The Internet of Things (IoT): Computational Modeling in Congested and Contested Environments

Nandi O. Leslie, Anthony Martone, and Michael Weisman

U.S. Army Research Laboratory, Adelphi Laboratory Center

2800 Powder Mill Rd

Adelphi, MD 20783, USA

Email: Nandi.O.Leslie.Ctr@mail.mil

Introduction

The breadth and magnitude of the challenges facing computation for human-cyber-physical systems—examples include Internet of Things (IoT) and Internet of Battle Things (IoBT) (Kott et al., 2016)—is staggering. According to CISCO, there will be over 500 billion entities in an IoT environment by 2030 (CISCO, 2016). Many of these entities were previously unidentified as requiring networked communications (Trappe et al., 2015), and network security and system resilience concepts have often been excluded from their system design (Kott et al., 2016). On the battlefield, future warfighter operations and missions will rely more heavily on networked devices designed with autonomous cognitive decision-making capabilities to perform a broad range of tasks, including cognitive sensing, communicating with human warfighters, conducting operations in congested environments, and securely processing and communicating data to other autonomous agents (Adler et al., 2016; Kott et al., 2016). Furthermore, the modeling challenges related to autonomous cyber defense of IoT and IoBT[1] in contested and congested environments are complex. That is, the confidentiality, integrity and availability of IoT data combined with proper device functioning—this process is a key aspect of cyber resilience—can each be manipulated by a skilled adversary. Specifically, intrusion detection plays a key role IoT security and resilience: enhancing cyber systems for IoT in regularly congested environments so that the systems autonomously switch functionalities is a rich research area. Detecting cyberattacks that leverage novel methods to exploit previously unidentified (i.e., "zero-day" attacks) is infeasible with signature-based detection models for intrusion prevention or detection systems (IPS/IDS).

Here, we narrow our focus to some of the modeling challenges related resilient wireless communications for connected the radio frequency (RF) sensors in contested and congested environments. In this paper, we propose a framework for an autonomous cyber defense agent (ACDA) to enhance intrusion detection for IoT which, in addition to cybersecurity quantification, directly impacts cyber-risk and resilience assessments. This notional agent acts in a mediatory role between the entities in the IoT and some host: monitoring the network traffic with distributed or centralized network controls, preventing cyberattacks with IPS such as firewalls, and detecting anomalous and malicious traffic with anomaly-based, semi-supervised, and unsupervised learning algorithms. For example, the Routing Protocol for Low-Power and Lossy Networks (RPL) is a standard routing protocol for the IoT, and it is known that wireless sensor networks (WSN) using IPv6 over Low-powered Wireless Personal Area Networks (6LoWPAN)—entities, such as cognitive RF sensors—are vulnerable to cyberattacks originating on the Internet and from within the network (Raza et al., 2013). Motivated by Raza et al. (2013), we define a notional framework for an ACDA (see Figure 1) that places an IDS in the router and lightweight IDS modules (Yu and Leslie, 2017) on the entities in the network. Using a semi-supervised learning model, we show that an IDS can effectively detect botnet traffic for the datasets considered that the algorithm has not previously detected. This is critical for our understanding of the applicability of anomaly-based detection algorithms in IDS deployed in an IoBT environment.

[1] For the remainder of this paper, we will use the acronym IoT, if IoT, IoBT, or industrial IoT can be used interchangeably.

