Computational Intelligence-enabled Cybersecurity for the Internet of Things

Shanshan Zhao, Shancang Li, Senior Member, IEEE, Lianyong Qi, and Li Xu, Fellow, IEEE

Abstract—The computational intelligence (CI) based technologies play key roles in campaigning cybersecurity challenges in complex systems such as the Internet of Things (IoT), cyber-physical-systems (CPS), etc. The current IoT is facing increasingly security issues, such as vulnerabilities of IoT systems, malware detection, data security concerns, personal and public physical safety risk, privacy issues, data storage management following the exponential growth of IoT devices. This work aims at investigating the applicability of computational intelligence techniques in cybersecurity for IoT, including CI-enabled cybersecurity and privacy solutions, cyber defense technologies, intrusion detection techniques, and data security in IoT. This paper also attempts to provide new research directions and trends for the increasingly IoT security issues using computational intelligence technologies.

Index Terms—Computational Intelligence, Cybersecurity, Digital Forensics, Internet of Things (IoT).

I. INTRODUCTION

The computational intelligence (CI) techniques can enable smart devices to learn a specific task from data or behaviours, which covers evolving computation, neural networks, fuzzy logic, learning theory, probabilistic, and similar computational models [1], [2], [3]. The CI techniques promise considerable benefits for the Internet of Things (IoT) that connects billion of smart devices. The CI exponentially enlarged the size of IoT market and reach up to £217,832.19 million by the end of 2020 [4]. The combination of CI and IoT can create new markets and opportunities and they are highly unlikely to lose ground in the nearest future. However, the increasing IoT also leads increased security concerns and challenges. The CI techniques are promising to help IoT systems to detect security patterns from data and learn to adjust their behavior to avoid potential cyberthreats [5].

Typically, cybersecurity analysis in IoT mainly focus on two kinds of approaches [6]: (1) analysis driven, which relies on rules determined by security and policies; and (2) artificial intelligence (AI) and machine learning (ML), which relies on AI techniques such as un/supervised machine learning techniques to detect rare or anomalous patterns. In complex IoT systems, we expect that the security analysis can be conducted in automated and timely manner, which requires intelligent learning techniques to be developed to assist IoT to analyse massive data/events/behaviours and identify/response cyberthreats in an efficient way. In the past few years, CI techniques are adopted to improve cybersecurity, including learning to detect suspicious behaviours, stopping cyberattacks, etc.; However, the cybercriminals are also adopting the advantages of CI to exploit vulnerabilities and conduct cyberattacks, for example, DeedLocker is a malware assisted by CI techniques that can automatically identify, analyze and steal data from social media [7].

Similarly to the AI and ML, the CI is a new automated information processing technology with minimal human intervention to solve complicated problems, which have attracted huge of research efforts due to two facts: (1) The CI-based algorithms are computationally intensive, the advances in new techniques, such as GPU, big data, cloud computing, etc., have significantly boost the development of CI algorithms; (2) The availability of huge amount of data make it possible to train fine models, which makes the data analytic platform available and significantly improved the effectiveness of CI. The advances in computing and data make CI better in many applications than human. In the heterogeneous 5G-IoT environment, the increasing number of IoT devices increases vulnerability for various spoofing attack. Existing security techniques (such as authentication, encryption, etc.) are facing many challenges in such a complex dynamic environment, including significant security overhead, low reliability, as well as difficulty in pre-designing authentication model, providing continuous protections, and learning time-varying attributes. The CI is promising in solving the cybersecurity in IoT, in which if a device can be trained to find new malware and anomalies, however it requires: (1) access to large volume of data; more malware samples can train ‘smarter’ model; (2) CI-based data processing techniques that enhance the performance of data analysis; (3) security expertise to monitor CI model can accurately identify the real threats and provide insights.

