Liu, and Danyi Li, "Integrating Blockchain for Data Sharing and Collaboration in Mobile Healthcare Applications," IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, October 8-13, 2017, Montreal, QC, Canada.
- 18) Deepak Tosh, Sachin Shetty, Xueping Liang, Charles A. Kamhoua, and Laurent Njilla, "Consensus Protocols for Blockchain-based Data Provenance: Challenges and Opportunities," 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON) (Short Paper), NY, October 19-21, 2017.
- 19) Deepak Tosh, Iman Vakulinia, Sachin Shetty, Shamik Sengupta, Charles A. Kamhoua, Laurent Njilla, and Kevin Kwiat, "Three Layer Game Theoretic Decision Framework for Cyber-Investment and Cyber-Insurance," GameSec 2017, 8th Conference on Decision and Game Theory for Security, Vienna, Austria, October 23-25, 2017.
- 20) Xueping Liang, Juan Zhao, Sachin Shetty, and Danyi Li, "Towards Data Assurance and Resilience in IoT Using Distributed Ledger," IEEE Milcom, October 23-25, 2017, Baltimore, MD.
- 21) Juan Zhao, Sachin Shetty, and Jan Wei Pan, "Feature-Based Transfer Learning for Network Security," IEEE Milcom, October 23-25, 2017, Baltimore, MD.
- 22) Laila Almutairi and Sachin Shetty, "Generalized Stochastic Petri Net Model Based Security Risk Assessment of Software Defined Networks," IEEE Milcom, October 23-25, 2017, Baltimore, MD.
- 23) Iman Vakulinia, Deepak Tosh, and Shamik Sengupta, "3-way Game Model for Privacy-Preserving Cybersecurity Information Exchange Framework," IEEE Milcom, October 23-25, 2017, Baltimore, MD.
- 24) Ali, A., Hsieh, G. and Khan, M., "Malware Analysis with Finite Ordered Trees," Mathematical Association of America, Maryland-District of Columbia-Virginia Section, Newport News, VA, November 17-18, 2017.
- 25) Mamta Kumari, George Hsieh, and Christopher A. Okonkwo, "Deep Learning Approach to Malware Multi-Class Classification Using Image Processing Techniques," 2017 International

Conference on Computational Science and Computational Intelligence (CSCI'17), Las Vegas, NV, December 14-16, 2017.

26) Cheryl Hinds and Peter Fenton, "ESeizure: A Mobile App for Digital Forensics First Responders," 2017 International Conference on Computational Science and Computational Intelligence (CSCI'17), Las Vegas, NV, December 14-16, 2017.

27) Jonathan Graham, Cheryl Hinds, and Sandi Samuel, "Hunting Malware: An Example using Gh0st," 2017 International Conference on Computational Science and Computational Intelligence (CSCI'17), Las Vegas, NV, December 14-16, 2017.

28) Xueping Liang, Sachin Shetty, Juan Zhao, Daniel Bowden, Danyi Li and Jihong Liu, "Towards Decentralized Accountability and Self-Sovereignty in Healthcare Systems," The 19th International Conference on Information and Communications Security (ICICS 2017), Beijing, China, Dec 2017.

29) Xueping Liang, Sachin Shetty, Deepak Tosh, Peter Foytik and Lingchen Zhang, "Towards a Trusted and Privacy Preserving Membership Service in Distributed Ledger using Intel Software Guard Extensions," The 19th International Conference on Information and Communications Security (ICICS 2017), Beijing, China, Dec 2017.

30) Ali, Abdinur, Hu, Yen-Hung, Hsieh, Chung-Chu (George), and Khan, Mushtaq (2017), "A Comparative Study on Machine Learning Algorithms for Network Defense," Virginia Journal of Science: Vol. 68: No. 3, Article 1. Available at:
<https://digitalcommons.odu.edu/vjs/vol68/iss3/1>

31) Hellen Maziku, Sachin Shetty, Dong Jin, Charles A. Kamhoua, Laurent Njilla, Kevin Kwiat, "Diversity Modeling to Evaluate Security of Multiple SDN Controllers," 2018 International Conference on Computing, Networking and Communications (ICNC): Communications and Information Security Symposium - Communications and Information Security, Maui, Hawaii, March 5-8, 2018.

32) Ali, A., Hsieh, G. and Khan, M., "Machine Learning Approaches for Malware Detection," Mathematical Association of America, Maryland-District of Columbia-Virginia Section, Newport News, VA, April 13-14, 2018.