IoBT in Contested Environments: Cybersecurity Modeling

In a contested environment, IoT security requires successful autonomous cyber defense which is in part characterized by detecting cyberattacks and protecting against them. Similarly, network intrusion detection systems (NIDS) and related algorithms, of necessity, must be fast, of low computational complexity, and have some degree of autonomy with advanced cognition, where decision-making normally reserved for human analysts is automatically performed by IoT NIDS (Kott et al., 2016; Shearer et al., 2017; Yu and Leslie, 2017). One of the simplest unsupervised learning algorithms is the k-means algorithm (Lloyd, 1982)—it has several successful applications to anomaly detection models in NIDS for large-scale enterprise networks (Wang & Paschalidis, 2015)—and machine learning models with similar properties will be critical to IoT and IoBT operations in adversarial environments (Chang et al., 2013; Raza et al., 2013; Kott et al., 2016). The k-means algorithm is a clustering algorithm characterized by two iterative steps that produce a Voronoi tessellation of the feature space: (i) assign training examples to closest centroid; and (ii) (re)compute the mean of each centroid which is the mean of the training examples assigned to it. Because the k-means algorithm is very simple and many of its implementation challenges are shared among other unsupervised learning algorithms, we use k-means as a representative for anomaly-based detection models in NIDS to stage our discussion of IoT security modeling, namely, whether the challenges of cyber security and resilience modeling for large-scale networks have parallels in IoT and IoBT. For example, anomaly-based detection models tend to have high false positive rates (often misclassifying normal activity as an attack), and signature-based NIDS tend to have high false negative rates (misclassifying actual attacks).

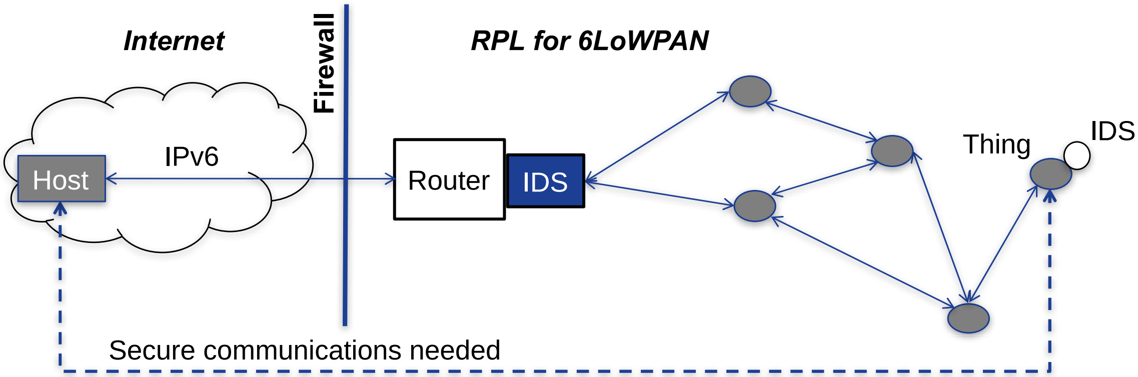

Figure 1. Notional ACDA framework.

We further explore these modeling challenges with the Czech Technical University (CTU)-13 botnet packet capture (pcap) scenarios which provide a labeled, real dataset with background, normal, and malicious traffic, where each of the thirteen pcap scenarios is characterized by a unique type of botnet (Garcia et al., 2014). In this study, we focus on three of the thirteen CTU-13 botnet scenarios (see Table 1 for some details about CTU-13 scenarios), where each of the scenarios is a different type of distributed denial of service (DDoS) attack: IDs 4, 10, and 11 (Gu et al., 2008).

Table 1. Characteristics of three of the CTU-13 botnet pcap scenarios: IDs 4, 10, and 11 are different types of DDoS attacks (Gu et al., 2008).

	ID 4	**ID 10**	**ID 11**
Brief description of DDoS type	UDP & ICMP	UDP	ICMP
Duration (hrs.)	4.21	4.75	0.26
Number of flows	4,238,045	5,180,852	40,836
Number of bots	1	10	3

We implement a semi-supervised learning algorithm for a NIDS based on the k-means algorithm, where we use the labels from the training set only in the centroid initialization phase of the algorithm—this centroid initialization phase occurs before the two main iterative steps of this nearest neighbor model. In addition, prior to model implementation, we pre-process the IP flow data and characterize each example by the following nine categorical and quantitative features: duration (in hours), protocols (e.g., TCP, UDP), server IP address, server port, client IP address, client port, total number of packets, total bytes, and number of client bytes. Additional pre-processing steps are converting categorical features to quantitative features with one-hot encoding (Garcia et al., 2014), and then standardizing the feature space. We follow by performing principal component analysis (PCA) for dimension reduction of the feature space such that 99% of the variance is explained by the principal components resulting in as much as a 20-dimensional feature space for CTU-13 Scenario ID 11. Using 4-fold cross validation, we present the k-means prediction performance results for the testing dataset in Table 2, where $k = 2$ for the normal and malicious clusters.