The main security challenges in the IoT includes: 1) Insufficient testing and updating, it is a very difficult task for IoT systems to offer firmware updates and enough support for legacy linux kernels. This makes cyber vulnerabilities exposed to cyberattacks in IoT. For example, the Mirai malware can conduct attacks using this kind of vulnerabilities. The IoT device vendors must proper test their products before launching into the public, and the IoT devices need to be updated regularly. 2) IoT malware and ransomware will continue to rise in the following years, and the security solutions should

Dr. Shanshan Zhao is with School of Software, Shanxi Agricultural University, China; Dr. Zhao also is with the Department of Engineering, Design and Mathematics, University of the West of England, Bristol BS16 1QY, UK. Email: Shanshan.Zhao@uwe.ac.uk.

Dr. Shancang Li is with the Department of Computer Science and Creative Technologies, University of the West of England, Bristol BS16 1QY, UK. Email: Shancang.Li@uwe.ac.uk.

Prof. Lianyong Qi is with School of Information Science and Engineering, Qufu Normal University, Rizhao, China. Email: lianyongqi@gmail.com.

Prof. Li Xu is with Department of Information Technology, Old Dominion University, USA. Email: lxu@odu.edu.
be respond to in a timely and automated manner; 3) IoT botnets aiming at cryptocurrency, with the emerging blockchain technologies; 4) Data security and privacy.

There is a lack of clear definition computational intelligence, which is the intelligence of a smart devices that capable of performing computations and can perform any intellectual task with minimal manually intervention. Typically, the CI contains following five main principles, fuzzy logic, neural networks, evolutionary computation, learning theory, and probabilistic methods. From the viewpoint of security, the CI covers following topics: cybersecurity issues, cyber defense technologies, new security challenges and CI solutions, and intrusion detection approaches using main CI methods.

II. BACKGROUND AND RELATED WORKS

In IoT environment, the IoT devices (e.g. sensors) are typically limited with resources, in terms of computation, storage, memory, power etc.; hence, new technologies that can match the needs for resource-constrained devices in IoT are necessary. Considerable research efforts on cyberthreat intelligence have been conducted in the past few years and a number of sophisticated techniques have been developed that can perform cybersecurity anomaly detection. In IoT environment, the CI enabled cybersecurity solutions need to be more flexible and efficient.

In broadly, computational intelligence algorithms have been used in IoT security solutions, i.e., malware detection, cyberthreats identification, suspicious behavior monitor, intrusion detection, stopping cyberattackers, etc. The CI techniques can enable the IoT upgrade its cybersecurity capabilities and protect IoT applications and users.

An IoT systems involves large number of devices, services, applications, and users, it may face a large number of ordinary and innocent cyber attempted attacks everyday, such as customers mis-entering password, etc., which need automated systems to filter out and truly dangerous signal from the more-easily-addressed noise. The CI can offer operational deficiencies and potential operating expense savings.

In smart home, the CI enabled techniques also brings threats to cybersecurity: it is reported that the CI techniques are used to develop techniques for unlocking doors and transferring money using devices such as Alexia, Siri, and Google Assistant in smart home without the knowledge of the smart home users. It can image that the CI techniques, such as financial sectors, pricing algorithms, smart environments, can be used by attackers targeting on IoT applications.

It is clear that the cyberthreats and attacks grow in volume and complexity, even worse the CI could be used by attackers to develop more powerful attack tools, including malware, ransomwares, CI-enabled attack kits, and more [8]. In industry, CI can help make cyberthreats detection quicker and make IoT systems stay ahead of threats. The using of CI techniques, it is possible to provide deeper security and simplify the process for security analysis. However, the cybercriminals can also use CI techniques to develop new threats that might be more difficult to identify. In this case, the organizations need to well design the security strategies and use data-centric security models [9]. However, cybercriminals are also using more intelligent tools to commit cybercrimes. It is possible that the CI are used to make IoT malware turn into a weapon. Dilek et al. reviewed the CI based techniques in cybercrime [10], including the vulnerabilities and threats identification in intrusion detection systems using the most recent CI techniques. In the IoT environment, the existing human intervention based methods againsting cyberattacks are not enough, and intelligent and automated techniques combating cyberattacks, malwares, ransomwares, has become a requirement.