- 33) Vernon-Bido, D., Grigoryan, G. Kavak, H., and Padilla, J., "Assessing the Impact of Cyberloafing on Cyber Risk," SpringSim Conference, Baltimore, MD, April 15-18, 2018.
- 34) Kavak, H., Padilla, J., Diallo, S., and Lynch, C., "Big Data, Agents, and Machine Learning: Towards a Data-Driven Agent-based Modeling Approach," 2018 SpringSim Conference, Baltimore, MD, April 15-18, 2018.
- 35) Sachin Shetty, Michael McShane, Jay Kesan, Linfeng Zhang, Charles Kamhoua and Laurent Njilla, "Reducing Informational Disadvantages to Improve Cyber Risk Management," The Geneva Papers on Risk and Insurance. Springer, 43 (2), April 2018, pp 224-238.
- 36) Ali, A., Hsieh, G. and Khan, M., "Web Mining based on Machine Learning Techniques," Society of Industrial & Applied Mathematics-Old Dominion University, Norfolk, VA, April 21, 2018.
- 37) Marco Gammara, Sachin Shetty, Oscar Gonzalez, David M. Nicol, Charles Kamhoua, Laurent Njilla, "Analysis of Stepping Stone Attacks in Dynamic Vulnerability Graphs," accepted for IEEE ICC 2018, May 20-24, 2018, Kansas City, MO.
- 38) Kamrul Hasan, Sachin Shetty, Amin Hassanzadeh, Malek Ben Salem and Jay Chen "Modeling Cost of Countermeasures in Software Defined Networking-enabled Energy Delivery System," accepted for IEEE Conference on Communications and Network Security (CNS), 30 May-1 June 2018, Beijing China.
- 39) Leah Winkfield, Yen-Hung Hu, and Mary Ann Hoppa, "A Study of the Evolution of Secure Software Development Architecture," accepted for CISSE 2018 Conference, June 11-13, 2018, New Orleans, LA.
- 40) Debb, S.M., "Information security awareness and African American college students: An exploration of attitudes and behaviors," accepted for the 23rd Cyberpsychology, CyberTherapy and Social Networking Conference, June 26-28, 2018, Montreal, Canada.
- 41) Bivens, T., and Debb, S.M., "Exploring the negative impacts of #Fitspiration using GSR and eye tracking technology," (poster) accepted for the 23rd Cyberpsychology, CyberTherapy and Social Networking Conference, June 26-28, 2018, Montreal, Canada.

- 42) Debb, S.M., and Panigrahi, E., "Validation and cross-cultural comparison of the Online Self-Presentation Scale for use in the United States," accepted for the 23rd Cyberpsychology, CyberTherapy and Social Networking Conference, June 26-28, 2018, Montreal, Canada.
- 43) Deepak Tosh, Sachin Shetty, Peter Foytik, Charles Kamhoua, and Laurent Njilla, "CloudPoS: A Proof-of-Stake Consensus Design for Blockchain Integrated Cloud," submitted to the IEEE Cloud Conference 2018, July 2-7, 2018, San Francisco, CA.
- 44) Kavak, H., Vernon-Bido, D., and Padilla, J., "Fine-scale Prediction of People's Home Location using Social Media Footprints," accepted for the International Conference on Social Computing, Behavior-Cultural Modeling & Prediction and Behavior Representations in Modeling and Simulation, July 10-13, 2018, Washington, DC.
- 45) Brandon Johnson, George Hsieh, and Mary Ann Hoppa, "Automating Deployment of Gh0st Remote Access Trojan Malware," accepted for the Security and Management 2018 Conference, July 30 - August 2, 2018, Las Vegas, NV.
- 46) Cody Butler, George Thompson, George Hsieh, Mary Ann Hoppa, and Kevin S. Nauer, "Developing a UI and Automation Framework for a Cybersecurity Research and Experimentation Environment," accepted for the Security and Management 2018 Conference, July 30 - August 2, 2018, Las Vegas, NV.
- 47) Abdinur Ali, Yen-Hung Hu, Cheryl Hinds, and Jonathan Graham, "Diffusion Metrics of the AES Symmetric Cryptosystem," accepted for the FCS'18 Conference, July 30 - August 2, 2018, Las Vegas, NV.
- 48) Ali, Abdinur; Hu, Yen-Hung; Hsieh, Chung-Chu (George); and Khan, Mushtaq (2018) "Detailed Accuracy Analysis by Attack Class Based on Four Machine Learning Algorithms for Network Security," accepted for the ICAI'18 - The 20th International Conference on Artificial Intelligence, July 30 - August 2, 2018, Las Vegas, NV.
- 49) Padilla, J., Kavak, H., Vernon-Bido, D., Diallo, S., Shetty, S., and Gore, R. "Simulation for Cyber," accepted by the Journal of Defense Modeling and Simulation.
- 1) Xuebiao Yuchi and Sachin Shetty "Enabling Security-Aware Virtual Machine Placement in IaaS Clouds," 2015 MILCOM Conference, October 2015, Tampa FL.