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

Although our semi-supervised modeling approach for an ACDA is very simple, we are able to detect cyber intrusions with success (see Table 2 for prediction performance results), where FPR is 0.0 for scenario ID 4 and accuracy is above 0.97 for each botnet scenario considered. Nonetheless, modeling challenges exist even for detecting these botnets with DDoS attack which are masked by high accuracy results and low FPR (see Table 2). Low recall results can be detrimental to IoBT mission success for botnet scenario ID 4 (see Table 2) indicating a high number of false negatives. NIDS prediction performance must be prioritized to match IoBT mission needs. In addition, IoBT security must operate efficiently to secure mobile ad hoc networks (MANETs) and ensure that the computational resource utilization constraints are met (Chang et al., 2013; Leslie et al., 2017 b).

IoBT in Congested Environments: Cognitive Sensor Spectrum Sharing

IoBT security and resilience challenges are multi-faceted—its sensors and communications are interactive, adaptive, dynamically-configured, and goal-driven. These challenges include effectively and efficiently operating in a congested environment to develop situational awareness by collecting and refining data (Kott et al., 2016). The technology needed to develop situational awareness for sensors is an ongoing research challenge that has been given much attention from the RF sensor community in the application areas of cognitive radio and cognitive radar (Martone, 2014). Fueled by the ever-growing wireless communication industry and its need for more frequency bandwidth, regulatory institutions [such as the Federal Communications Commission (FCC)] are motivated to explore new spectrum access technologies (FCC, 2013; FCC, 2015). These technologies would allow radio and communications systems to effectively share the frequency spectrum and mitigate mutual RF interference. This technology is necessary for "blue force" radar and communication systems to coexist within the IoBT.

An enabling radar technology that could be leveraged for communication system spectrum sharing is the spectrum sensing, multi-objective optimization (SS-MO) technique (Martone et al., 2015 a). SS-MO is a bandwidth sharing approach, where the radar attempts to identify a sub-band (or channel), within an overall frequency band of interest, by passively sensing the electromagnetic spectrum. The goal of SS-MO is to simultaneously maximize radar performance while mitigating mutual interference. To maximize performance SS-MO uses multi-objective optimization to jointly maximize bandwidth and signal to interference plus noise (SINR) ratio. A maximum radar bandwidth is required to preserve radar range resolution and resolve closely spaced targets. SS-MO has been shown to significantly increase SINR (Martone et al., 2015 a; Martone et al., 2016), mitigate range sidelobes (Martone et al., 2015 b), and effectively share the frequency spectrum with communication systems (Martone et al., 2017).

Conclusion

Central to the task of an ACDA is defending mission-critical IoBT assets operating in a contested and/or congested environment so that they are secure and resilient—that is, able to recover from attacks to the physical and cyber environments. Furthermore, to promote IoT security and situational understanding, the risks of cyber-attacks must be assessed with predictive computational modeling (Leslie et al., 2017 b). To communicate and collaborate on the battlefield and ensure cyber situational awareness and understanding for IoBT entities (e.g., cognitive RF sensors), enabling technologies and cognitive techniques (such as SS-MO), provide spectrum access for RF sensors and optimized performance in congested electromagnetic environments. Researchers will need to continue to fuse these existing technologies with intrusion detection for cyber-risk, security, and resilience modeling for IoBT.

References

Adler, E., Dietlein, C., Hedden, A., Martone, A., Mitchell, G., Zaghloul, A., & Govoni, M. A. (2016, May). Trends in radar: a US Army Research Laboratory perspective. In *SPIE Defense+ Security* (pp. 98290U-98290U). International Society for Optics and Photonics.

Chang, R. J., Harang, R. E., & Payer, G. S. (2013). *Extremely lightweight intrusion detection (ELIDe)* (No. ARL-CR-0730).

CISCO (2016, June). *The Internet of things* [web blog post]. Retrieved from https://www.cisco.com/c/dam/en/us/products/collateral/se/internet-of-things/at-a-glance-c45-731471.pdf.

