U.S. ARMY COMBAT CAPABILITIES DEVELOPMENT COMMAND

Using Machine Learning for Network Intrusion Detection

Dr. Nandi Leslie

Contractor: Raytheon Technologies

CCDC Army Research Laboratory

As of 24 Jun 2020

PROACTIVE ADVERSARIAL MODELING PROJECT OVERVIEW (FY18-25)

Goal & Objectives:
- Improve and complement attack detection
- Enhance prevention (e.g., deceptive network)
- Mitigate successful attacks (e.g., redirecting attackers)

Approach:
- Learn and predict adversaries' exploit preferences
- Intrusion forecasting on mobile ad hoc network (MANETs) supporting adaptive deception

Impact & Capability:
- Enhanced network robustness
- Defensive advantage against adversaries
- Deeper understanding of the adversaries' TTPs for network intrusions and reconnaissance

Key Stakeholders:
- FREEDOM ERP
- Cybersecurity and IoBT CRA
- Tech Transition partners: C5ISR Center/I2WD and S&TCD; and DARPA DSO SI3-CMD

GAPS & RESEARCH QUESTIONS

Limitations and gaps
- Previous work was mainly enterprise-focused for detecting enterprise attack campaigns
- Limited empirical data on adversarial TTPs in honeynets
- **Research questions?**
- Is it possible to develop predictive network-security and resilience models of adversarial network processes? What are the related constraints?

RECENT PROGRESS

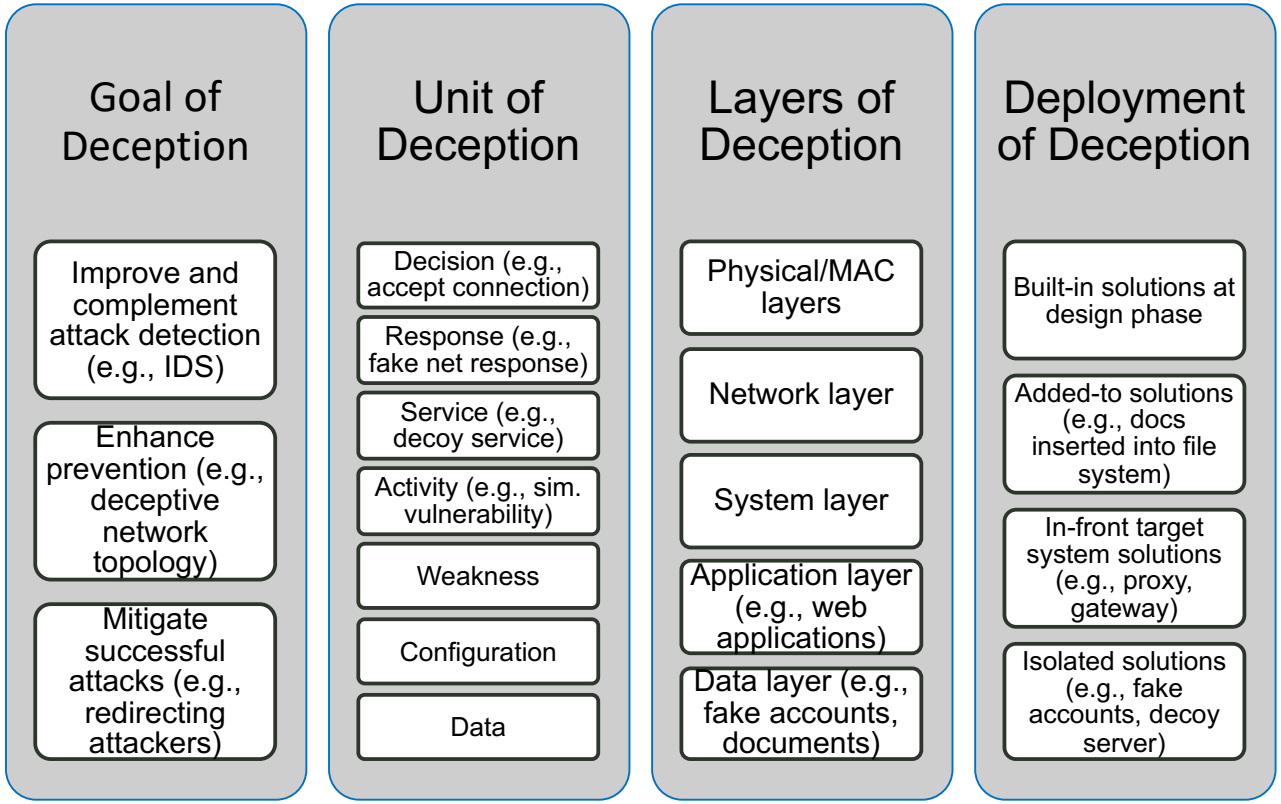
- Developed models to capture how adversaries learn details about the target's network (Albanese et al., '20)
- Developed neural machine translation (NMT) models to generate fake network traffic (Basu et al., '19)
- Demonstrated that we can uniformly learn the adversary's preferences using data from a modest number of deception strategies (Shi et al., '19)
- Formulated the IoBT domain as a graph-learning problem through an adversarial lens (Park et al., '19)

OUTLINE

- Provide background
- Present related work
- Provide data description for network traffic in CTU-13 Scenarios
- Present Semi-Supervised Learning for Exploits and Exploit Kits (SLEEK) algorithm
- Present metrics and SLEEK cases considered
- Analyze prediction results from SLEEK experiments with CTU-13 Scenarios
- Present conclusions

BACKGROUND

DIMENSIONS OF CYBER DECEPTION



Han, X., Kheir, N., & Balzarotti, D. (2018). Deception techniques in computer security: A research perspective. *ACM Computing Surveys (CSUR)*, *51*(4), 1-36.

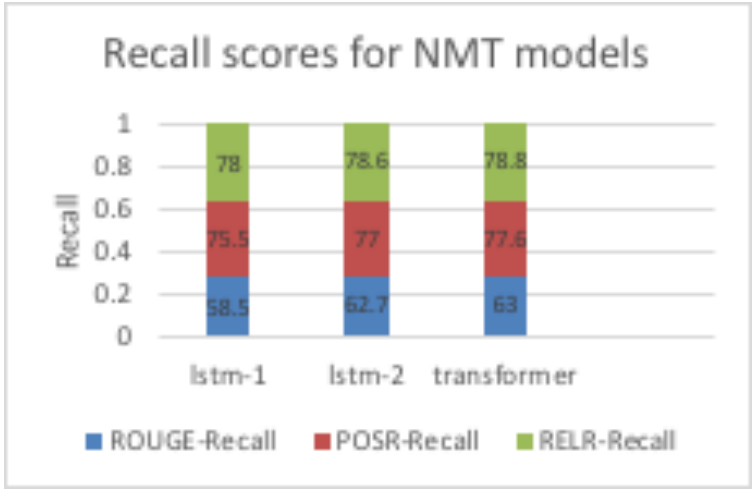
TECHNICAL APPROACH

- Develop novel approaches to disguise a mobile network and impair the attacker's decision-making with false information
- Identify cyber deception techniques relevant to tactical networks considering computational resource constraints
- Autonomously prioritize units and layers of deception and deployment of deception
- Learn the adversary's COAs and adaptively automate the deployment of deceptive measures
- Determine which deception strategies are optimal, using game theory

FY20 SELECTED 6.1 RESEARCH ACCOMPLISHMENTS (CYBER CRA LFD)

- **Goal:** Deceive the adversary into believing that honey users are real sys admins/users

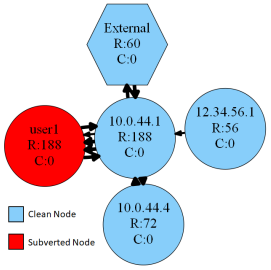
- **Accomplishments:**
  - Used neural machine translation (NMT) models to generate spearphishing e-mails
  - Proposed metrics that capture both topical relevance and the grammatical structure of text generated by NMT models
  - Found that the spearphishing filter performs poorly in distinguishing between real and automatically-generated content



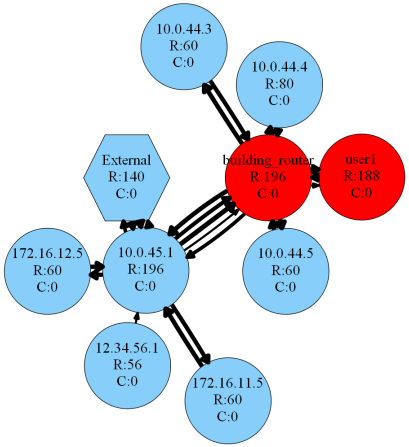
Basu, C., Venkatesan, S., Chiang, C. J., **Leslie, N. O.**, & **Kamhoua, C. A.** (2019). Generating Targeted E-mail Content at Scale Using Neural Machine Translation. In *DYnamic and Novel Advances in Machine Learning and Intelligent Cyber Security (DYNAMICS) Workshop*.