In [11], Ivkic et al. analysed the cybersecurity standards and quality management methods in smart IoT and business. Alansari et al. believed that the CI plays an essential role in the interpretation of big data in bio-informatics, such as DNS sequence analysis, medical big data, etc. The CI-based techniques can be used in complicated and computational expensive data analysis [12]. Similarly, in IIoT, Rehman et al. proposed a big data processing framework for applying CI at IoT device-end, where a large amount of resource-constrained smart sensors are included [13]. For IIoT data analysis, such as classification, categorization, etc., the CI can be used to perform accurate and operational and customer level indulgences, such as data sources, analytics tools, analytic techniques, requirements, industry analytic applications, etc.

In [14], Donno et al. reviewed security issues in cloud-based IoT environment from three different levels: physical (device), vialization, and application level. In their works, the cloud-based IoT security issues were addressed from two different scenarios: cloud-specific and cloud-generic.

Actually, the CI-based techniques are widely used in intrusion detection and prevention of cybersecurity in IoT. Table I summaries the CI techniques in cybersecurity intrusion detection and prevention. It is noted that in recent, evolutionary computation algorithms are widely used in intrusion detection and attack defense, learning algorithms are mainly used in structural security feature extraction.

In Section III, we will address how the CI-enabled techniques can be used in enforcing the security of IoT.

III. CI-ENABLED CYBERSECURITY FOR IOT

To address new features in IoT cybersecurity, we must consider that the CI-enabled security solutions should be designed to be able to handle security issues in IoT environment: security architecture, threats and incident analysis, security incident management and response, intrusion detection, malware analysis, data security, and more.

Figure 1 presents an IoT architecture for CI-enabled cybersecurity analysis, in which the bottom block shows cybersecurity concerns in IoT, and the top block addresses the CI algorithms, the architecture applies CI algorithms over IoT security concerns for security solution in IoT.

Figure 2 shows the details of CI-enabled IoT Cybersecurity, in which the CI techniques could be one or more computational models and tools that encompass elements of learning, adaption, and/or heuristic optimization that can help to solve cybersecurity problems that are difficult to solve using conventional computational algorithms.
The CI-enabled techniques can enhance existing cybersecurity systems and practices from following three levels:

- **Prevention and Protection.** It can be seen that traditional security solutions, such as IDS, struggle to keep up with security threats. The new cyberthreats such as fileless attacks, powershell attacks, are increasing emerging, which are stealthy and very difficult to detect. The CI-enabled security solutions can identify and detect malicious piece of intelligent code and get out of the IoT systems in a timely and automated way. Meanwhile, the CI-enabled techniques can be adapted to cyberdefence tools that can learn from labelled data, emerging threats, and create security strategies. The CI-enabled cybersecurity solutions can match the urgent needs to protect data and high-value assets, systems in IoT systems.

- **Threats identify and detection.** A number of CI (AI) assisted cyberthreats detection techniques have been developed, which use advanced machine learning algorithms (such as unsupervised learning algorithm, deep learning) to identify threat patterns in datasets. These patterns can be used to spot anomalies in cyber behaviours. CI algorithms can also be used in threat detection, such as deep learning can be applied to insider threats that difficult to spot. Some self-training CI algorithms (such as deep learning) can be used to labelled to identify insider threats that human being unable to differentiate from normal behaviour. The main security solution vendors, such as DarkTrace, Cylance, Symantec, etc. are using CI techniques to provide CI-based cyberthreats detection/prevention solutions.

- **Response.** The IoT systems generate huge amount of data to go through, the CI techniques can detect threats quickly and determine an accurate response. Using CI and ML algorithms, CI-enabled cybersecurity tools can conduct searching, security analysis, threats detection in an automated manner. By doing this, CI can also significantly reduce the intervention of human being. Since the CI enable these tools to learn from potential threats and update response strategy in a timely manner.

It is very difficult to implement an CI-driven IoT, which requires careful security design and strategic planning, specifically for critical industries, such as healthcare, public security,
smart cities, emergency settings, etc.