- 2) Turner, C. and Joseph, A., "A Wavelet and Mel-Frequency Cepstral Coefficients-Based Feature Extraction Method for Speaker Identification," *Procedia Computer Science*, Vol. 61, November 2015, pp. 416-421.
- 3) Ali, A., Hsieh, G. and Khan, M., "Performance Metrics for Machine Learning Algorithms," *Mathematical Association of America Conference*, Saint Mary's City, Maryland, November 6-7, 2015.
- 4) Turner, C., Turner, C., Richards, D. and Yan, J., "PingFloodViz: A Visualization Tool Targeted to a Non-Technical Cyber Defense Competition Audience," abstract submitted to *Complex Adaptive Systems Conference*, Nov 2-4, 2016, Los Angeles, CA.
- 5) Shetty, S., Hsieh, G., Padilla, J.J., Graham, J.M., Kamhoua, C.A. and Kwiat, K.A., "Cyber Intelligence and Cyber Defense," accepted (session presentation) by *CISSE 2016*, June 2016, Philadelphia, PA.
- 6) Padilla, J.J., Hsieh, G., Shetty, S., Graham, J.M., Kamhoua, C.A. and Kwiat, K.A., "Cyber Modeling, Simulation, Analysis and Decision Support," accepted (session presentation) by *CISSE 2016*, June 2016, Philadelphia, PA.
- 7) Hsieh, G., Graham, J.M., Padilla, J.J., Shetty, S., Kamhoua, C.A. and Kwiat, K.A., "Cyber Research Experimentation and Training Infrastructures," accepted (session presentation) by *CISSE 2016*, June 2016, Philadelphia, PA.
- 8) Bonas, R. and Graham, J., "Malware Classification: An integrated view," submitted in November 2015 to *CYBERSEC-16 Cybersecurity Symposium: Your Security, Your Future*, April 2016, Coeur d'Alene, Idaho.
- 9) Fields, T. and Graham, J., "Classifying Network Attack Data Using Random Forest," accepted by *CATA 2016*, April 2016, Las Vegas, NV.
- 10) Hinds, C., "Securing Smart Meters in the Smart Grid," accepted by *CYBERSEC-16 Cybersecurity Symposium: Your Security, Your Future*, April 2016, Coeur d'Alene, Idaho.
- 11) Hsieh, G., Kotut, L. and Ali, A., "Big Data Analytic Techniques for Cybersecurity Applications," accepted (poster) for *CYBERSEC-16 Cybersecurity Symposium*, April 2016, Coeur d'Alene, Idaho.

- 12) Kavak, H., Padilla, J., Vernon-Bido, D., Gore, R., and Diallo, S., "A characterization of cybersecurity simulation scenarios," accepted for SpringSim Conference 2016, April 2016, Pasadena, CA.
- 13) Vernon-Bido, D., Padilla, J., Kavak, H., Diallo, S. and Gore, R., "Towards modeling factors that enable an attacker," accepted for publication at SummerSim Conference 2016, July 2016, Montreal, Canada.
- 14) Vernon-Bido, D., Padilla, J., Kavak, H., Diallo, S. and Gore, R., "Modeling insider threat in the cyber security domain," under preparation.
- 15) Gore, R., Padilla, J. and Diallo, S., "Markov chain monte carlo modeling of cyber threats," under review, Journal of Defense Modeling and Simulation.
- 16) Gore, R., Diallo, S., Padilla, J. and Ezell, B. "Assessing Cyber-Incidents Using Machine Learning," under review, Journal of Defense Modeling and Simulation
- 17) Padilla, J., Kavak, H., Vernon-Bido, D., Diallo, S., and Gore, R., "An introduction to cyber simulation: dimensions, missions and applications," under preparation.
- 18) Paul Mcneil, Sachin Shetty, Divya Guntu and Gauree Barve, "SCREDDENT: Scalable Real-time Anomalies Detection and Notification of Targeted Malware in Mobile Devices," 7th International Conference on Ambient Systems, Networks and Technologies (ANT-2016)- International Workshop on Mobile Cloud Computing systems, Management, and Security, May 2016 (accepted).
- 19) Sai Kiran Mukkavilli, Sachin Shetty and Liang Hong, "Generation of Labeled Datasets to Quantify the Impact of Security Threats to Cloud Data Centers," Journal of Information Security, 2016 (accepted).
- 20) Taiwo Oyedare, Ashraf Al Sharah and Sachin Shetty, "A Reputation-Based Coalition Game to Prevent Smart Insider Jamming Attacks in MANETs," 14th International Conference on Wired & Wireless Internet Communications (WWIC 2016), May 2016 (accepted).
- 21) Sachin Shetty, Xuebiao Yuchi and Min Song, Moving Target Defense for Distributed Systems, Springer, ISBN 978-3-319-31031-2, April 2016.