FY20 SELECTED 6.2 RESEARCH ACCOMPLISHMENTS (CYBER CRA LFD)

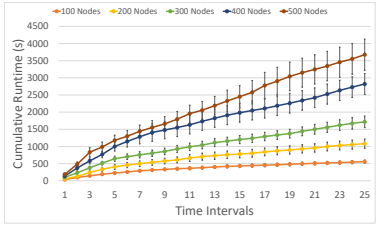
- **Goal**: Learn the adversary's COAs and adaptively automate the deployment of deceptive measures
- **Accomplishments:**
  - Developed a novel model to capture how stealthy adversaries acquire knowledge about the target network's topology and establish their foothold.
  - Quantified the cost and reward, from the adversary's perspective, of compromising individual nodes and maintaining control over those nodes.
  - Evaluated our model through simulations in the CyberVAN testbed
  - Demonstrated how our model can guide the deployment of defensive capabilities (e.g., honeypots) to influence the behavior of adversaries



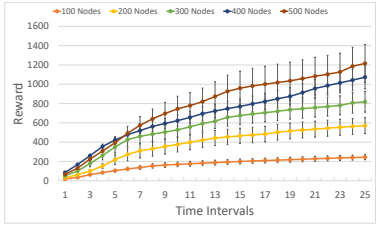
(a) $S' = \{\text{user1}\}$



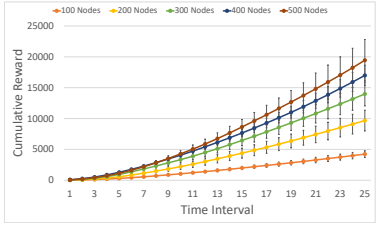
(b) $S' = \{\text{user1, building\_router}\}$



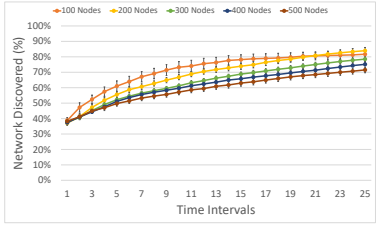
(a) Runtime



(b) Reward



(c) Cumulative Reward



(d) Percentage of network nodes discovered

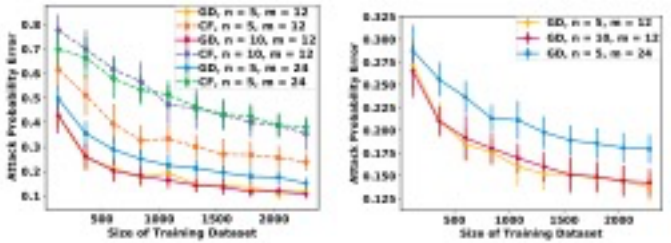
Albanese, M., Chadha, R., Chiang, C. J., **Kamhoua, C. A., Leslie, N. O.**, Pham, L., Venkatesan, S. (2020, accepted). A Quantitative Framework to Model Reconnaissance by Stealthy Attackers and Support Deception-Based Defenses. In *IEEE CNS.*

FY20 SELECTED 6.1/6.2 RESEARCH ACCOMPLISHMENTS (CYBER CRA)

- **Goal:**
  - Deception mitigates the defender's loss by misleading the attacker to make suboptimal decisions.
  - In order to formally reason about deception, we introduce the *feature deception game (FDG)*, a domain-independent game theoretic model and present a learning and planning framework.
- **Accomplishments:**
  - Demonstrated that we can uniformly learn the adversary's preferences using data from a modest number of deception strategies.
  - Proposed an approximation algorithm for finding the optimal deception strategy and show that the problem is NP-hard.
  - Performed extensive experiments to empirically validate our methods and results.

Feature	Observable value	Hidden value
Operating system	Windows 2016	RHEL 7
Service version	v1.2	v1.4
IP address	10.0.1.2	10.0.2.1
Open ports	22, 445	22, 1433
Round trip time for probes [Shamsi *et al.*, 2014]	16 ms	84 ms

Table 1: Some relevant features for cybersecurity



(a) Learning 1-layer score function (b) Learning 3-layer score function

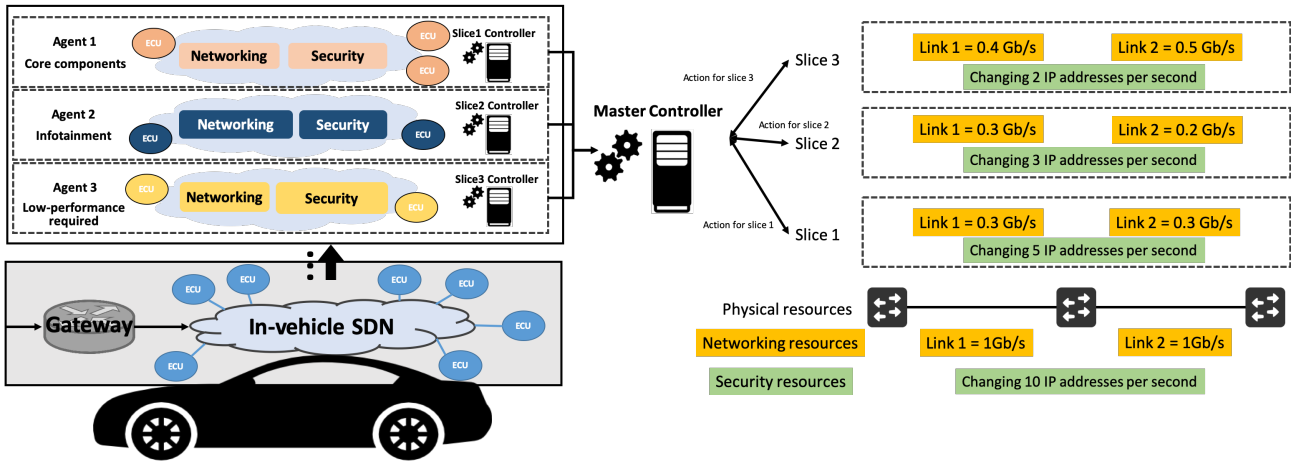
1. Milani, S., **Chan, K.**, Fang, F., **Leslie, N. O., & Kamhoua, C. A**. (2020, in-progress). Iterated Deception Games.
2. Shi, Z. R., Procaccia, A. D., **Chan, K. S.,** Venkatesan, S., Ben-Asher, N., **Leslie, N. O., Kamhoua, C. A.,** & Fang, F. (2019). Feature Deception Games. *IJCAI 2019 Workshop on Strategic Reasoning*.
3. Shi, Z. R., Procaccia, A. D., **Chan, K. S**., Venkatesan, S., Ben-Asher, N., **Leslie, N. O., Kamhoua, C. A.,** & Fang, F. (2019). Learning and Planning in Feature Deception Games. In *ACM EC 2019 Workshop on Learning in Presence of Strategic Behavior.* Phoenix, Arizona: ACM.

FY20 SELECTED 6.1 RESEARCH ACCOMPLISHMENTS

Goals: Examine the current state-of-the-art in vehicular communications security

Accomplishments:

- Developed a highly secure, resilient, and affordable MTD-based proactive defense mechanism, which achieves multiple objectives of minimizing system security vulnerabilities and defense cost while maximizing service availability.
- Proposed a multi-agent Deep Reinforcement Learning (mDRL)-based network slicing technique that can help determine two key resource management decisions: (1) link bandwidth allocation to meet quality-of-service requirements and (2) the frequency of triggering IP shuffling as an MTD operation.
- Applied this strategy in a tactical in-vehicle network that uses **software-defined networking (SDN) technology** to deploy the IP-shuffling-based MTD by changing IP addresses assigned to electronic control unit (ECU) nodes.



Yoon, S., Cho, J., Kim, D. S., **Moore, T. J., Nelson, F. F.**, Lim, H., **Leslie, N. O.**, & **Kamhoua, C. A.** (2020). Moving Target Defense for In-Vehicle Software Defined Networking: IP Shuffling in Network Slicing with Multiagent Deep Reinforcement Learning. In *SPIE, AI and ML for Multi Domain Operations Applications II.*

FY20 6.2 RESEARCH ACCOMPLISHMENTS: INTRA/INTER VEHICULAR NETWORKS

Goals:

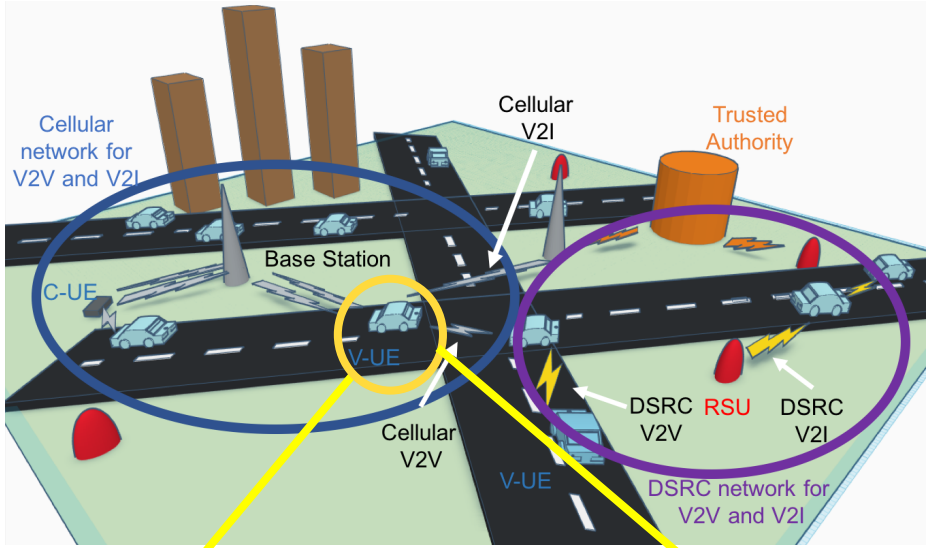
- Examine the current state-of-the-art in vehicular communications security
- Explore the use of 5.6 GHz band for cooperative Intelligent Transportation System (ITS)