IV. KEY ENABLING TECHNOLOGIES IN CI-ENABLED CYBERSECURITY

As mentioned above, CI is a group of computational models and tools that encompass elements of learning, adaption, and/or heuristic optimization [1]. In this section, we will go through the CI-enabled key cybersecurity technologies in IoT environment.

A. Algorithm

In cybersecurity analysis, CI-based algorithms include biologically inspired algorithms, artificial immune systems, image processing, data mining, natural language processing, reasoning and decision making algorithms. These algorithms are being used to develop cybersecurity applications. Since the security in IoT are expected to be proactive, detecting and preceding CI algorithms are in high demand. These CI algorithms, including data aggregation algorithms, CI learning algorithms, CNN for mobile vision applications, can help to improve in detecting cyberattacks, malware, cyberthreats, etc. in IoT. Figure 3 shows a CI-driven predictive cybersecurity framework, which solve security problems using CI algorithms to create IoT security solutions, and the feedback from IoT security indicators is iteratively updated in CI algorithms to further improve the security solutions.

1) CI Algorithm Packages: Many CI learning packages have been designed to meet specific applications, including the TensorFlow, Caffe, MXNet, and Pytorch. However, these packages are not suitable for IoT devices. Open problems include: First, to execute real-time tasks on the IoT device, many packages scarifies memory to reduce latency; however, memory on the IoT devices is also limited thus how to trade-off the latency and memory is a challenge; Second, having access to personalized, training on the IoT device is ideal while the training process usually requires huge computing resource. Therefore, how to implement a local training process with limited resources is another challenge; Third, the IoT device will need to handle multiple tasks which raises the problem of how to execute multiple tasks on a packages in the meantime need to be addressed.

2) Learning algorithms in IoT Applications: CI learning techniques have been widely used in the IoT applications, including CI-enabled biometric authentication, healthcare, intrusion detection, IoT data analysis platform, etc.

B. Devices and User Access control

To protect IoT users from cyberthreats, we need all content that users access is monitored and updated in time. It needs the system be able to intercept SSL-encrypted communications between IoT devices. In general, the Sandboxing is used to analyse such zero-day threats, which executes suspicious apps in a virtual machine sandbox by mimicking the status and then decide if it is safe or not.

C. CI-enabled Malware and Threats Detection for IoT

It is reported that machine learning based techniques can detect 85% of cyber-attacks using active learning in [6]. CI-enabled malware detection and cyberthreats prevention, using CI techniques, an IoT system can detect malware by providing detailed attributes or behaviors to analyse false positives/negatives and improve model prediction accuracy. In IoT environment, it is very difficult to conduct malware classification without the use of CI-based techniques, which can offer a solution to the diminishing feasibility of data classification without automation. By accessing patterns within the data, CI classifiers are able to effectively label data as malicious, thus increasing the level of security and the ability of administrators to more effectively monitor the complex IoT systems. The classification can extremely beneficial to the area of malware detection. Since the model relies on a pool of training data in order to shape its predictions, this data serves
as a target for potential adversaries who may targeting the classifier’s underlying algorithm. Mathematically, given viable training data, CI classifiers are able to predict classification labels with a high degree of accuracy.

Figure 4 shows an example of a CI-enabled Intrusion detection system. It can be seen that the it uses unsupervised learning algorithms extract pattern signatures, which can be used in intrusion, anomaly, and misuse detection. Security alarm will be report based on the detection results. Meanwhile, the feedback from detection results will be return to the security feature extraction stage for further improve the pattern signature.