Eskin, E., Arnold, A., Prerau, M., Portnoy, L., & Stolfo, S. (2002). A geometric framework for unsupervised anomaly detection: Detecting intrusions in unlabeled data. *Applications of data mining in computer security*, 6, 77-102.

Federal Communication Commission. (2003, December). *Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies* (No. FCC 03-322). Washington, DC: FCC.

Federal Communication Commission. (2015, April). *Amendment of the commission's rules with regard to commercial operations in the 3550-3650 MHz band* (No. FCC 15-47). Washington, DC: FCC.

Garcia, S., Grill, M., Stiborek, J., & Zunino, A. (2014). An empirical comparison of botnet detection methods. *Computers & security, 45*, 100-123.

Gu, G., Perdisci, R., Zhang, J., & Lee, W. (2008, July). BotMiner: Clustering Analysis of Network Traffic for Protocol-and Structure-Independent Botnet Detection. In *USENIX security symposium* (Vol. 5, No. 2, pp. 139-154).

Kott, A., Swami, A., & West, B. J. (2016). The Internet of Battle Things. *Computer*, *49*(12), 70-75.

Leslie, N. O., Harang, R. E., Knachel, L. P., & Kott, A. (2017 a). Statistical models for the number of successful cyber intrusions. *The Journal of Defense Modeling and Simulation: Applications, Methodology, and Technology,* DOI: 10.1177/1548512917715342.

Leslie, N. O., Marvel, L. M., Edwards, J., Comroe, K., Shearer, G., & Knachel, L. (2017 b). Modeling approaches for intrusion detection and prevention system return on investment. In *SPIE Defense+ Security* (pp. 1018502-1018502). International Society for Optics and Photonics.

Lloyd, S. (1982). Least squares quantization in PCM. *IEEE transactions on information theory*, 28(2), 129-137.

Martone, A. F. (2014). Cognitive radar demystified. *URSI Bulletin*, (350), 10-22.

Martone, A. F., Gallagher, K. A., Sherbondy, K. D., Ranney, K. I., Dogaru, T. V., Mazzaro, G. J., & Narayanan, R. M. (2015 a). Adaptable bandwidth for harmonic step-frequency radar. *International Journal of Antennas and Propagation*.

Martone, A.F., Sherbondy, K.D., Ranney, K.I., Dogaru, T.V. (2015 b). Passive sensing for adaptable radar bandwidth. *Proc. of the 2015 IEEE Int. Radar Conf., Arlington, Va, May 2015*, pp. 280 - 285.

Martone, A., Ranney, K., & Sherbondy, K. (2016, May). Genetic algorithm for adaptable radar bandwidth. In *Radar Conference (RadarConf), 2016 IEEE* (pp. 1-6). IEEE.

Martone, A. F., Gallagher, K. A., Sherbondy, A., Hedden, C., Dietlein, (2017, in press). Adaptable waveform design for enhanced detection of moving targets. *IET Radar, Sonar & Navigation Journal*.

Raza, S., Wallgren, L., & Voigt, T. (2013). SVELTE: Real-time intrusion detection in the Internet of Things. Ad hoc networks, 11(8), 2661-2674.

Shearer, G., Leslie, N.O., Ritchey, P., Braun, T., & Nelson, F. (2017, October). IDS Alert Prioritization through Supervised Learning. In *Proceedings of the NATO Specialists' Meeting on Predictive Analytics and Analysis in the Cyber Domain, 10-11 October 2017, Sibiu, Romania*. NATO.

Trappe, W., Howard, R., & Moore, R. S. (2015). Low-energy security: Limits and opportunities in the internet of things. *IEEE Security & Privacy*, 13(1), 14-21.

Wang, J., & Paschalidis, I. C. (2015). Statistical traffic anomaly detection in time-varying communication networks. *IEEE Transactions on Control of Network Systems*, 2(2), 100-111.

Yu, K., & Leslie, N.O. (2017, October). FAST-D: Malware and Intrusion Detection for Mobile Ad Hoc Networks (MANETs). In *NATO Specialist Meeting IST-145 on Predictive Analytics and Analysis in the Cyber Domain*. NATO.