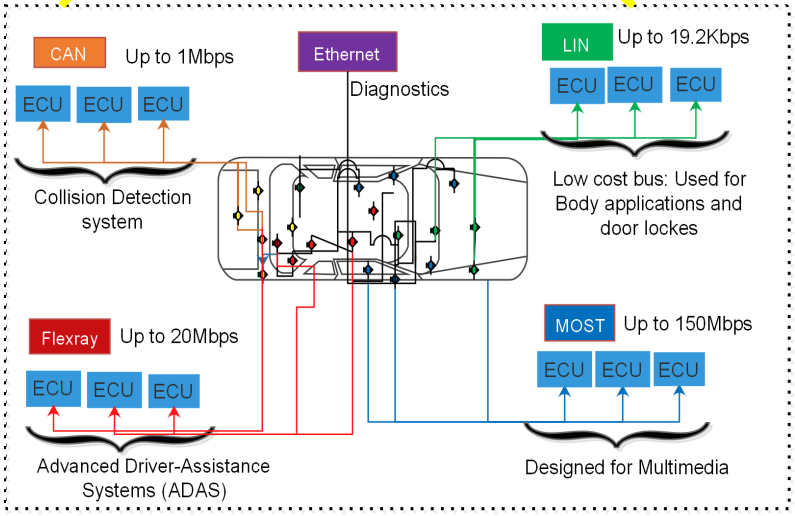
Accomplishments:

- Identified threat scenarios in the 2 paradigms for vehicular communications:
  - Cellular vehicle-to-everything (cV2X) and
  - Vehicular ad hoc networks (VANET)
- Compared the most common anomaly detection techniques used in these vehicular networks
- Examined machine learning and knowledge-based methods for anomaly-based intrusion detection systems in ITS

In-progress research:

- Currently using tools SUMO and NS-3 for simulation for vehicular routes
- Create malicious routes and vehicles in network and nodes (e.g., vehicles and RSUs) that are able to observe traffic
- Make changes to multi-hop protocol so that malicious nodes see unwanted traffic and target false nodes.



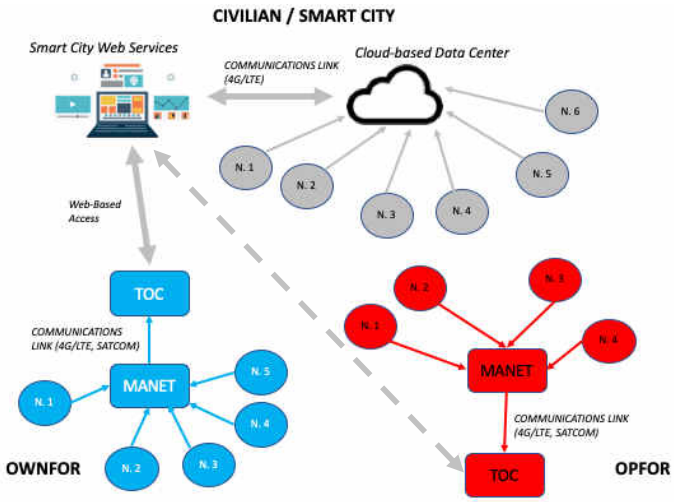


Dayal, A., **Leslie, N. O., Kamhoua, C. A.,** Marojevik, V., & Reed, J. (2020, submitted). Taxonomy of Anomaly Based Intrusion Detection Systems in Vehicular Communications. In IEEE Vehicular Technology Magazine.

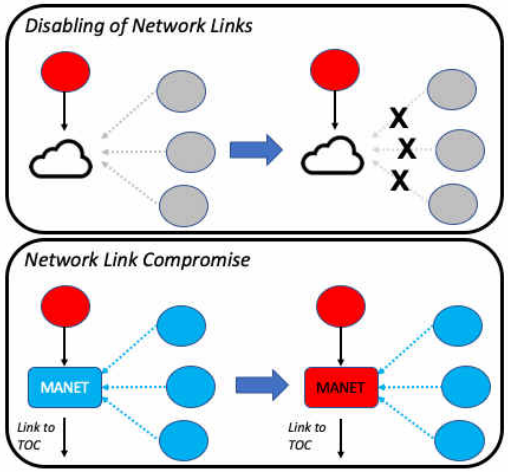
FY20 SELECTED 6.1 RESEARCH ACCOMPLISHMENTS (IOBT CRA)

- **Goal:** Advance current IoBT efforts with a collection of prior-developed cybersecurity techniques

- **Accomplishments:**
  - Reviewed for applicability to IoBT operational environments:
    - Diverse asset ownership
    - Degraded networking infrastructure
    - Adversarial activities
  - Covered research techniques focused on two themes:
    - Supporting trust assessment for known/unknown IoT assets
    - Ensuring continued trust of known IoT assets and IoBT systems



Configuration of OWNFOR (Blue) and OPFOR (Red) assets within a Smart City environment (Grey)

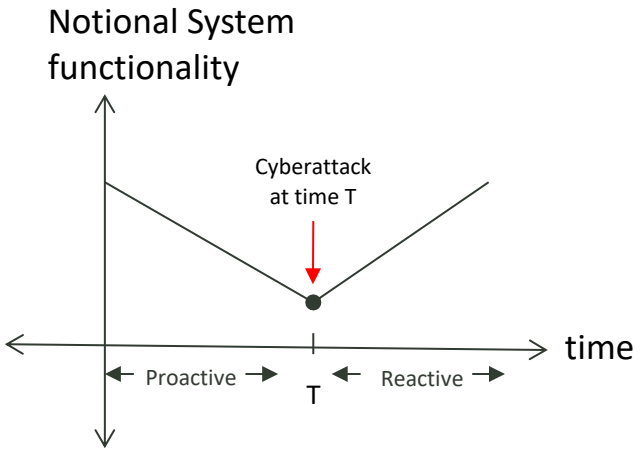


Types of interaction between OPFOR IoT assets and Civilian + OWNFOR IoT networks

Agadakos, I., Ciocarlie, G. F., Copos, B., Emmi, M., **George, J., Leslie, N., & Michaelis, J.** (2019). Application of Trust Assessment Techniques to IoBT Systems. In *MILCOM 2019-2019 IEEE Military Communications Conference (MILCOM)* (pp. 833-840). IEEE.

WHY IS NETWORK RESILIENCE IMPORTANT?

- **Botnet attacks represent a serious threat to commercial and governmental networks**
- **Cyber-physical systems (CPS), including Internet of Things (IoT) have severe results if there are failures**
  - Increased risk of cyberattacks
  - Energy network (smart grid)
  - Transportation systems and large industrial facilities
- **Proactive techniques for network resilience include redundancy and compartmentalization**
  - Redundancy allows to tolerate attacks to a certain extent
  - Compartmentalization attempts to restrict the cyberattack locally and prevent its expansion across the entire network
  - Configuration and set-up of intrusion detection and prevention systems (IDS/IPS)
- **Reactive techniques follow this high-level, three-step approach**
  - Detecting an attack
  - Mitigating its impacts
  - Restoring a system's usual operation

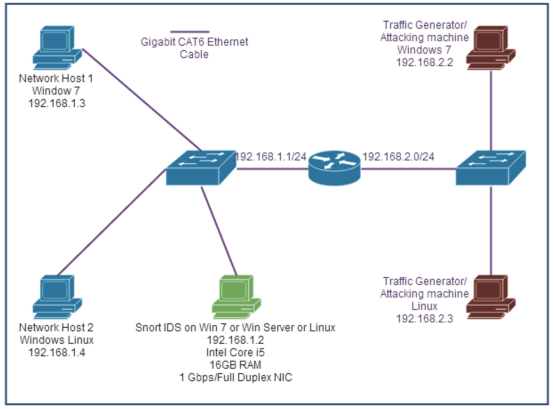


Ganin et al., 2017. Resilience and efficiency in transportation networks. *Science Advances.*
Linkov et al., 2018. Risk and resilience must be independently managed. *Nature.*

NIDS BACKGROUND

Intrusion Detection System (IDS). "…a real-time intrusion-detection expert system that aims to detect a wide range of security violations ranging from attempted break-ins by outsiders to system penetrations and abuses by insiders." (Denning, 1987)

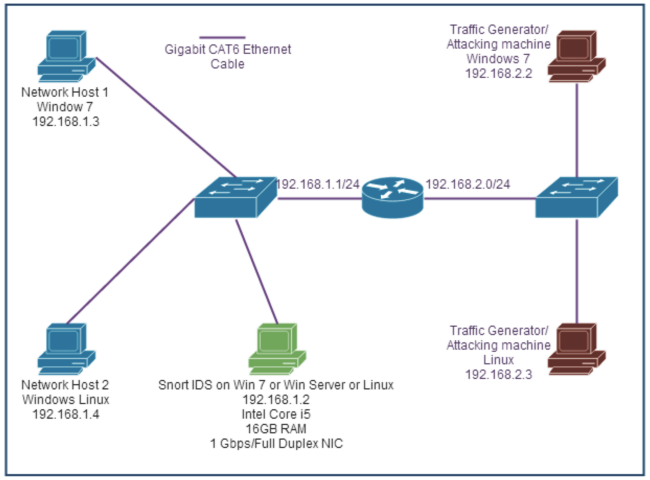
- **Detection models and algorithms**
  - Signature-based
  - Anomaly-based
- **Host-based IDS (HIDS)**
- **Network IDS (NIDS)**



A typical test-bed for Snort NIDS evaluation (Karim et al., 2017 *Computers*)