D. CI-enabled Anomaly Detection

Anomaly detection is a key enabling technique in IoT security. Similar to intrusion detection system, an anomaly detection mainly aims to offer following key security functionalities: monitoring, detecting, analyzing, and responding to unauthorized traffic. Since IoT covers a huge number of smart devices running rich operating systems and have a massive number of security solutions. However, a large volume of resource-constrained IoT devices are unable to run computational expensive security solutions, which makes it difficult for IoT devices to protect themselves and users against cyberattacks. The CI techniques are expected to improve the security together with lightweight cryptography. In IoT, each devices are generating or collecting a large volume of data, the CI techniques can be employed to analyse the security issues, device behaviours, and usage patterns, data traffic signatures, to help to spot and block abnormal activity and potentially vulnerabilities and threats in IoT. Figure 5 shows an example of anomaly detection using CI packages.

Basically, CI based behavioral analysis is one of the biggest trends in detecting abnormal activities in IoT. The basic idea is to aggregate into an IoT cloud server data from all IoT devices, and then analyze to determine patterns and spot malicious behavior. CI techniques can learn features from these activities and pick up on the abnormal traffic. CI techniques are very promising but it is still in its infancy and has a long way to go.

E. Applications of CI techniques to combating cybercrimes

As discussed above, the CI enables us to design automated cybersecurity solutions. CI techniques are also very promising on assessing security risks and developing security measures for cybersecurity strategies in IoT: (1) CI techniques can discover the essence of intelligence in the huge amount of data created in IoT and develop intelligent analysis systems; (2) CI can find security features modelling methods for solving complex problems that cannot be solved by existing methods. The IoT application of CI to cybersecurity covers following key areas:

- Smart home analysis;
- Big data analysis;
- E-Discovery;
- Smart and Connected Health.

1) CI-enabled Forensics: It is noted that in large-scale IoT systems, increasingly large amount of data are generated and needs to be analysed. Since in such widely distributed systems, the cyber security analysis/investigation, human capacity is simply not possible anymore. The forensic investigation will be conducted over IoT devices, mobile devices, computers, cloud, etc., which makes it very impossible to use traditional forensic investigation in a timely manner. The CI can enable forensic investigation identify artefacts, analyse them, and provide evidence in order to reconstruct what has happened. The forensics is a purely fact-based area, however, CI can enable. For example, pattern recognition, learning algorithms, classification methods. The CI will be very useful for social systems related forensic investigation. The CI can make DF tools trustworthy and efficient. E-Discovery, etc.
V. RESEARCH CHALLENGES AND FUTURE TRENDS

As discussed above, CI can help IoT combat cyberattacks growing in volume and complexity. In the IoT context, many research efforts being conducted on how CI helps solve the cybersecurity challenges. However, there are still many challenges need to be addressed. This section summarises the research challenges and future research trends.

A. Research Challenges

CI algorithms are increasingly being used in a combined approach with related technologies, ranging from advanced analytics and IoT to robotics, edge computing, and more. The main technical challenges are summarised as following.

1) CI-enabled Cybersecurity Architecture: The architecture of CI-enabled cybersecurity is one of the most significant challenges. The IoT combines a number of technologies, such as 5G-IoT, cybersecurity, and CI, which makes it very challenging to applying CI as a “whole” to solve tackles how CI helps solve security challenges.

2) CI Algorithms and Tools: A noticeable shortage of accessible CI algorithms, tools, methods and teaching materials for incorporating verification into cybersecurity solutions.

3) CI-Enabled Data Mining in Cybersecurity: A properly trained cybersecurity CI model can be an important addition to the IoT security solutions. In recent, a number of AI/ML techniques have already been exploited for preventing all cyberthreats. Security visualization in IoT is a key challenge, which consists following main topics:
   - Security labelling of large data volume;
   - Efficient data analysis tools;
   - Keywords extraction;
   - Information coordination, requirement sharing, between different entities in IoT;
   - Correlating the breaches/events.

4) Cognitive security with IoT device: By combining the CI and human intelligence, CI algorithms, including machine-learning algorithms, deep-learning networks, can enable IoT devices more secure and smarter over time. The cognitive IoT devices can learn with each interaction between threats and provide actionable insights. By doing this, the security system over IoT can respond to the threats with greater confidence and quicker. However, there a few challenges need to be addressed: 1) One major challenge is that CI is a broad concept, to select right CI techniques is difficult; 2) How it interacts with other transformational technologies.