- 22) Ashraf Al Sharah, Taiwo Oyedare and Sachin Shetty, "Detecting and Mitigating Smart Insider Jamming Attacks in MANETs Using Reputation Based and Coalition Game," Journal of Computer Networks and Communications, Hindawi Publishing Corporation [Accepted].
- 23) Ashraf Al Sharah, Taiwo Oyedare and Sachin Shetty, "Colluding Jamming Attack on a Grand Coalition by Aggrieved Nodes, Journal of Communications and Network," Scientific Research Publishing [Accepted].
- 24) Xuebiao Yuchi, Sachin Shetty, Charles Kamhoua, Kevin Kwiat, "Evaluating Cloud Network's Robustness by Scalable Network Diversity Modeling," under PA review.
- 25) Xuebiao Yuchi, Sachin Shetty, Charles Kamhoua, Kevin Kwiat, "Network Aware Resource Allocation in Cloud Data Center," under PA review.
- 26) Hellen Maziku, Sachin Shetty, David Nicol Malcolm, "Security Risk Assessment for SDN-enabled Smart Grids," submitted to Security and Communication Networks Journal, Wiley Publishing.
- 27) Kamrul Hasan, Taiwo Oyedare, Sachin Shetty, "Cross Layer Attacks on GSM Mobile Networks Using Software Defined Radios," submitted to Vehicular Technology Conference.
- 28) Ali, A., Hsieh, G. and Khan, M., "Statistical Learning Methods for Cyber Security," accepted by Society of Industrial and Applied Mathematics and Old Dominion University Mathematics Awareness Conference, Norfolk, Virginia, April 9, 2016.
- 29) Ali, A., Hsieh, G. and Khan, M., "Hardware Security Vulnerabilities and Cryptanalysis of Modern Data Encryption Algorithms," accepted for Mathematical Association of America, Maryland-District of Columbia-Virginia Section, Germantown, MD, April 15 -16, 2016.
- 30) Turner, C., Jeremiah, R., Joseph, A., "A Study of Snort Rule Performance, Categorization and Statistics," abstract submitted to Complex Adaptive Systems Conference, Nov 2-4, 2016, Los Angeles, CA.
- 31) Jeremiah, R., Turner, C., Richards, D., Joseph, A. "Intrusion Detection and Visualization of Password Compromise Through Linux Security Auditd Daemon," abstract submitted to Complex Adaptive Systems Conference, Nov 2-4, 2016, Los Angeles, CA.

32) Esther Amullen, Sachin Shetty and L.H. Keel, “Secured Formation Control for Multi-agent Systems under DoS Attacks”, Homeland Security Science and Technology Conference, 2016.

PRESENTATIONS AND TALKS

- 1) Center of Excellence in Cyber Security, HBCU Centers of Excellence Outreach to ARFL, February 2, 2021 (virtual).
- 2) Center of Excellence in Cyber Security at Norfolk State University, meeting with PACT at University of Arizona, March 18, 2021 (virtual).
- 3) Centers of Excellence Teams Meeting on COVID-19 impact with Ms. Evelyn Kent, March 23, 2021 (virtual).
- 4) Center of Excellence in Cyber Security at Norfolk State University, meeting with MITRE Bridging Innovation Program, March 29, 2021 (virtual).
- 5) Dr. Sachin Shetty gave a talk at Argonne National Lab’s cybersecurity seminar series on October 2, 2020.
- 6) Dr. Sachin Shetty gave a talk at Pacific National Lab’s cybersecurity seminar series on November 17, 2020.
- 7) CCI Virtual Summer Academy on AI Assurance, August 10 - 14, 2020. Drs. Sachin Shetty and Yen-Hung (Frank) Hu of the COE each presented a talk.
- 8) Sandia Tracer FIRE Virtual Training Workshop, August 12 - 13, 2020. Dr. George Hsieh participated in the workshop to learn Tracer FIRE and discuss collaboration plans.
- 9) The COE held its Annual Program Review and Advisory Board Meeting on May 28, 2020 (virtual).
- 1) Dr. Laurent Njilla, AFRL Program Manager for the COE, visited the COE personnel at Norfolk State University and Old Dominion University on March 10-11, 2020 to review and discuss key project matters as we wrapped up 1Q2020.