NIDS CHALLENGES

- **Anomaly-based NIDS tend to have high false positive rates (FPR)**
  - High FPRs may cause personnel to disregard those tools
  - Results in unreported network breaches
- **Signature-based NIDS may have high false-negative rates**
  - Misclassify cyberattacks with unknown exploits as benign
  - Examples of NIDS include Snort and Bro



A typical test-bed for Snort NIDS evaluation (Karim et al., 2017 *Computers*)

RELATED WORK

- **Garcia *et al.* (2014) examined 5 Scenarios (i.e., Scenarios 1-2, 6, and 8-9) in the CTU-13 Botnet Packet Capture (pcap) Scenarios**
- **Both of their botnet detection methods use only the NetFlow files in CTU-13**
- **Tradeoff in prediction accuracy results across many detection methods considered (Garcia et al., 2014)**
  - **Scenarios 1 and 2:** If precision > 0.8, then TPR < 0.1 for those anomaly detection methods considered
  - **Scenario 6:** Best prediction performance for methods is precision > 0.79 and TPR < 0.71. For example, one model resulted in precision = 0.9 with TPR = 0.7
  - **Scenarios 8 and 9:** Best-performing models resulted in precision = 0.5 and TPR = 1.0. For all other models considered in Garcia et al. (2014), TPR < 0.3

ID	Bot	Characteristic	Total flows	Botnet flows	Normal flows	Background flows
1	Neris	IRC, SPAM, Click Fraud	2,824,636	39,933	30,387	2,754,316
2	Neris	IRC, SPAM, Click Fraud, FTP	1,808,122	18,839	9,120	1,780,163
3	RBot	IRC, Port Scan, US	4,710,638	26,759	116,887	4,566,992
4	RBot	IRC, DDOS, US	1,121,076	1,719	25,268	1,094,089
5	Virut	SPAM, Port Scan, HTTP	129,832	695	4,679	124,458
6	Mentri	Port Scan	585,919	4,431	7,494	573,994
7	Sogou	HTTP	144,077	37	1,677	142,363
8	Merli	Port Scan	2,954,230	5,052	72,822	2,876,356
9	Neris	IRC, SPAM, Port Scan, Click Fraud	2,753,884	17,880	43,340	2,692,664
10	Rbot	IRC, DDOS, US	1,309,791	106,315	15,847	1,187,629
11	RBot	IRC, DDOS, US	107,251	8,161	2,718	96,372
12	NSIS.ay	PP	325,471	2,143	7,628	315,700
13	Virut	SPAM, PS, HTTP	1,925,149	38,791	31,939	1,854,419

RELATED WORK

- **Leslie, Martone, and Weisman (2018) developed a semi-supervised NIDS algorithm**
- **We assessed 3 Scenarios in CTU-13 dataset using the NetFlow files to accurately detect botnet behaviors using K-means clustering algorithm**
- **We examined Scenarios 4, 10, and 11 in CTU-13**
  - Each of these 3 Scenarios use the IRC protocol to perform DDoS attacks
  - Scenario 4 has a lower prevalence of botnet flows than Scenarios 10 and 11

ID	Bot	Characteristic	Total flows	Botnet flows	Normal flows	Background flows
1	Neris	IRC, SPAM, Click Fraud	2,824,636	39,933	30,387	2,754,316
2	Neris	IRC, SPAM, Click Fraud, FTP	1,808,122	18,839	9,120	1,780,163
3	RBot	IRC, Port Scan, US	4,710,638	26,759	116,887	4,566,992
4	RBot	IRC, DDOS, US	1,121,076	1,719	25,268	1,094,089
5	Virut	SPAM, Port Scan, HTTP	129,832	695	4,679	124,458
6	Mentri	Port Scan	585,919	4,431	7,494	573,994
7	Sogou	HTTP	144,077	37	1,677	142,363
8	Merli	Port Scan	2,954,230	5,052	72,822	2,876,356
9	Neris	IRC, SPAM, Port Scan, Click Fraud	2,753,884	17,880	43,340	2,692,664
10	Rbot	IRC, DDOS, US	1,309,791	106,315	15,847	1,187,629
11	RBot	IRC, DDOS, US	107,251	8,161	2,718	96,372
12	NSIS.ay	PP	325,471	2,143	7,628	315,700
13	Virut	SPAM, PS, HTTP	1,925,149	38,791	31,939	1,854,419

Leslie et al. (2018). The Internet of Things (IoT): Computational Modeling in Congested and Contested Environments. In *Proceedings of the NATO IST-152 Workshop on Autonomous Agents for Cyber Defence.*

RELATED WORK

- **Leslie et al. (2018) showed their K-means based algorithms for IDS have precision values above 0.8, and FPR below 0.02 for CTU-13 scenarios**
- **We showed better performance for Scenarios 10 and 11 than Scenario 4 in CTU-13**
- **Our model yielded better prediction performance results than many previous methods published for these Scenarios**

Table 2. The k-means performance results for 3 of the CTU-13 botnet pcap scenarios characterized in Table 1: IDs 4, 10, and 11.

	ID 4	ID 10	ID 11
accuracy	1.00	0.97	0.97
precision	0.98	0.85	0.82
recall	0.26	0.90	0.89
FPR	0.0	0.02	0.02

Leslie et al. (2018). The Internet of Things (IoT): Computational Modeling in Congested and Contested Environments. In *Proceedings of the NATO IST-152 Workshop on Autonomous Agents for Cyber Defence.*

DATA DESCRIPTION FOR CTU-13 BOTNET SCENARIOS

- **García (2013) at the Czech Tech University (CTU) published online thirteen botnet scenarios, CTU-13**
- **Each scenario includes**
  - Botnet pcap file
  - Labeled NetFlow file
  - README file, with the capture time line and the original malware executable binary from 2011 data
- **García (2013) was not possible to publish the complete pcap file with the background and normal packets because they contain private information**

ID	Bot	Characteristic	Total flows	Botnet flows	Normal flows	Background flows
1	Neris	IRC, SPAM, Click Fraud	2,824,636	39,933	30,387	2,754,316
2	Neris	IRC, SPAM, Click Fraud, FTP	1,808,122	18,839	9,120	1,780,163
3	RBot	IRC, Port Scan, US	4,710,638	26,759	116,887	4,566,992
4	RBot	IRC, DDOS, US	1,121,076	1,719	25,268	1,094,089
5	Virut	SPAM, Port Scan, HTTP	129,832	695	4,679	124,458
6	Mentri	Port Scan	585,919	4,431	7,494	573,994
7	Sogou	HTTP	144,077	37	1,677	142,363
8	Merli	Port Scan	2,954,230	5,052	72,822	2,876,356
9	Neris	IRC, SPAM, Port Scan, Click Fraud	2,753,884	17,880	43,340	2,692,664
10	Rbot	IRC, DDOS, US	1,309,791	106,315	15,847	1,187,629
11	RBot	IRC, DDOS, US	107,251	8,161	2,718	96,372
12	NSIS.ay	PP	325,471	2,143	7,628	315,700
13	Virut	SPAM, PS, HTTP	1,925,149	38,791	31,939	1,854,419

DATA DESCRIPTION FOR CTU-13 BOTNET SCENARIOS

- **Implemented data from thirteen botnet pcap scenarios called CTU-13 into SLEEK**
- **IRC, P2P, or HTTP protocols used in these pcap scenarios**
- **Botnets are characterized by either**
  - Sending SPAM
  - Performing port scan (PS)
  - Performing click fraud (CF)
  - Performing distributed denial of service (DDoS)

ID	Bot	Characteristic	Total flows	Botnet flows	Normal flows	Background flows
1	Neris	IRC, SPAM, Click Fraud	2,824,636	39,933	30,387	2,754,316
2	Neris	IRC, SPAM, Click Fraud, FTP	1,808,122	18,839	9,120	1,780,163
3	RBot	IRC, Port Scan, US	4,710,638	26,759	116,887	4,566,992
4	RBot	IRC, DDOS, US	1,121,076	1,719	25,268	1,094,089
5	Virut	SPAM, Port Scan, HTTP	129,832	695	4,679	124,458
6	Mentri	Port Scan	585,919	4,431	7,494	573,994
7	Sogou	HTTP	144,077	37	1,677	142,363
8	Merli	Port Scan	2,954,230	5,052	72,822	2,876,356
9	Neris	IRC, SPAM, Port Scan, Click Fraud	2,753,884	17,880	43,340	2,692,664
10	Rbot	IRC, DDOS, US	1,309,791	106,315	15,847	1,187,629
11	RBot	IRC, DDOS, US	107,251	8,161	2,718	96,372
12	NSIS.ay	PP	325,471	2,143	7,628	315,700
13	Virut	SPAM, PS, HTTP	1,925,149	38,791	31,939	1,854,419

EXAMPLE SESSION FROM NETFLOW FILE

- **Sessions are labeled as botnet, background, or normal**
- Selected features are categorical (see arrows), and others are numerical
- Ports can be considered either categorical or numerical

Start Time	*Duration*	Protocol	*Source IP Address*	*Source Port*	*Direction*	*Destination IP Address*	*Destination Port*	*State*	*sToS*	dToS	*Total Packets*	*Total Bytes*	Source Bytes	Label
2011/08/18 16.700994	0	icmp	147.32.84.165	0x0005	->	147.32.96.69		RED	0		1	1066	1066	**Flow=From-Botnet-V51-1-ICMP**

EXAMPLE SESSION FROM NETFLOW FILE

- Sessions are labeled as botnet, background, or normal
- **Selected features are categorical (see arrows), and others are numerical**
- Ports can be considered either categorical or numerical