5) Efficient CI Algorithms in Cybersecurity: CI algorithms have been widely used in user activity recognition, cyberthreats detection and preventing, cyber anomaly detection, etc. The cybersecurity intelligence gather from complex IoT environment needs efficient and timely CI algorithms. Since in
the cybersecurity analysis, a huge volume of unlabelled data makes it very difficult to train learning models using CI algorithms. For other data analysis, including cyber vulnerabilities exploitation, cyberthreats and attacks detection and prevention, cybersecurity defense, efficient CI algorithms are needed.

6) CI-enabled Malware Detection and Classification: A number of CI-enabled malware detection and classification solutions have been developed in the past few years. To detect malware with learning classifiers in IoT environment still a key challenge due to the evolving and diversity of malware in IoT systems. A good thing is CI based pattern techniques and classifiers are remarkably improved which benefits the development of efficient malware/intrusion detection, classification, and prevention in cybersecurity solutions.

7) CI vs GDPR: The General Data Protection Regulation (GDPR) requires customer data need to be protected, however how it will tackle CI-enabled automation in cybersecurity is not clear yet. The GDPR prohibits fully automated individual decision-making and profiling, which makes it a bit challenging technically. Under GDPR, additional security measures need to be developed to meet compliance requirements.

Many other challenges, such as dense heterogeneous deployment of networks, multiple radio access, and full-duplex transmission at the same time over 5G, etc., are still to be addressed.

B. Research Trends

The evolving computation intelligence techniques is still in its infancy and there are many unresolved research challenges as mentioned above. In this section, we address the above challenges and review the recent research trends and provide a categorization of CI-enabled cybersecurity for IoT.

1) Privacy-preserving data aggregation techniques, which focuses on ensuring data privacy in IoT environment during data aggregation in resource-constrained IoT devices, such as intelligent sensors, smart meter, etc. This scheme can provide authentic reports to the users, Message authentication code (MAC), Incremental hashing function (IHf), data slice, homomorphic encryption (He), Secret sharing, etc. are used to provide this.

2) In 5G-IoT environment, the dense moving of edge devices makes it very challenging to localise threats monitoring, detection, and protection for IoT nodes as well as offering powerful proximity-based authentication and identity management. This includes secure operations in heterogeneous environment and dynamic adaptation of security measures. Facing this issues, dynamically adjust its overall security level will be ready to respond adequately to any security compromises without creating disruptions hampering safe and uninterrupted system operations.

3) CI-enabled cyber behaviour profiling defense intelligence, CI enabled cyberthreat intelligence is one of emerging areas of focus in information security. Specifically, in cyber intrusion detection methods, attackers aim at obtaining better insight predictive power on the further behaviors, the CI techniques will play a key role in cyber behavioural modelling to use less training time and utilizes the benefits of ensemble learning to better model temporal relationships in data.

4) CI-enabled Cyber defences, aiming at defending IoT using CI techniques. The CI-enabled cyber defense can enable IoT detect vulnerabilities and perform response actions against cyberthreats and attacks. It can strengthen security strategies/tools make certain defensive aspects of cybersecurity more wide-reaching and effective.

5) New CI algorithms to adapt over time and make it easier to respond to cybersecurity threats. The learning algorithms can keep the tools updated with latest malware features and can enable it to detect the new generation of malware, cyberthreats, and cyberattacks. To investigate evolving cybersecurity techniques using CI and develop dynamic approaches is still research trends.

VI. CONCLUSION

With the development of IoT, CI emerges since it has the potential to significantly improve the dynamic cybersecurity features in the complex systems. In this paper, we survey the state-of-the-art of computational intelligence enabled cybersecurity challenges and research trends in the IoT. To achieve resilience and make IoT more secure, CI and cybersecurity based techniques should be considered in the design of IoT. Different from other survey papers, in this paper we highlight the main challenges that CI-enabled cybersecurity solutions are facing and new potential research trends.

REFERENCES