- 2) At the meeting with ARL/ARO visitors on February 28, 2020 at NSU, eight briefings and demonstrations were presented by the COE faculty and student researcher team at NSU covering a variety of major research topics in CASE-V platform, emulated attackers, cyber threat intelligence, machine learning for malware classification, security and robustness of machine learning, forensic and incident response training platform, virtual lab environment for Hyperledger based blockchain technology, human performance and behavior in the digital age, and fair and transparent decision making.
- 3) At the meeting with NSU President Dr. Javaune Adams-Gaston on February 11, 2020, Dr. Scott Debb presented a briefing titled "Human Performance and Behavior in the Digital Age" and Dr. George Hsieh presented an update on the COE project.
- 4) Dr. Jules Chenou presented the paper titled "Radial Basis Function Network: Its Robustness and Ability to Mitigate Adversarial Examples," at CSCI'19 Conf., Las Vegas, NV, December 5-7, 2019.
- 5) Tonya Fields presented the paper titled "Mitigating Drift in Time Series Data with Noise Augmentation," at CSCI'19 Conf., Las Vegas, NV, December 5-7, 2019.
- 6) Neil Balram, a graduate RA for the COE at NSU, presented the paper titled "Static Malware Analysis using Machine Learning Algorithms on APT1 Dataset with String and PE Header Features," at CSCI'19 Conf., Las Vegas, NV, December 5-7, 2019.
- 7) Kwesi Elliot, a graduate RA (spring 2019) for the COE at NSU presented the paper titled "A Comparison of Machine Learning Algorithms in Keystroke Dynamics," at CSCI'19 Conf., Las Vegas, NV, December 5-7, 2019.
- 8) Dr. Lijun Qian, PI/PD for the CREDIT Center at PVAMU visited NSU on November 15, 2019 to discuss research collaboration with the COE in Cyber Security.
- 9) Dr. Laurent Njilla, AFRL Program Manager for the COE, visited the COE personnel at Norfolk State University and Old Dominion University on September 27, 2019, to review and discuss key project matters as we wrapped up 3Q2019.
- 10) Dr. George Hsieh gave a presentation titled "Data Science @ Center of Excellence in Cyber Security" at the meeting with the ACC A29A Data Analytics team on September 24, 2019 at NSU.

- 11) Dr. George Hsieh hosted a tour and discussion of the NSU Cybersecurity Complex (including the COE facility and research infrastructure) for an ODU team led by Dr. Michael Wu, Director of Center for Cybersecurity Education and Research, on September 12, 2019. The ODU team was very impressed by what they saw, and plan to adopt some of the design and implementation ideas for its new cybersecurity research and education facility being developed.
- 12) Dr. George Hsieh hosted a tour of the COE facility and research infrastructure for the Department of Navy HBCU/MI Program personnel led by Mr. Anthony Smith Jr. on September 4, 2019.
- 13) Dr. George Hsieh gave a presentation titled "Center of Excellence in Cyber Security @ NSU" to Dr. Javaune Adams-Gaston, President of NSU, on August 8, 2019.
- 14) Dr. Jules Chenou and Tonya Fields presented the paper titled "Increasing the Robustness of Deep Learning with Colored Noise Augmentation," at the Security and Management 2019 Conf., Las Vegas, NV, July 29 - August 1, 2019.
- 15) Charlotte Dawson and Alex Shappie presented the paper titled "Personality and cybersecurity behaviors: The role of the 'Big Five'" at CYPSY24 2019 Conf., Norfolk State University, Norfolk, VA, June 24-26, 2019.
- 16) Daniel Schaffer presented the paper titled "Confirmatory factor analysis to analyze Instagram Questionnaire factor structure fit across cultural settings" at CYPSY24 2019 Conf., Norfolk State University, Norfolk, VA, June 24-26, 2019.
- 17) Drs. Mary Ann Hoppa and George Hsieh presented the paper titled "Twitter Open Source Intelligence (OSINT): automated open source intelligence collection, analysis and visualization tool" at CYPSY24 2019 Conf., Norfolk State University, Norfolk, VA, June 24-26, 2019.
- 18) The 24th CyberPsychology, CyberTherapy and Social Networking (CyPsy24) Conference was held on June 24-26, 2019 at NSU. The COE hosted a facility tour to discuss potential collaboration opportunities with key participants.
- 19) Drs. George Hsieh and Mary Ann Hoppa attended the OUSD HBCU/MI Workshop in Durham NC on June 12-13, 2019. They presented a poster about the COE.
- 20) The COE held its Fourth Annual Program Review and Advisory Board Meeting on June 6, 2019, at NSU.