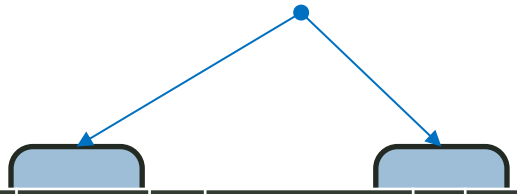
Original feature space

Start Time	*Duration*	Protocol	*Source IP Address*	*Source Port*	*Direction*	*Destination IP Address*	*Destination Port*	*State*	*sToS*	dToS	*Total Packets*	*Total Bytes*	Source Bytes	Label
2011/08/18 16.700994	0	icmp	147.32.84.165	0x0005	->	147.32.96.69		RED	0		1	1066	1066	**Flow=From-Botnet-V51-1-ICMP**

Categorical features

EXAMPLE SESSION FROM NETFLOW FILE

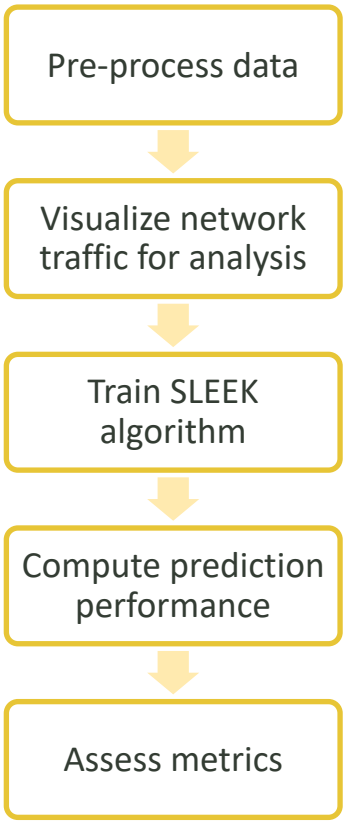
- Sessions are labeled as botnet, background, or normal
- Selected features are categorical (see arrows), and others are numerical
- **Ports can be considered either categorical or numerical**



Start Time	*Duration*	Protocol	*Source IP Address*	*Source Port*	*Direction*	*Destination IP Address*	*Destination Port*	*State*	*sToS*	*dToS*	*Total Packets*	*Total Bytes*	Source Bytes	Label
2011/08/18 16.700994	0	icmp	147.32.84.165	0x0005	->	147.32.96.69		RED	0		1	1066	1066	**Flow=From-Botnet-V51-1-ICMP**

SLEEK MODELING METHODOLOGY

- **Pre-process IP traffic data**
  - Cross validate input data
  - Compute IP distance metrics
  - Convert other categorical features to numerical
- **Visualize labelled network as a colored graph**
  - **Blue nodes** are IP addresses sending benign traffic
  - **Red nodes** represent IP addresses sending botnet traffic
  - Links represent network sessions between nodes
- **Implement clustering and classification machine learning algorithms into SLEEK**
  - K-means and Gaussian Mixture Model (GMM) algorithms for clustering with semi-supervised approach
  - k-Nearest Neighbor (k-NN) algorithm for classification has best results
- **Examine SLEEK's prediction performance for identifying botnets in testing phase**
- **Measure the significance of features in metrics results**



PRE-PROCESSING NETWORK TRAFFIC DATA

- **Cross Validation**
  - Divide each Scenario in CTU-13 into training and testing datasets using K-fold cross validation, where K = 5
- **Feature Space Modification and Normalization**
  - Temporarily modify training set by removing feature vectors with IP addresses, X'
  - Convert other categorical feature vectors to numerical values with one-hot encoding
  - Normalize modified training data without IP addresses, $\widehat{X}$'
  - Find 2 centroids for malicious, C^+, and benign, C^-, samples in the modified training set, $\widehat{X}$'

CONVERTING CATEGORICAL FEATURES TO NUMERICAL

- Label all sessions with “Botnet” or “Normal”
- Eliminate sessions labeled as “Background” traffic
- Convert categorical features to numerical features with one-hot encoding for all categorical features except IP addresses

Training Examples	Protocols
1	TCP
2	UDP
3	ICMP
⋮	
k	TCP

One-hot Encoding →

Training Examples	TCP	UDP	ICMP
1	1/3	0	0
2	0	1/3	0
3	0	0	1/3
⋮			
k	1/3	0	0

METRIC FOR IP ADDRESSES IN FEATURE SPACE

- **Let x and y be 32-bit IPv4 addresses, where each byte of x is represented by x_j such that $x = x_1.x_2.x_3.x_4$**
- **Define distance D(x, y) between IP addresses as**

$$D(x, y) = \sum_{i=1}^{4} a^{4-i}(x_i \neq y_i), \text{ where a > 1 constant}$$

x =	147.	32.	84.	165
y =	132.	38.	67.	128

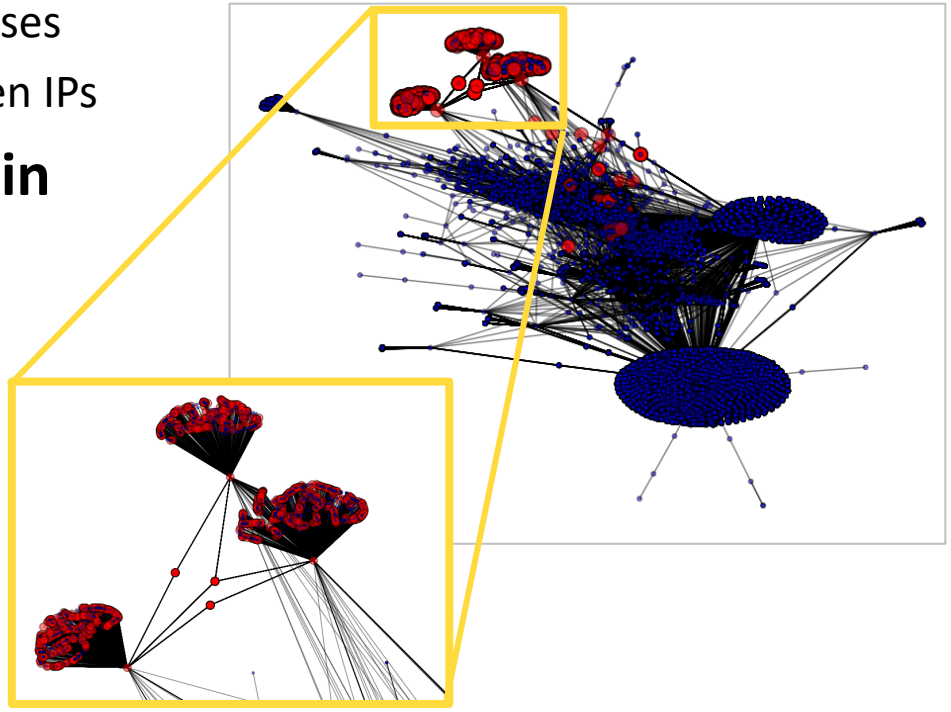
Let a = 2.
Then, distance D = 8 + 4 + 2 + 1
D = 15

x =	147.	32.	84.	165
y =	147.	32.	70.	130

Let a = 2.
Distance between IP addresses is
D = 2 + 1
D = 3

VISUALIZING NETWORK TRAFFIC

- **Use SLEEK visualization module**
  - Red nodes are malicious IP addresses
  - Blue nodes are normal IP addresses
  - Links signify connections between IPs
- **Scenario 12 Characteristics in CTU-13**
  - Use P2P protocol
  - Synchronization attack
  - Botnet flows are 67.97% of total sessions
  - 3 Bots in network

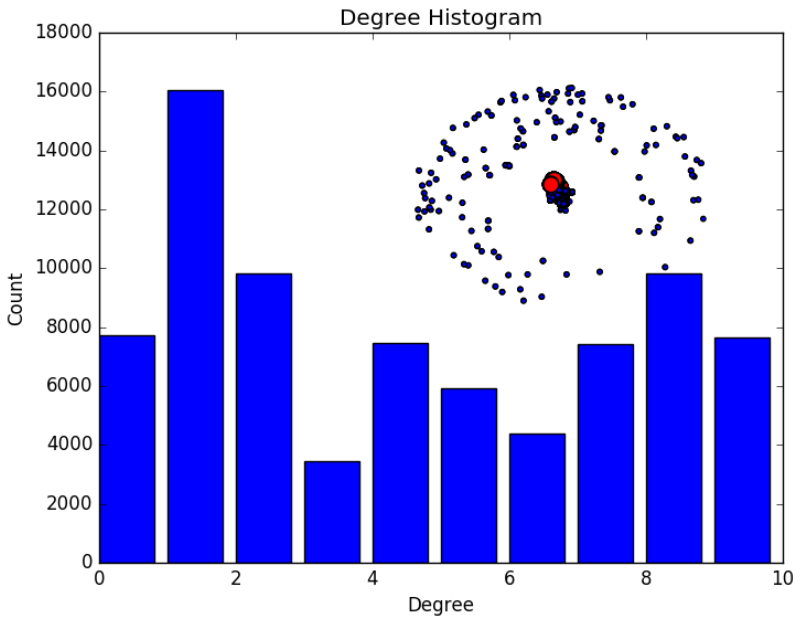


SLEEK NETWORK VISUALIZATION

Leslie, N. O. Using Semi-Supervised Learning for Flow-Based Network Intrusion Detection. In *Proceedings of the ICCRTS, 6-9 November 2018, Pensacola, FL*.

GRAPH DEGREE HISTOGRAM

- **Implemented network traffic data into SLEEK graph analysis module**
- **Present degree histogram with network graph inset**
- **Bi-modal distribution in degree histogram of network sessions data**
- **Scenario 12 Characteristics in CTU-13**
  - Use P2P protocol
  - Synchronization attack
  - Botnet flows are 67.97% of total sessions
  - 3 Bots in network



SLEEK: A MACHINE LEARNING-BASED APPROACH TO NIDS

- **Developed K-means clustering algorithm in a semi-supervised approach to assess SLEEK for the CTU-13 datasets**
- **Developed a GMM-based algorithm to assess SLEEK predictions for intrusion detection for CTU-13**
- **Developed a k-NN-based algorithm to assess SLEEK predictions for intrusion detection**
  - Assign each network session to the majority class (i.e., benign, botnet) of its closest neighbors, where *k* is a parameter
  - Best performing algorithm
    - Across 13 Scenarios in CTU-13
    - Each Scenario of CTU-13 was implemented into 3 configurations of IP distance metric for SLEEK

PREDICTION METRICS

- **CTU-13 Scenarios have a high prevalence of benign IP traffic**
  - True positive rate (TPR) and False positive rate (FPR) are often presented to show prediction performance of NIDS
  - These metrics alone are insufficient metrics for full view of NIDS prediction performance
- **Analyzed prediction metrics from test data, including**
  - Accuracy
  - Precision is a valuable metric from information theory that is a "true positive accuracy measure," $\frac{TP}{TP+FP}$
  - Recall is another prediction performance metric also known as true positive rate, $\frac{TP}{TP+FN}$
  - FPR

CASES IMPLEMENTED IN SLEEK

- Case 1. Exclude IP addresses from the feature space entirely
- Case 2. Include IP distance metric

METRICS RESULTS FOR SLEEK PREDICTIONS: CASE 1

- **Implemented Case 1 without IP Distance Metric in SLEEK module**
- **High prevalence of negatives in data**
  - Accuracy is inadequate for these cases
  - Precision and recall are better prediction performance metrics
- **SLEEK makes excellent predictions for most Scenarios in CTU-13 for this case**
  - SLEEK precision > 0.7
  - SLEEK accuracy > 0.9
  - If precision > 0.7, place green check

CTU-13 Scenarios		Botnet Type	Accuracy	Precision	Recall	FPR
✓	1	IRC, SPAM, CF	0.9922	0.7996	0.6160	0.00227
✓	2	IRC, SPAM, CF, FTP	0.9962	0.9431	0.7250	0.00057
✓	3	IRC, PS, US	0.9981	0.8834	0.7664	0.00058
	4	IRC, DDOS, US	0.9973	0.0505	0.0058	0.00040
	5	SPAM, PS, HTTP	0.9902	0.6786	0.3149	0.00177
✓	6	PS	0.9992	0.9712	0.9816	0.00048
	7	HTTP	0.9990	0.5000	0.0769	0.00008
	8	PS	0.9981	0.5314	0.7447	0.00137
✓	9	IRC, SPAM, PS, CF	0.9651	0.8249	0.7696	0.01588
✓	10	IRC, DDOS, US	0.9994	0.9995	0.9929	0.00004
✓	11	IRC, DDOS, US	0.9989	0.9988	0.9871	0.00010
	12	P2P	0.9937	0.5766	0.1820	0.00090
✓	13	SPAM, PS, HTTP	0.9902	0.8350	0.6590	0.00276

METRICS RESULTS FOR SLEEK PREDICTIONS: CASE 2

- **Implemented Case 2 with novel IP distance metric in SLEEK module**
- **SLEEK performs very well for several CTU-13 Botnet scenarios**
  - SLEEK doesn't detect botnets in most Scenarios for this case
  - SLEEK performs very well at detecting botnets performing DDoS attacks
  - 3 Scenarios in CTU-13 have precision > 0.7 & accuracy > 0.9
  - If precision > 0.7, place green check
  - SLEEK had no predicted positives for selected Scenarios

	CTU-13 Scenarios	Botnet Type	Accuracy	Precision	Recall	FPR
	1	IRC, SPAM, CF	0.9855	0.0000	0.0000	0.00227
	2	IRC, SPAM, CF, FTP	0.9884	nan	0.0000	0.00057
✓	3	IRC, PS, US	0.9943	0.7778	0.0013	0.00058
✓	4	IRC, DDOS, US	0.9975	0.0000	0.0000	0.00040
	5	SPAM, PS, HTTP	0.9930	nan	0.0000	0.00177
	6	PS	0.9917	nan	0	0.00048
	7	HTTP	0.9994	nan	0.0000	0.00008
	8	PS	0.9979	nan	0.0000	0.00137
	9	IRC, SPAM, PS, CF	0.9104	0.0000	0.0000	0.01588
✓	10	IRC, DDOS, US	0.9991	0.9999	0.9886	0.00004
✓	11	IRC, DDOS, US	0.9983	0.9988	0.9786	0.00010
	12	P2P	0.9931	0.0667	0.0023	0.00090
	13	SPAM, PS, HTTP	0.9792	nan	0.0000	0.00276

COMPARISON OF MACHINE LEARNING ALGORITHMS FOR DETECTING SPAM/CF

Scenario ID 2 (IRC, SPAM, CF, FTP)	k-NN	K-means	GMM
Precision	0.9431	0.0821	0.0003
Recall	0.7250	0.9799	0.0201
FPR	0.00057	0.1284	0.8693

- SLEEK (Case 1) with K-means algorithm has a higher recall for Scenario 2 than k-NN and GMM
- However, k-NN has a higher F_1-measure for Scenario 2.

Scenario ID 9 (IRC, SPAM, PS, CF)	k-NN	K-means	GMM
Precision	0.8249	0.1876	0.0583
Recall	0.7696	0.5015	0.4985
FPR	0.01588	0.2112	0.7825

- SLEEK (Case 1) with k-NN algorithm has better prediction performance than K-means and GMM for Scenario 9

COMPARISON OF MACHINE LEARNING ALGORITHMS FOR DETECTING DDOS ATTACKS

CTU-13 Datasets: DDoS Scenario 10	k-NN	K-means	GMM
Precision	0.9995	0.8046	0.0064
Recall	0.9929	0.8167	0.0690
FPR	0.00004	0.0175	0.9512

CTU-13 Datasets: DDoS Scenario 11	k-NN	K-means	GMM
Precision	0.9988	0.7888	0.5169
Recall	0.9871	0.8177	0.9843
FPR	0.00010	0.0181	0.0758

For both DDoS scenarios in CTU-13, IDs 10 and 11, SLEEK (Case 1) with k-NN algorithm has better prediction performance than than it has with K-means and GMM algorithms.

DISCUSSION

- **Distance metrics on the source and destination IP addresses features greatly impact results**
- **SLEEK performs exceptionally well with k-NN algorithm at detecting cyberattacks that in 13 Scenarios of CTU-13**
  - Perform port scans
  - Perform click fraud
  - Send spam
  - Use IRC protocol
- **SLEEK Case 1 with k-NN algorithm has the best prediction results**
  - SLEEK with K-means and GMM have poor performance results

SLEEK Case 1 has high # of FN for Scenario 5

Scenario 5 implemented in SLEEK Case 2 has no predicted positives

CTU-13 Scenarios	Botnet Type	Case 1 Accuracy	Case 2 Accuracy
1	IRC, SPAM, CF	**0.9922**	0.9855
2	IRC, SPAM, CF, FTP	**0.9962**	0.9884
3	IRC, PS, US	**0.9981**	0.9943
4	IRC, DDOS, US	**0.9973**	0.9975
5	SPAM, PS, HTTP	0.9902	0.9930
6	PS	**0.9992**	0.9917
7	HTTP	0.9990	**0.9994**
8	PS	**0.9981**	0.9979
9	IRC, SPAM, PS, CF	**0.9651**	0.9104
10	IRC, DDOS, US	**0.9994**	0.9991
11	IRC, DDOS, US	**0.9989**	0.9983
12	P2P	**0.9937**	0.9931
13	SPAM, PS, HTTP	**0.9902**	0.9792

CONCLUSION

- Experiments show quite accurate prediction performance results for SLEEK, a collection of NIDS algorithms
- NetFlow files from other data sources can be implemented with minimal level of effort
- Metrics for pre-processing the feature space are easily configurable
- Network visualization is easily configurable to include animation over time and labeling
- Existing signatures from signature-based NIDS like Snort can be implemented into SLEEK

WAY AHEAD

- Start integration of RF and cyber deception techniques between SEDD and CISD (FY20)
- Incorporate computational algorithms for honeynet allocation (FY20)
- Implement machine learning algorithms for predicting adversaries' preferences for network intrusions (FY20)
- Implement adaptive honeynet configuration algorithm (FY21)
- Develop and implement software-defined networking (SDN) approaches for cyber deception (FY22-23)
- Integrate RF and cyber deception techniques between SEDD and CISD (FY24)

SUMMARY

Main Technical Accomplishment FY20

- Preliminary implementation of location deception demo in Python, CyberVAN, and sdt3d tools
- Developed scalable algorithm for POSG for cyber deception
- Demonstrated that one can learn the adversary's preferences using data from a modest number of deception strategies
- Three books published by IEEE Press

Other Accomplishments in FY20

- Presentations: 20+ presentations
- Conference paper or journal article submissions: 50+ papers
- Engagements with broader S&T community: DARPA DSO SI3-CMD, Cybersecurity CRA, IoBT CRA