- 21) Dr. Sachin Shetty was invited to give a talk titled “Cyber Risk Scoring and Mitigation,” Turbo Transition, US Transcom, June 2019.
 - 22) Kwesi Elliot presented the paper titled “The use of Keystroke Dynamics as a Primary means of Authentication” at World Congress of Undergraduate Research, Carl von Ossietzky University, Oldenburg, Germany, May 22-25, 2019.
 - 23) Drs. Mary Ann Hoppa and Yen-Hung Hu attended the Inaugural Annual HBCU Blockchain Summit on May 3-4, 2019 at Morgan State University in collaboration with Ripple, a Silicon Valley technology firm. This summit was the largest gathering of HBCU faculty from around the country to explore and discuss blockchain technology, cryptocurrencies, and use cases.
 - 24) Dr. Yen-Hung (Frank) Hu presented the paper titled “Machine Learning Techniques for Classifying Malicious API Calls and N-Grams in Kaggle Data-set” at IEEE SoutheastCon 2019, Huntsville, AL, April 11-14, 2019.
- 1) Mr. Hamdi Kavak (ODU, GRA) presented the paper titled “The Spread of Wi-Fi Router Malware Revisited” at the Communications and Network Symposium at Spring Simulation Conference, Virginia Beach, VA, April 2017.
 - 2) Dr. Abdinur Ali presented the paper titled “Security of Anonymous Networks” at the Society of Industrial and Applied Mathematics and Old Dominion University Mathematics Awareness Conference, Norfolk, VA, April 15, 2017.
 - 3) Dr. George Hsieh presented a talk titled “Center of Excellence in Cyber Security” at ODU’s School of Business on April 19, 2017.
 - 4) Dr. Abdinur Ali presented the paper titled “Application of Cycle-Free Graphs for Data Security in Large Distributed Systems” at the Meeting for Mathematical Association of America, Maryland-District of Columbia-Virginia Section, Frostburg, MD, April 28-29, 2017.
 - 5) Dr. Yen-Hung Hu presented the paper titled “Big Data Content Mining Techniques for Detecting Cyberbullying in Social Media” at the 2017 Virginia Academy of Science Annual Meeting, May 17-19, 2017, Richmond, VA.

- 6) Dr. Yen-Hung Hu presented the paper titled "Review of secured mobile application development process in the industry" at the 2017 Virginia Academy of Science Annual Meeting, May 17-19, 2017, Richmond, VA.
- 7) Dr. Deepak Tosh presented the paper titled "Risk Management using Cyber-Threat Information Sharing and Cyber-Insurance" at the EAI International Conference on Game Theory for Networks, May 2017, Knoxville, TN.
- 8) Dr. Sachin Shetty presented the paper titled "ProvChain: A Blockchain-based Data Provenance Architecture in Cloud Environment with Enhanced Privacy and Availability" at IEEE CCGRiD, May 2017, Madrid, Spain.
- 9) Dr. Sachin Shetty presented the paper titled "Security Implications of Blockchain Cloud with Analysis of Block Withholding Attack" at IEEE CCGRiD, May 2017, Madrid, Spain.
- 10) Dr. Sachin Shetty gave a talk during the Blockchain Symposium at AFRL/Rome on June 1, 2017.
- 11) Dr. Sachin Shetty presented the paper "A SDN based Framework for Guaranteeing Security and Performance in Information Centric Cloud Networks" at IEEE Cloud conference, June 2017, Honolulu, Hawaii.
- 12) Dr. Sachin Shetty presented the paper "Man in the Cloud (MITC) Defender: SGX-based User Credential Protection for Synchronization applications in Cloud Computing Platform" at IEEE Cloud conference, June 2017, Honolulu, Hawaii.
- 13) Dr. Scott Debb presented the poster titled "Examining the culture of information security awareness for post-millennial African-American undergraduate students" at the 22nd CyberPsychology, CyberTherapy and Social Networking Conference, Wolverhampton, U.K, June 2017.
- 14) Dr. Sachin Shetty gave a talk on "Cyber Risk Scoring and Mitigation" at Scott Air Force Base, Shiloh, IL.
- 15) Dr. Sachin Shetty gave a talk on "Security Metrics for Cyber Insurance" at ECE Department, ODU.