Back-up Slides

ALGORITHM FOR IP ADDRESS METRIC

Input: X_{tr}, original training set; and the 2 centroids in $\widehat{X}'$, the modified/normalized training set (without IPs), of the examples labelled as malicious, C^+, and benign, C^-

Output: $\widehat{X^*}$ a modified feature space with 4 additional feature vectors concatenated with $\widehat{X}'$ to represent the IP metric for distances between

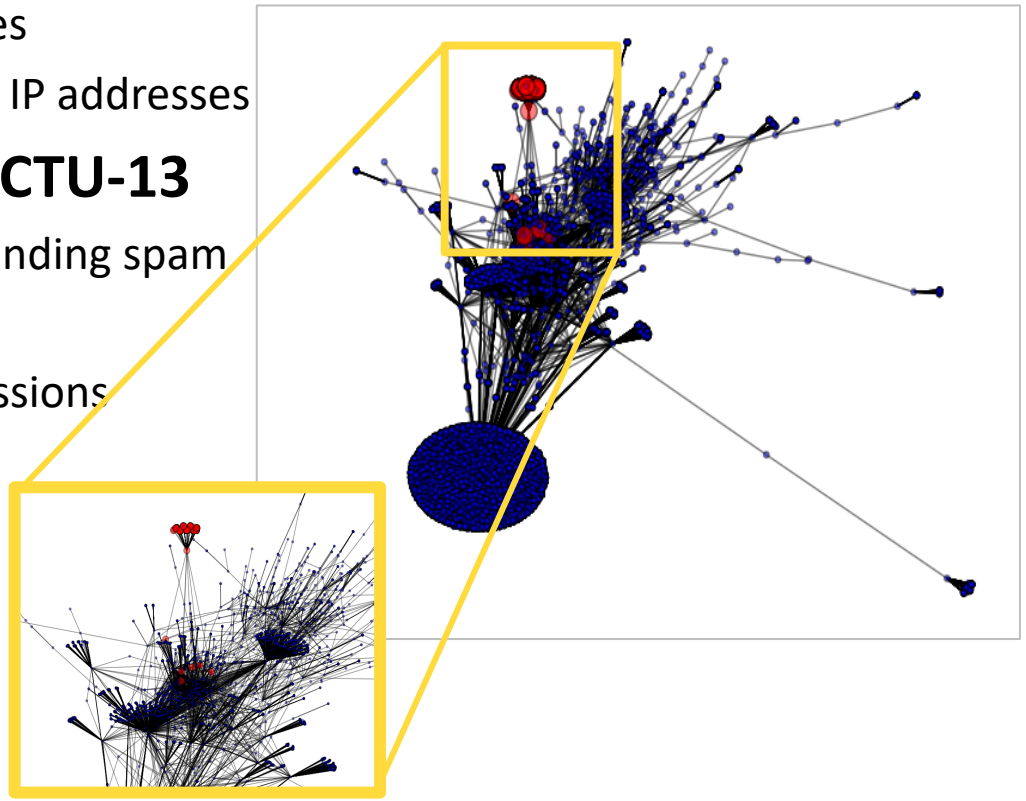
a) Source (destination) IP addresses in training data and

b) Source (destination) IP of the training sample closest to each centroid

1. Find sample in training set, $x \in \widehat{X}'$, with the minimum L2-distance to the malicious centroid, C^+.
2. Repeat Step 1 for C^-
3. Find the associated training example in X that maps to these 2 samples in $\widehat{X}'$ (found in Step 1) that are closest to the 2 centroids
4. Compute the IP metric for distances between feature vectors for source and destination IP addresses in original training set, X, and training samples of the modified training data identified in Step 3
5. Repeat Step 4 for the malicious centroid, C^+
6. Create four additional feature vectors with IP distances computed in Steps 4 and 5 and concatenate with the modified feature space, $\widehat{X}'$
7. Normalize feature space for test data with mean and standard deviation from the modified training data
8. Modify test data with distances between its feature vectors for source and destination IP addresses and the malicious and benign centroids of the normalized/modified training set using

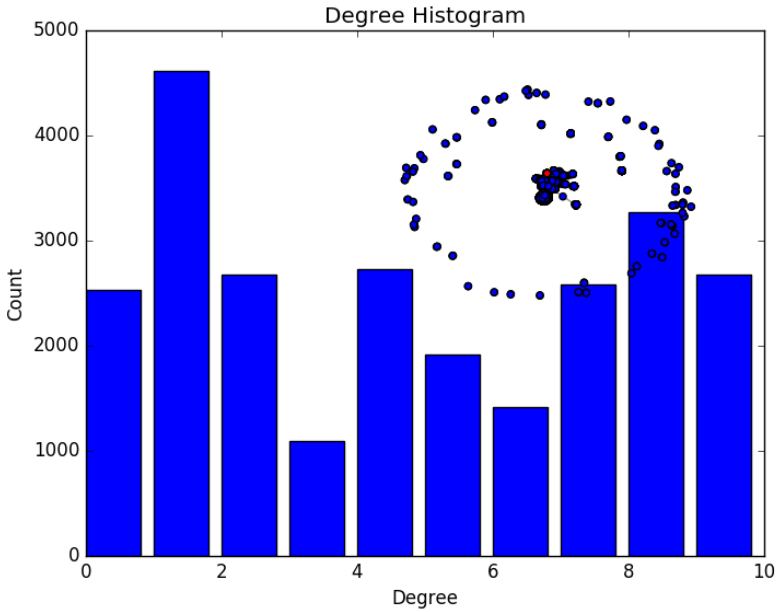
VISUALIZING NETWORK TRAFFIC

- **Developed the SLEEK visualization module**
  - Red nodes are malicious IP addresses
  - Blue nodes are normal IP addresses
  - Links signify connections between IP addresses
- **Scenario 5 Characteristics in CTU-13**
  - Use HTTP for port scanning and sending spam
  - Scan web proxies
  - Botnet flows are 0.54% of total sessions



GRAPH DEGREE HISTOGRAM

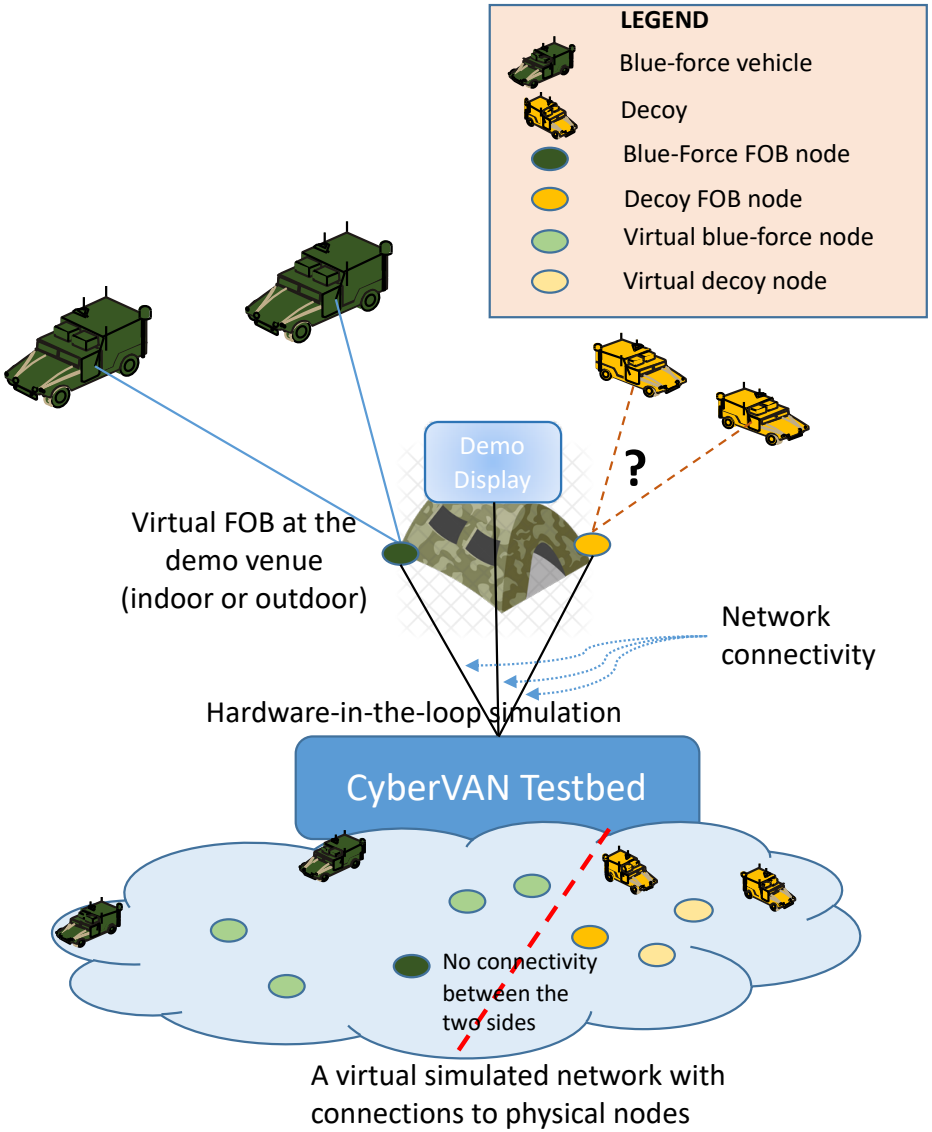
- **Implemented Scenarios in the SLEEK graph analysis module**
- **Present degree histogram with network graph inset**
- **Bi-modal distribution in degree histogram of network sessions data**
- **Scenario 5 Characteristics in CTU-13**
  - Use HTTP for port scanning and sending spam
  - Scan web proxies
  - Botnet flows are 0.54% of total sessions



PLANNING FOR THE OUTDOOR DEMO

Questions:

- **RF communications**
  - What RF links on the blue-force vehicles that the on-board devices will use to communicate with the devices at the FOB?
    - Range, bandwidth, link stability of the RF links to support reachback to the FOB
  - What RF links are to be used on the decoy devices?
    - Other than ISM band push-to-talk, any other RF links on decoys supporting reachback to the FOB such that remote monitoring or cyber deception may be performed?
- **Decoy devices (single board computers?)**
  - If RF links on decoys support reachback to the FOB, is it possible to run software on the decoys?
- **Demo logistics**
  - What are the key points to be demonstrated?



PROACTIVE ADVERSARIAL MODELING HEILMEIER CATECHISM

1. What are you trying to do? Articulate your objectives using absolutely no jargon. What is the problem? Why is it hard?

- Army platforms/dismounts are vulnerable to adversary detection, classification, identification, geolocation, and kinetic/non-kinetic targeting
- Lack cyber deception or decoys
- Limited resiliency in congested CEMA environment

2. How is it done today, and what are the limits of current practice?

- Current cyber deception strategies suffer from the following limitations:
- Mainly enterprise-focused geared towards detecting enterprise attack campaigns;
- Focused on business applications mimicry to deliver technologies that deceive and contain a near-peer or peer adversary; and
- Assumption that attackers act alone and ignore the coalitions among attackers.

3. What's new in your approach and why do you think it will be successful?

- Commercial cyber deception products are not designed to fulfill several important Army objectives, including:
- Mission Resilience
- Resource-constrained environments (e.g., size, weight, power, run time, memory usage)
- Vehicle and tactical networks
- Simple to deploy and maintain. Does not require subject matter experts in the field
- Enable Multi-Domain Operations

4. Who cares?

- TRADOC Pamphlet 525-3-1, The U.S. Army in Multi-Domain Operations 2028 (Dec 2018) identifies deception as being necessary for robust lines of communication.
- Moreover, the publication entitled, Army Support to Military Deception, FM 3-13.4 (February 2019) states,
    - Deception applies to all levels of warfare, across the range of military operations, and is conducted during all phases of military operations.
    - When properly integrated with operations security (OPSEC) and other information-related capabilities (IRCs), deception can be a decisive tool in altering how the enemy views, analyzes, decides, and acts in response to friendly military operation.

PROACTIVE ADVERSARIAL MODELING HEILMEIER CATECHISM(CONT.)

5. If you're successful, what difference will it make? What impact will success have? How will it be measured?

- Developing adaptive cyber deception approaches enhances network resilience and better positions the Army to face adversaries capable of deploying automated cyberattacks
- Adaptive cyber deception provides the Army with a defensive advantage against adversaries with the following solutions
- Hiding mission critical assets through camouflage;
- Misrepresenting a system with obfuscation techniques; and
- Luring the enemy to expend resources on fake nodes, including decoys and honeynets, while real systems remain safe and continue to execute mission critical tasks.
- Using computational algorithms and modeling approaches for adaptive, autonomous deception, the Army will gain a deeper understanding of the adversaries' tactics, techniques, and procedures (TTPs) for network intrusions and reconnaissance
- An example metric is the amount of generation time for a dynamic honeynet from initial adversarial network characterization

6. What are the risks and the payoffs?

Risk	Mitigation
Most graph problems suffer from a combinatorial explosion of the number of states with the network growth	Use of heuristic search algorithms that quickly converge and find optimum policies that are scalable
Centralized solutions are not appropriate for ad hoc tactical networks that are distributed systems	Develop distributed algorithms that converge to the optimum policies by local estimates, sharing with neighboring networked devices, and iterative computation

7. How long will it take? ERP Project end date is FY25.

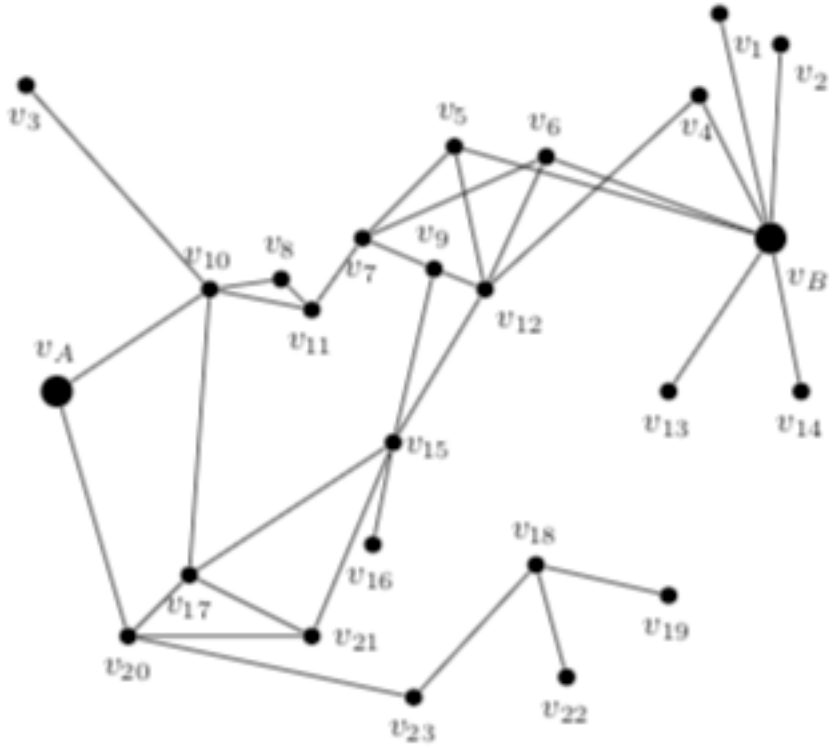
8. What are the mid-term and final "exams" to check for success? How will progress be measured?

- Develop machine learning algorithms to improve predictions of the number or occurrence of network attacks for effectively implementing adaptive, dynamic honeynets (FY20 Q4).
- Assess prediction performance results and sensitivity analysis for machine learning forecasting algorithm implementation (FY21 Q1).

Community Detection Modeling with Machine Learning

IOBT NETWORK AS A GRAPH (1/2)

(a) The IoBT. [4]



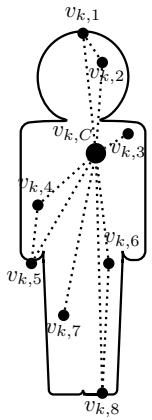
(b) A graph representation of the IoBT.

- **The IoBT, an application of IoT, can be represented as a graph.**

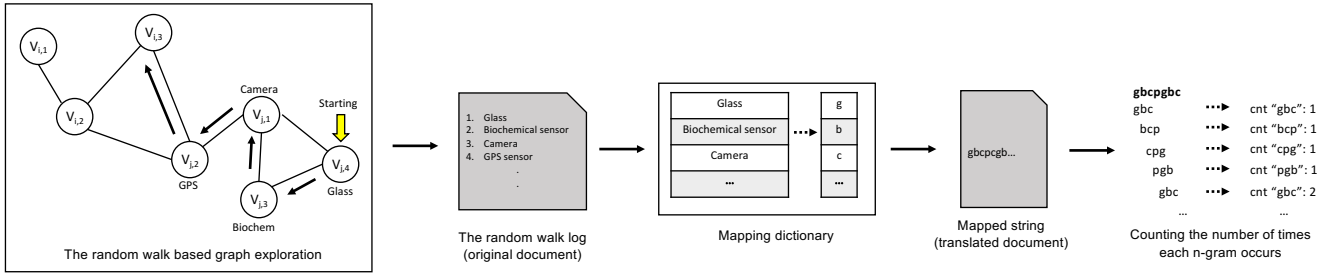
Park, J., Mohaisen, A., Kamhoua, C. A., Weisman, M. J., Leslie, N. O., & Njilla, L. (2019). Cyber Deception in the Internet of Battlefield Things: Techniques, Instances, and Assessments. In *20th World Conference on Information Security Applications (WISA 2019), August 21-24, 2019. Maison, Jeju, Korea*.

FY19 SELECTED RESEARCH ACCOMPLISHMENTS

- **Goal:** Deceive the adversary into misunderstanding the defender's IoBT network activities
- **Accomplishments:**
  - Formulated the IoBT domain as a graph learning problem from an adversarial point of view
  - Introduced various tools through which an adversary can learn the graph starting with partial prior knowledge
  - Developed machine learning algorithms to show that an adversary can learn high-level information from low-level graph structures (i.e., number of soldiers, their proximity, and the number (and type of) assets in the network
  - Developed a powerful n-gram based algorithm to obtain features from random walks on the underlying graph representation of IoBT
  - Provided microscopic & macroscopic approaches that manipulate the underlying IoBT graph structure to introduce uncertainty in the adversary's learning
  - Successfully demonstrated our approach's effectiveness through various analyses and evaluations.



Body Area Network



Adversary's strategy using random walk and n-grams

Park, J., Mohaisen, A., **Kamhoua, C. A., Weisman, M. J., Leslie, N. O.**, & Njilla, L. (2019). Cyber Deception in the Internet of Battlefield Things: Techniques, Instances, and Assessments. In *20th World Conference on Information Security Applications (WISA 2019),* August 21-24, 2019.