- 16) Dr. George Hsieh presented a talk titled “Cyber Analysis, Simulation and Experimentation for Cyber Training” at 2017 Workshop on Employees' Cybersecurity Behavior, July 14, 2017, ODU.
- 17) Dr. George Hsieh presented the paper titled “Building a Gh0st Malware Experimentation Environment” at 2017 International Conference on Security and Management, Las Vegas, NV, July 17-20, 2017.
- 18) Dr. George Hsieh presented the paper titled “Building a Virtual Enterprise Network Environment for APT Experimentation” at 2017 International Conference on Security and Management, Las Vegas, NV, July 17-20, 2017.
- 19) Dr. Cheryl Hinds presented the paper titled “Thematic Graphical Authorization: Graphical Authentication using Thematic Images on Mobile Devices” at 2017 International Conference on Security and Management, Las Vegas, NV, July 17-20, 2017.
- 20) Dr. George Hsieh presented a tech talk titled “Cyber Analysis, Simulation and Experimentation” at Sandia National Labs on July 27, 2017, Albuquerque, NM.
- 21) Dr. George Hsieh attended the Naval Opportunity Awareness Workshop, September 21-22, 2017, Florida A&M University, Tallahassee, FL.
- 22) Dr. Deepak Tosh presented the paper “Three Layer Game Theoretic Decision Framework for Cyber-Investment and Cyber-Insurance” at GameSec conference, October 2017, Vienna, Austria.
- 23) Dr. Deepak Tosh presented the paper “Consensus Protocols for Blockchain-based Data Provenance: Challenges and Opportunities” at IEEE UEMCON conference, October 2017, New York, NY.
- 24) Dr. George Hsieh presented a talk titled “Big Data Analytics for Cyber Analysis, Simulation and Experimentation” at the Department of Mathematics and Statistics, ODU, October 24, 2017.
- 25) Dr. George Hsieh presented a talk titled “SCHEV Assessment Briefing - Cybersecurity” at an assessment meeting with State Council of Higher Education for Virginia, October 26, 2017, NSU.

- 26) Dr. Abdinur Ali presented a talk titled “Malware Analysis with Finite Ordered Trees” at the Mathematical Association of America, Maryland-District of Columbia-Virginia Section, Newport News, VA, November 17-18, 2017.
- 27) Mamta Kumari, a graduate student in the MS.CSC program at NSU, presented the paper titled “Deep Learning Approach to Malware Multi-Class Classification Using Image Processing Techniques” at 2017 International Conference on Computational Science and Computational Intelligence (CSCI’17), Las Vegas, NV, December 14-16, 2017.
- 28) Peter Fenton, a graduate student in the MS in Cybersecurity program at NSU, presented the paper titled “ESeizure: A Mobile App for Digital Forensics First Responders” at 2017 International Conference on Computational Science and Computational Intelligence (CSCI’17), Las Vegas, NV, December 14-16, 2017.
- 29) Sandi Samuel, a graduate student in the MS.CSC program presented the paper titled “Hunting Malware: An Example using Gh0st” at 2017 International Conference on Computational Science and Computational Intelligence (CSCI’17), Las Vegas, NV, December 14-16, 2017.
- 30) Drs. Tim Kroecker and Laurent Njilla visited ODU and NSU on January 31 - February 2, 2018. They met with COE faculty and students to discuss research and program related topics. They also toured the new Cybersecurity Complex at NSU and met with students who were interested in the COEs summer internship program at DoD facilities. In addition, accompanied by Dr. George Hsieh, they met with Dr. Leroy Hamilton, Interim Provost and Vice President for Academic Affairs at NSU to discuss topics of mutual interests.
- 31) Dr. Jonathan Graham attended the Virginia Cyber Range Executive Committee meeting at Virginia Military Institute, February 23, 2018, Lexington, VA.
- 32) Dr. Deepak Tosh presented a talk titled “Redefining Cybersecurity in the Blockchain Era” at Dept. of Math and Computer Science, Elizabeth City State University, March 2018, NC.
- 33) Drs. Tim Kroecker and Laurent Njilla visited ODU and NSU on April 2-5, 2018. They met with COE faculty and students to discuss research and program related topics.
- 34) During the ribbon-cutting event for the new Cybersecurity Complex at NSU on April 23, 2018, the COE hosted tours for Governor Ralph Northam of the Commonwealth of Virginia, and

other guests including the personnel from SPAWAR. Ms. Evelyn Kent participated in the event including the ribbon-cutting ceremony as a special guest.

35) Dr. Abdinur Ali presented a talk titled “Machine Learning Approaches for Malware Detection,” Mathematical Association of America, Maryland-District of Columbia-Virginia Section, Lexington, VA, April 13-14, 2018.

36) Dr. Abdinur Ali presented a talk titled “Web Mining based on Machine Learning Techniques” at Society of Industrial & Applied Mathematics Conference, ODU, Norfolk, VA, April 21, 2018.

37) Gayane Grigoryan presented the paper entitled “Assessing the Impact of Cyberloafing on Cyber Risk” at the SpringSim 18 conference in Baltimore, MD, April 2018.

38) Hamdi Kavak presented the paper entitled “Big Data, Agents, and Machine Learning: Towards a Data-Driven Agent-based Modeling Approach” at the SpringSim 18 Conference, Baltimore, MD, April 2018.

39) Jose Padilla presented the simulation platform and its impact on rapid modeling to members of the US Naval Special Warfare Development Group (NSWDG).

LIST OF SYMBOLS, ABBREVIATIONS, AND ACRONYMS

AI	Artificial Intelligence
AP	Access Point
API	Application Programming Interface
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BP	Belief Propagation
BPSK	Binary Phase Shift Keying
BS	Base Station
CAP	Computational Access Point
CG	Code Generator
CNN	Convolutional Neural Network
CO	Carbon Monoxide
COTS	Commercial Off-The-Shelf
CREDIT	The Center of Excellence in Research and Education for Big Military Data Intelligence
DL	Deep Learning
DMAT	Distributed Multi-dimensional Array Toolkit
DNN	Deep Neural Networks
DoD	Department of Defense
DRF	Dominant Resource Fairness
DST	Dempster-Shafer Theory
DSTS	Dynamic Series-Time Structure
DSmT	Dezert-Smarandache Theory
ECC	Elliptic Curve Cryptosystem
ETSI	European Telecommunications Standards Institute
FFT	Fast Fourier Transform
GA	Global Array
GACA	Genetic Algorithm based Computation Algorithm
GPU	Graphics Processing Unit

GRU	Gated Recurrent Unit
HARNN	Hierarchical Attention RNN
HDFS	Hadoop Distributed File System
HPC	High Performance Computing
IBC	ID-based cryptography
IoT	Internet of Things
JTORA	Joint Task Off Loading and Resource Allocation
LDM	Large Data Management
LIWC	Linguistic Inquiry and Word Count
LOD	Level of Detail
LOED	Leave-One-Event-Out
LSTM	Long Short-Term Memory
MCC	Mobile Cloud Computing
MDP	Markov Decision Process
MEC	Mobile Edge Computing
MES	Mobile Edge Computing Server
MINLP	Mixed Integer Non-Linear Programming
ML	Machine Learning
MSE	Mean Square Error
MTFNN	Multi-Task learning based Feedforward Neural Network
MTL	Multi-Task learning
MU	Mobile User
NLP	Natural Language Processing
NPB	NASA Parallel Benchmark
OFDMA	Orthogonal Frequency Division Multiple Access
PAAS	Platform as a Service
PAC	Probably Approximately Correct
PCA	Principal Components Analysis
PSO	Particle Swarm Optimization

PUF	Physically Unclonable Functions
PVAMU	Prairie View A&M University
QoS	Quality-of-Service
RDD	Resilient Distributed Dataset
RFID	Radio Frequency IDentification
RLTBB	Reformulation Linearization Technique-based Branch-and-Bound
RCNN	Recurrent CNN
RNN	Recurrent Neural Network
SAC	Seismic Analytics Cloud
SBB	Spatial Branch and Bound
SBT	Simple Build Tool
SEG	Society of Exploration Geophysicists
SGD	Stochastic Gradient Descent
SINR	Signal-to-Interference-plus-Noise Ratio
SOM	Self-Organizing Map
SQL	Structured Query Language
TDSL	Two-path Deep Semi-supervised Learning
TF-IDF	Term Frequency-Inverse Document Frequency
THECB	Texas Higher Education Coordinating Board
TPU	Tensor Processing Unit
UAV	Unmanned Aerial Vehicle
VM	Virtual Machine
WLAN	Wireless Local Area Network
WSN	Wireless Sensor Network